Design and Impact of Data Caps

Wei Dai
School of Information and Computer Science
University of California, Irvine
Irvine, CA, US
daiw1@uci.edu

Scott Jordan
School of Information and Computer Science
University of California, Irvine
Irvine, CA, US
sjordan@uci.edu

Abstract— **We propose models to evaluate the impact of data caps upon subscribers and Internet Service Providers (ISPs). These models incorporate the time users devote to Internet applications and the opportunity cost of a user's free time. We consider a monopoly ISP that maximizes profit by setting tier prices, tier rates, network capacity, data caps, and overage charges. We illustrate which users are affected by data caps, and the resulting impact upon both users and ISPs. We show that an ISP will increase the tier rate and decrease the tier price when data caps are used to maximize profit. We give numerical results to illustrate the increase in ISP profit when data caps are used, and the corresponding changes in user and social welfare.**

Index Terms — **Charging, pricing, business models**

I. INTRODUCTION

In recent years, it has become common for wireless Internet Service Providers (ISPs) in the United States to place caps on the monthly usage of cellular data plans. Some wireline ISPs have also started placing caps on monthly usage of their broadband service offerings. The data caps often differ by the tier of the plan, and are often in the range from 50GB to 500GB per month [1]. The consequences of exceeding the cap differ by ISP; some charge an overage charge per unit volume over the cap, some reduce the throughput of violators, and some issue warnings and/or upgrade subscribers to a higher tier.

ISPs commonly claim that caps benefit most users. They cite statistics [2] that show that a small percentage of users consume a high percentage of network capacity, typically because these subscribers are heavy users of video streaming or file sharing [3]. The ISPs claim that flat-rate pricing, in which all subscribers to a tier pay the same amount independent of usage, is unfair to the majority of users [4]. They further claim that caps affect only a small percentage of heavy users [5], and that caps result in lower tier prices than would be offered without caps. Finally, ISPs claim that caps increase the incentive for ISPs to add capacity to the network, since the incremental capacity will benefit a broader set of users [6].

In contrast, many public interest groups claim that caps hurt most users. They claim that caps discourage the use of certain applications, including video streaming, and that this is often intended to protect the ISP's other services from competition [7]. They further claim that caps encourage a climate of scarcity, and that ISPs can increase their profit through the use of caps principally because of a lack of consumer choice in broadband providers [8]. Finally, public interest groups often claim that caps and their corresponding overage charges do not correspond to the cost for network capacity, and that the use of caps may decrease an ISP's incentive to add capacity [9].

There is a vigorous debate over the use of caps. Some public interest groups have called for government oversight [10]; in particular some have asked the US Federal Communications Commission (FCC) to investigate AT&T's broadband data caps [11]. A US Senate bill, the Data Cap Integrity Act, would require the FCC to evaluate data caps to determine whether they reasonably limit network congestion without unnecessarily restricting Internet use [12].

However, there is little academic literature that addresses the impact of data caps. Papak et al. [13] proposes an economic model of flat rate pricing as a form of bundling, where some bandwidth hogs exist. Their analysis shows for a monopoly service provider with negligible marginal costs that flat rate pricing almost always maximizes profit, even when there are some buyers with disproportionately large usage. While the authors suggest that the model might be used to explore the effects of data caps, the model presented does not incorporate data caps, nor does it model user willingness-to-pay or network congestion. Minne [14] explores ISP motivations for using data caps. It argues that heavy users are often profitable for ISPs, and that data caps may be a method for ISPs to price gouge and to protect an ISP's video business. However, no mathematical model of usage is proposed. Waterman et al. [15] expresses similar concerns that data caps may be anticompetitive behavior in the online television market. However, caps are not the focus of the paper and no model is proposed. Chetty et al. [16] focuses on the impact of data caps on subscribers. In a study of 12 households in South Africa, they find that uncertainties related to caps pose substantial challenges. However, again no mathematical model of usage is proposed.

In this paper, we propose models to evaluate the impact of data caps upon subscribers and ISPs. The model includes the critical elements of both Internet architecture and economic motivations. In section II, user utility is represented as a function of the time devoted per month to Internet applications, performance, and a user's relative utility for high bandwidth applications. User willingness-to-pay is expressed as utility minus the opportunity cost of the time devoted, which depends on the income of the user. We model ISP profit maximization by considering a monopolist that sets tier prices, tier rates, network capacity, data caps, and overage charges. In section III, we then examine which users are affected. We show how users fall into three categories: those unaffected by a cap, those who are capped but do not choose to exceed the cap, and those who exceed the cap and pay overage charges. In section IV, we analyze the impact of data caps on subscribers and ISPs. We examine a monopolist's use of caps, and compare the optimal tier rates, tier prices, and network capacity without caps to the

same quantities when caps are added. We show that an ISP will increase the tier rate and decrease the tier price when an ISP sets caps and overage charges to maximize its profit. Finally, in section V we give numerical results to illustrate how the tier rate, tier price, cap, and overage charges vary with the standard deviation in Internet usage amongst subscribers. We also illustrate the increase in ISP profit when caps are used, the corresponding change in user surplus, and the change in social welfare. The major contributions of this paper are:

- Novel user utility models: In contrast to previous literature that models user utility solely as a function of bandwidth, we propose a novel utility function that incorporates the time users devote to Internet applications, and the opportunity cost of a user's free time, thus differentiating light and heavy users on an economic basis.
- First model of ISP cap design: We present what we believe is the first model in the academic literature of how an ISP may set data caps and overage charges.
- First mathematical model of the impact of caps on subscribers: We characterize which users benefit from or are hurt by caps.

II. CAP MODEL FORMULATION

In this section, we introduce utility functions for subscriber usage, and we introduce ISP tier and cap models for profit maximization. We consider two interconnected problems separated by time scale. On a time scale of days, broadband Internet subscribers choose how much time to devote to Internet applications. On a time scale of months, subscribers choose what tier to subscribe to, and ISPs choose tier rates, tier prices, data caps, overage charges and network capacity.

A. Short term model

The dominant applications on North American fixed access broadband Internet access networks are video streaming, web browsing, and peer-to-peer file sharing, which together account for approximately 85% of download traffic volume [2]. Although there are different data caps in different tiers, the caps on lower tiers are generally high enough not to impact subscriber use of web browsing and email [1]. We thus focus on the impact of data caps on the use of video streaming (including video streaming done within peer-to-peer file sharing).

We posit that the perceived utility by user i for video streaming should be a function U_i of the time devoted to video streaming per month, the performance of video streaming, and a user's relative utility for video streaming. Denote t_i as the time (in seconds per month) that user i devotes to video streaming; normal economic assumptions are that a user's utility is an increasing concave function $V(t_i)$ of time devoted [17]. With respect to performance, video streaming is commonly classified as a semi-elastic application; we thus model a component of user utility by a sigmoid function $Q(x_i)$ of the throughput x_i (in bits per second) experienced by video streaming applications [18], normalized so that $Q(\infty)=1$. User i's utility for video streaming relative to other users is modeled using a scale factor v_i. The interaction between these three factors has not been studied; we model user i's utility for video streaming (in dollars per month) as the product:

$$U_i = v_i V(t_i) Q(x_i)$$

User i's willingness-to-pay for video streaming also depends on how the user values leisure time. The scale factor v_i should be increasing with this value. However, the time devoted to video streaming is also likely to be viewed as an opportunity cost. Denote p_i^t as the opportunity cost (in dollars per second) of user i's time, which is usually estimated to be between 20 to 50 percent of user i's income [19]. We model user i's willingness-to-pay for video steaming, (in dollars per month) as:

$$W_i = U_i - p_i^t t_i \qquad (1)$$

Most ISPs have started designing and marketing tiers on the basis of the applications they are intended for. We focus here on the decision by a user whether to subscribe to the tier designed for video streaming (hereafter referred to as the premium tier) or to a lower tier (hereafter referred to as the basic tier). Denote the price to upgrade from the basic tier to the premium tier by P. Denote an ISP's data cap (i.e. the maximum allowed number of bytes downloaded per month without incurring an overage charge) in the premium tier by C, and the price per byte charged for usage above the cap by p^o. Since the download volume consumed by applications other than video streaming is assumed to be much less than a typical cap [1] [2], user i's overage charge is $p^o \max(0, x_i t_i - C)$. User i is assumed to choose the time devoted to video streaming so as to maximize surplus, S_i, defined as the difference between willingness-to-pay and cost:

$$\max_{t_i} S_i = W_i - p^o \max\left(0, x_i t_i - C\right) - P \qquad (2)$$

If user i is not capped, the marginal utility from video streaming is equal to the user's valuation of time, p_i^t, or equivalently:

$$t_i = V'^{-1}\left(p_i^t / v_i Q(x_i)\right) < C/x_i, \quad \partial W_i / \partial t_i = 0 \qquad (3)$$

If user i is capped but not paying an overage charge, the marginal utility is larger than the user's valuation of time but smaller than the sum of this and the overage charge (per unit time), or equivalently:

$$t_i = C/x_i, \quad 0 < \partial W_i / \partial t_i < p^o x_i \qquad (4)$$

Finally, if user i is capped and paying an overage charge, the marginal willingness-to-pay is equal to the overage charge:

$$t_i = V'^{-1}\left(\left(p_i^t + p^o x_i\right) / v_i Q(x_i)\right) > C/x_i, \quad \partial W_i / \partial t_i = p^o x_i \quad (5)$$

We now turn to the relationship between traffic and performance. Denote the total downstream video streaming traffic (in bits per month) on the bottleneck link within the access network by $\lambda = \sum_i x_i t_i$.

As is common, we model the bottleneck link using an M/M/1/K queue to estimate the average delay d and loss l as a function of the traffic λ and the capacity μ. Most video streaming uses TCP or TCP-friendly protocols [20]. The throughput can be expressed as the minimum of the premium tier rate, denoted by X, and the limit that TCP places on the flow, denoted by $TCP(d,l)$ [21].

$$x_i = \min\left(X, TCP(d,l)\right)$$

B. Long term model

In the long term, say on a time scale of months, users seek to maximize their surplus by making the optimal Internet

subscription decision. User i will upgrade to the premium tier as long as her willingness-to-pay for video streaming is larger than the sum of the upgrade price P and overage charge $p^o \max(0, x_i t_i - C)$, i.e. if $S_i > 0$. Thus, user i's subscription choice can be expressed by:

$$T_i = \begin{cases} 1, & \text{if } W_i \geq P + p^o \max\left(0, x_i t_i - C\right) \\ 0, & \text{if } W_i < P + p^o \max\left(0, x_i t_i - C\right) \end{cases} \quad (6)$$

where T_i is an indicator of whether user i subscribes to the premium tier, and where t_i is determined by (3)-(5). The total number of the subscribers in the premium tier is:

$$N^p = \sum_i T_i$$

In the United States and many other countries, it is common that only one or two ISPs offer wireline broadband services [22]. In the remainder of the paper, we consider one ISP that monopolizes the market, which is a reasonable starting point given that there is no academic literature on how an ISP may set data caps and overage charges. A monopoly ISP is presumed to maximize its profit from video streaming by controlling the parameters P, X, C, p^o and μ:

$$\max_{P, X, C, p^o, \mu} PN^p + p^o O - K(\mu) \quad (7)$$

where the total amount of data above the cap is:

$$O = \sum_i \max\left(0, x_i t_i - C\right)$$

and $K(\mu)$ is the ISP's variable cost per month for capacity required to accommodate user demand.

It has been shown that a monopoly ISP can achieve near-optimal profit by setting network capacity so that the network load remains below a threshold ρ^{th} [1], i.e. $\mu = \lambda/\rho^{th}$, and so that video streaming throughput is constrained by the tier rate, i.e. $x_i = X$ [23]. This approach reduces the ISP's profit maximization problem in (7) to:

$$\max_{P, X, C, p^o} PN^p + p^o \sum_i \max\left(0, X t_i - C\right) - K\left(\frac{p^\mu}{\rho^{th}} \sum_i X t_i\right) \quad (8)$$

III. AFFECTED USERS

User i may or may not be capped depending on the value placed on video streaming, v_i, and the value placed on time, p_i^t. We partition Internet subscribers into four groups: users not subscribing to the premium tier, users in the premium tier who are not capped, users in the premium tier who are capped but not paying an overage charge, and users in the premium tier who are capped and paying overage charges.

We first focus on users who are indifferent between the basic and premium tiers, henceforth referred to as *marginal users*. According to (1) (2) and (6), if user i is a marginal user, the value placed on video streaming, v_i, and the value placed on time, p_i^t, should satisfy:

$$v_i V\left(t_i\right) Q\left(X\right) - p_i^t t_i = P + p^o \max\left(0, X t_i - C\right) \quad (9)$$

Marginal users who are capped satisfy (3) and (9). Denote $v^{1,u}\left(C, P, X, p^o, p_i^t\right)$ as the solution to the fixed point equation in v_i resulting from (3) and (9). There is a unique solution for v_i because by (1) and (3) $dW_i/dv_i = V(t_i)Q(X) > 0$ for $v_i > 0$, which

---

[1] A commonly discussed choice for ρ^{th} is 0.7.

---

makes the left side of (9) an increasing function of v_i. Similarly, marginal users who are capped but not paying an overage charge satisfy (4) and (9); denote $v^{1,c}\left(C, P, X, p^o, p_i^t\right)$ as the solution to the fixed point from these equations. Finally, marginal users who are capped and paying overage charges satisfy (5) and (9); denote $v^{1,o}\left(C, P, X, p^o, p_i^t\right)$ as the solution to the fixed point from these equations.

Thus the marginal users lie on the curve:

$$v_i \; \square \; v^1\left(C, P, X, p^o, p_i^t\right) = \begin{cases} v^{1,u}\left(C, P, X, p^o, p_i^t\right), & \text{if } t_i < C/x_i \\ v^{1,c}\left(C, P, X, p^o, p_i^t\right), & \text{if } t_i = C/x_i \\ v^{1,o}\left(C, P, X, p^o, p_i^t\right), & \text{if } t_i > C/x_i \end{cases}$$

We can use this curve to partition all Internet subscribers. Users who do not subscribe to the premium tier, denoted by G_b, place a smaller value on video streaming than do marginal users:

$$G_b = \left\{ i : v_i < v^1\left(C, P, X, p_i^t\right) \right\}$$

Users in the premium tier who are not capped satisfy (3), i.e.

$$v_i < \frac{p_i^t}{V'\left(C/X\right)Q\left(X\right)} \; \square \; v^2\left(C, X, p_i^t\right)$$

Denote the set of such users:

$$G_u = \left\{ i : v^1\left(C, P, X, p^o, p_i^t\right) < v_i < v^2\left(C, X, p_i^t\right) \right\}$$

Users in the premium tier who are capped but not paying an overage charge satisfy (4), i.e.

$$v^2\left(C, X, p_i^t\right) \leq v_i \leq \frac{p_i^t + p^o X}{V'\left(C/X\right)Q\left(X\right)} \; \square \; v^3\left(C, X, p^o, p_i^t\right)$$

Denote the set of such users:

$$G_c = \left\{ i : \max\left(v^1\left(C, P, X, p^o, p_i^t\right), v^2\left(C, X, p_i^t\right)\right) < v_i < v^3\left(C, X, p^o, p_i^t\right) \right\}$$

Finally, the set of the users in the premium tier who are capped and paying overage charges satisfy (5), i.e.

$$v_i > v^3\left(C, X, p^o, p_i^t\right)$$

Denote the set of such users:

$$G_o = \left\{ i : v_i > \max\left(v^1\left(C, X, p^o, p_i^t\right), v^3\left(C, X, p^o, p_i^t\right)\right) \right\}$$

These four sets define a partition of Internet subscribers on the basis of (v_i, p_i^t). However, it is more revealing to use $(v_i/p_i^t, p_i^t)$ as the basis, as v_i/p_i^t determines the amount of time that user i devotes to video streaming absent a data cap. As illustrated in Fig. 1, the functions v^1, v^2 and v^3 form the boundaries of the four sets. Users with a small relative value on streaming, v_i/p_i^t, and/or a small income (and hence a small p_i^t) do not subscribe to the premium tier. Users with a small relative value on streaming, v_i/p_i^t, but a larger income (and hence a larger p_i^t) subscribe to the premium tier but are not capped due to their low interest in streaming. Users with a moderate relative value on streaming, v_i/p_i^t, and moderate or high incomes subscribe to the premium tier and are capped. Users with a high relative value on streaming, v_i/p_i^t, and/or high incomes subscribe to the premium tier and are willing to pay overage charges.
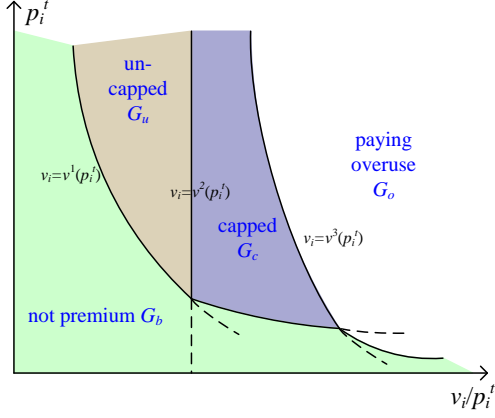
3

Figure 1. Partition of Internet subscribers based on $(v_i/p_i^t, p_i^t)$.

IV. IMPACT OF CAP

In this section, we analyze the impact of data caps on subscribers and ISPs. We wish to compare the optimal tier rates, tier prices, and network capacity without caps to the same quantities when caps are added.

A. Optimal pricing plan without data caps

We start by characterizing the rate, price and network capacity, without data caps, that maximize an ISP's profit. In a given market, denote the joint distribution of users' value placed on video steaming and users' value placed on time by $f_{v,p}(v, p^t)$. The ISP profit maximization problem in (8), without data caps (i.e., $p^o = 0$ or $C = \infty$), becomes:

$$\max_{P,X} Profit = PN \int_{p^t=0}^{\infty} \int_{v=v^1(\infty,P,X,0,p^t)}^{\infty} f_{v,p}(v, p^t) dv dp^t - K\left(\frac{N}{\rho^{th}} \int_{p^t=0}^{\infty} \int_{v=v^1((\infty,P,X,0,p^t))}^{\infty} Xt f_{v,p}(v, p^t) dv dp^t\right) \quad (10)$$

where $t = V'^{-1}\left(p^t/vQ(X)\right)$ and N is the total number of users.

Theorem 1: In the solution to the profit maximization problem (10), the optimal price P and rate X satisfy the following first order conditions:

$$\frac{\partial Profit}{\partial P} = N \int_{p^t=0}^{\infty} \int_{v=v^1(\infty,P,X,0,p^t)}^{\infty} f_{v,p}(v, p^t) dv dp^t$$
$$- N \int_{p^t=0}^{\infty} \left(P - \frac{p^\mu X t^1}{\rho^{th}}\right) \frac{f_{v,p}(v^1, p^t)}{Q(X)V(t^1)} dp^t = 0$$

$$\frac{\partial Profit}{\partial X} = N \int_{p^t=0}^{\infty} \left(P - \frac{p^\mu X t^1}{\rho^{th}}\right) f_{v,p}(v^1, p^t) \frac{v^1 Q'(X)}{Q(X)} dp^t -$$
$$N \frac{p^\mu}{\rho^{th}} \int_{p^t=0}^{\infty} \int_{v=v^1(\infty,P,X,0,p^t)}^{\infty} \left(t + X \frac{\partial t}{\partial X}\right) f_{v,p}(v, p^t) dv dp^t = 0$$

where t^1 is the amount of data consumed by a marginal user, i.e. $v_i = v^1(\infty, P, X, 0, p^t)$; $p^\mu = dK(\mu)/d\mu$ denotes the marginal network cost.

B. Data caps that ensure heavy users pay for their usage

We now turn to the effect of adding a data cap. We do so in two steps. First, we consider the case in which an ISP institutes caps merely in order to ensure that heavy users pay an amount

equal to the cost of their usage. This case is interesting in its own right, as some ISPs claim this is the purpose of their data caps [6]. In the next subsection, we consider the case in which an ISP uses caps to maximize profit.

Suppose that an ISP imputes a cost to user i equal to $p^\mu t_i X/\rho^{th}$, on the basis that user i's usage is $t_i X$, and that this requires incremental capacity $t_i X/\rho^{th}$ at an incremental cost per unit capacity p^μ. Then given the optimal price P and rate X as calculated in (10), we presume in this subsection that the goal of the ISP is to set a data cap C and overage charge p^o so that:

$$P - p^\mu t_i X \big/ \rho^{th} \geq 0 \text{ for all } i$$

We examine a simple method of achieving this goal: $C = P/p^\mu$ and $p^o = p^\mu/\rho^{th}$, which we henceforth refer to as the *heavy users cap*. Under this policy, premium tier subscribers with usage greater than or equal to $C = t_i X$ will pay an amount $P + p^o(t_i X - C)$, which exactly equals their imputed cost, i.e. $p^\mu t_i X/\rho^{th}$.

The impact of such a cap on all users is illustrated in Fig. 2. Under the pricing plan without data caps from (10), users above and to the right of the black curve subscribe to the premium tier. Under the *heavy users cap*, users in the grey, blue and white regions subscribe to the premium tier. The red region corresponds to users who drop the premium tier when data caps are added, because they have high valuations on video streaming but low incomes. Subscribers with moderate valuations on video streaming and high incomes (the grey region) are unaffected by the cap. Users with high valuations on video streaming and high incomes (the blue and white regions) have lower surplus after a heavy user cap is added. The ISP only earns a positive profit from un-capped users.
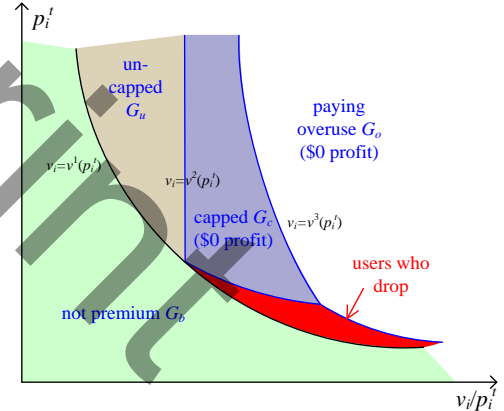

Figure 2. The impact of a cap designed to ensure that heavy users pay an amount equal to the cost of their usage.

C. Data caps to maximize ISP profit

We now consider the case in which an ISP sets caps and overage charges to maximize its profit. The pricing plan derived in the previous subsection does not maximize profit, since the cap and overage charge were only intended to ensure that heavy users pay for their usage.

Using the joint distribution $f_{v,p}(v, p^t)$, the revenue and associated capacity can be expressed as a sum over sets of users:

$$Rev = N \int_{(v,p^t)\in G_u \cup G_c} P f_{v,p}(v, p^t) dv dp^t$$
$$+ N \int_{(v,p^t)\in G_o} \left(P + p^o(Xt - C)\right) f_{v,p}(v, p^t) dv dp^t$$

$$\mu = \frac{N}{\rho^{th}} \int_{(v,p^t) \in G_u \cup G_o} Xtf_{v,p}\left(v, p^t\right) dv dp^t$$

$$+ \frac{N}{\rho^{th}} \int_{(v,p^t) \in G_c} Cf_{v,p}\left(v, p^t\right) dv dp^t$$

and the ISP profit (before fixed costs are subtracted) in (8) can be expressed as:

$$Profit = Rev - K\left(\mu\right) \qquad (11)$$

Unfortunately, a closed form characterization of the optimal tier rate, tier price, cap, and overage charge (henceforth referred to as the *profit-maximizing cap*) is difficult to obtain. We can, however, compare the tier rate and price to those without data caps, and we can compare the cap and overage charge to those in the *heavy users cap*:

Theorem 2: Set P and X using (10), $C = P/p^\mu$, and $p^o = p^\mu/\rho^{th}$. Then:

$$\frac{\partial Profit}{\partial P} \leq 0, \frac{\partial Profit}{\partial X} \geq 0, \frac{\partial Profit}{\partial C} \leq 0, \frac{\partial Profit}{\partial p^o} \geq 0$$

Proof: Omitted due to the space constraint.

Based on the results in theorem 2, we can summarize how ISPs might change the parameters in the pricing plan starting from the *heavy users cap* plan presented in the previous subsection: the ISP has the incentive to reduce price P, increase rate X, increase overage charge p^o above p^μ/ρ^{th}, and reduce cap C from P/p^μ. Thus an ISP that uses caps to maximize profit will have smaller caps and higher overage charges than one that uses caps only to ensure that heavy users pay for their usage.

In this case, all subscribers will benefit from the reduced price P. Capped subscribers will be hurt by the cap and the overage charge. A user who is not capped will benefit from the increased tier rate. However, the effect of the increased rate X upon capped users is less clear because an increase in tier rate also decreases the amount of video streaming a user can do before running into the cap. If user i is capped but does not pay an overage charge, the derivative of user surplus with respect to rate X can be expressed as:

$$\frac{\partial S_i}{\partial X} = v_i Q'\left(X\right) V\left(\frac{C}{X}\right) - \frac{C}{X^2}\left(v_i Q\left(X\right) V'\left(\frac{C}{X}\right) - p_i^t\right)$$

So, $\partial S_i/\partial X > 0 \Leftrightarrow p_i^t > v_i Q(X) V'(C/X) - v_i Q'(X) V(C/X) X^2/C$

Similarly, if subscriber i is capped and paying an overage charge, $\partial S_i/\partial X > 0 \Leftrightarrow v_i Q'\left(X\right) V\left(t_i\right) > p^o t_i$. In both cases, some users will benefit from the increased tier rate and some will be hurt, according to a user's valuation of video streaming and valuation of his time.

We would like to understand the total impact of all these changes upon each user. As above, the impact depends on a user's (v_i, p_i^t). The impact of such a cap on all users is illustrated in Fig. 3. Under the *profit-maximizing cap*, there is again a set of users (shaded in dark brown) who would have subscribed to the premium tier had caps not been present, because they have high valuations on video streaming but low incomes. However, now there is also a set of users (shaded in dark grey) with low valuations on video streaming and high incomes that would not subscribe to the premium tier without data caps but will subscribe with data caps because of the reduction in price and increase in tier rate. Subscribers with

moderate valuations on video streaming and moderate to high incomes (shaded in dark grey and light grey) benefit from a data cap; these users value their time a lot, but do not spend a lot on video streaming and consequently also benefit from the reduction in price and increase in tier rate. In contrast, subscribers with moderate valuations on video streaming and low to moderate incomes (shaded in dark brown and light brown) are hurt by the optimal data cap; the effect of the cap and overage charges outweigh the reduction in the tier price.
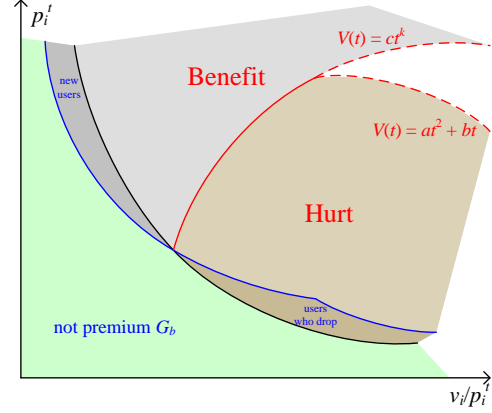

Figure 3. Comparing video streaming surplus utility in the presence and absence of profit-maximizing cap under $V(t) = at^2 + bt$ and $V(t) = ct^k$.

Finally, there is a set of heavy users with high valuations on video streaming and moderate to high incomes. These users may either benefit or be hurt by caps, depending on the shape of function $V(t)$. We examine two special cases here. In the case of a linear demand curve [24] ($V(t) = at^2 + bt$, $0 < t < -b/2a = t^{max}$, $a < 0$, $b > 0$), these heavy users benefit from the data cap, because the benefit from video streaming outweighs the overage charge. In contrast in the case of a constant elasticity demand curve [24] ($V(t) = ct^k$, $0 < t < t^{max}$, $0 < k < 1$), these heavy users are hurt by the data cap, because the overage charge outweighs the benefit from video streaming.

V. NUMERICAL RESULTS

In this section, we evaluate the impact of the data cap on the pricing plan, users, the ISP and social welfare. We set the following parameters: $(v/p^t, p^t)$ follows a multivariate lognormal distribution with parameters set to match demand and income statistics in [19] [25] and [26]; $V(t) = ct^k$ with parameters set to match the video streaming statistics in [25]; performance $Q(x)$ given in [18]; $N = 20000$ users; load threshold $\rho^{th} = 0.7$ [23]; $p^\mu = \$10$/Mbps/month [27].

In Fig. 4, we plot the ISP's premium tier rate X and upgrade price P with and without data caps. The price depends strongly on the distribution of the scale factor v_i for users' utility for video streaming, and hence it is plotted as a function of the shape parameter (denoted σ) of the lognormal distribution. Both the mean and variance of v_i increase with σ, while the median is fixed, reflecting a higher proportion of heavy users. Without a data cap, the price increases rapidly with the proportion of heavy users, due to both the higher willingness-to-pay of heavy users and the much higher usage (and therefore cost) of heavy users. The tier rate decreases slightly with the proportion of heavy users, in an attempt to maintain acceptable performance and cost. When profit-maximizing data caps are

used, the tier rate increases moderately and the price decreases substantially, as predicted by the analytical results above. The prices both with and without caps would fall if there are competing ISPs.
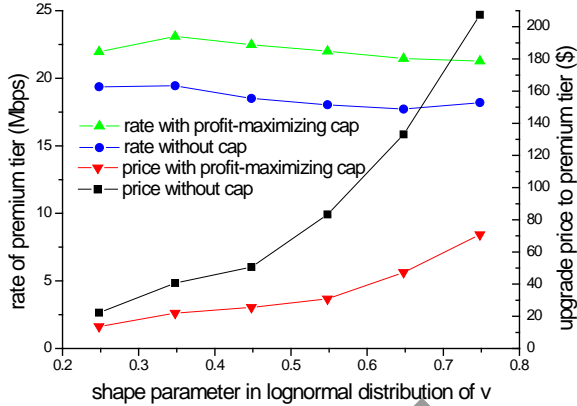

Figure 4. Tier rate and price versus proportion of heavy users.

The decrease in tier price when an ISP uses data caps does not necessarily decrease the total price paid by the subscriber, since there are also overage charges. In Fig. 5, we illustrate the caps and overage charges under the profit-maximizing cap and the less aggressive heavy users cap discussed above. Under the heavy user cap policy, the caps are quite high and the overage charges are quite low, since the cap and overage charges are only intended to recover the imputed cost from heavy users. In contrast, under the profit-maximizing cap policy, the caps are quite low and the overage charges quite high. Current wireline ISP caps are roughly in this range, e.g. AT&T offers pricing plans with a cap of 150 or 250 GB per month and an overage charge of $10 for an additional 50GB [1].
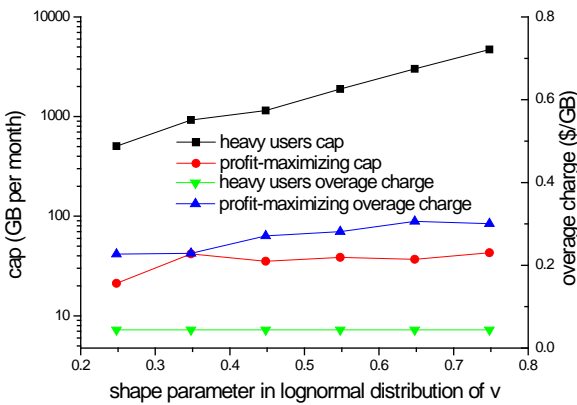

Figure 5. Data caps and overage charges versus proportion of heavy users.

To understand the impact upon different types of users, in Fig. 6, we investigate which users are better off with and without data caps as a function of (v_i / p_i^t, p_i^t) when σ=0.548. Users in the green region are indifferent, since they do not subscribe to the premium tier in either case. Users in the blue region have a larger surplus when profit-maximizing data caps are used, since the benefit of decreased prices and increased

rates outweigh the impact of the caps. Users in the brown region have a smaller surplus when profit-maximizing data caps are used, since the impact of the caps outweighs the benefit of decreased prices and increased rates.
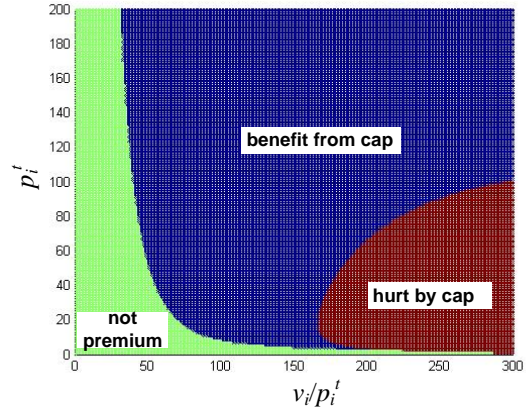

Figure 6. User surplus utility under profit-maximizing caps.

In Fig. 7, we give the ISP profit, user welfare (defined as $\sum_i S_i$), and social welfare (defined as user welfare plus ISP profit) resulting from each plan. Without data caps, ISP profit increases as the proportion of heavy users increases, due to increases in subscriptions to the premium tier. When profit-maximizing data caps are used, ISP profit further increases substantially. The increase reflects the new overage charges, minus some reductions due to lower tier prices and some changes in premium tier subscriptions. The revenue and profit are both *more* sensitive to the proportion of heavy users when data caps are used, because of the strong correlation between the proportion of heavy users and revenue raised through overage charges.
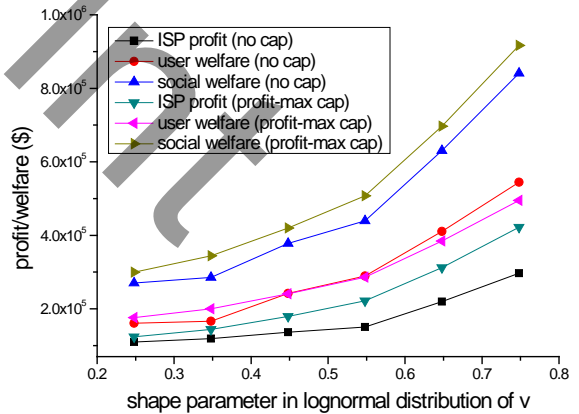

Figure 7. Welfare versus proportion of heavy users.

User welfare increases with the proportion of heavy users, since heavy users have high surpluses. When profit-maximizing data caps are present, user welfare decreases when there is a high proportion of heavy users. Social welfare might decease or increase when data caps are used, depending on parameters. In this plot, it increases slightly. Social welfare also depends on the distribution of wealth in the society. In Fig. 8, we plot it versus the shape parameter of the lognormal

distribution for a user's value on time, which is also proportional to income. Social welfare from video streaming increases with wealth inequality, primarily since the mean income is increasing. Of greater interest is that use of data caps is observed here to increase slightly when inequality is low but to decrease slightly when inequality is high. We warn, however, that social welfare also depends on the shape of the utility function, and that different utility functions may result in different conclusions about changes in social welfare.
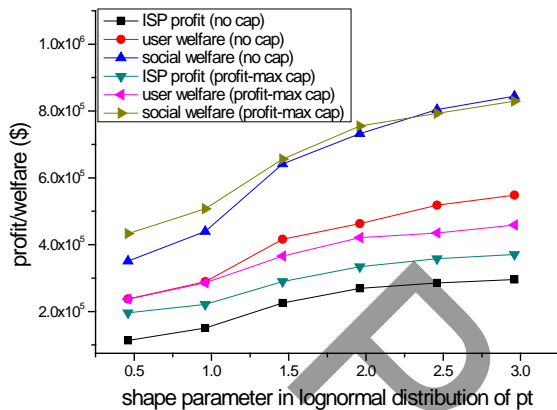

Figure 8. Welfare versus the proportion of wealthy users.

VI. CONCLUSION

This paper proposed novel user utility models that incorporate the time users devote to Internet applications, and the opportunity cost of a user's leisure time, thus differentiating light and heavy users on an economic basis. It considers a monopoly ISP that maximizes profit by setting tier prices, tier rates, network capacity, data caps, and overage charges; thereby presenting what we believe is the first model in the academic literature of how an ISP may set data caps and overage charges. We show that users with a small relative value on streaming and a large income subscribe to the premium tier but are not capped due to their low interest in streaming, users with a moderate relative value on streaming and moderate or high incomes subscribe to the premium tier and are capped, and that users with a high relative value on streaming and/or high incomes subscribe to the premium tier and are willing to pay overage charges.

Analytical and numerical results show that the ISP will increase the tier rate and decrease the tier price when data caps are used to maximize profit. The ISP will also set smaller caps and higher overage charges than when caps are used only to ensure that heavy users pay for their usage. As a result, light users benefit from data caps because of the increased tier rate and reduced tier price, while heavy users are hurt by the caps and overage charges. User welfare and social welfare may increase or decrease depending on the shape of the market density function and user utility function.

REFERENCES

[1] Stacey Higginbotham, "Which ISPs are capping your broadband, and why?" *Article posted at Gigaom.com*, Oct. 2012, available at http://gigaom.com/2012/10/01/data-caps-chart/

[2] Sandvine Incorporated, "Global Internet phenomena report 2012," *Technical report*, Fall 2012.

[3] ACLP, Advanced Communications Law&Policy Institute at New York Law School, "A Primer on Data Consumption: Trends& Emerging Business Models," *Released Technical Report*, Feb. 2012.

[4] George S. Ford, "A Most Egregious Act? The Impact on Consumers of Usage-Based Pricing," *Phoenix Center Report*, May. 2012.

[5] Richard Bennett, "Comcast Raises Invisible Data Cap," *The Information Technology & Innovation Foundation Report*, May. 2012.

[6] Roger Yu, "Cable companies cap data use for revenue," *news at USA Today*, available at http://www.usatoday.com/story/tech/2012/10/01/internet-data-cap/1595683/

[7] Joel Kelsey, "Comcast Should Eliminate Punitive Data Caps Altogether," *Free Press Report*, May. 2012.

[8] Hibah Hussain, Danielle Kehl, Patrick Lucey, "The Destructive Power of Data Caps," *Free Press Report*, Dec. 2012.

[9] Hibah Hussain, Danielle Kehl, Benjamin Lennett, Patrick Lucey, "Capping the Nation's Broadband Future?" *New America Foundation's Open Technology Institute Report*, Dec. 2012.

[10] Andrew Odlyzko, Bill St. Arnaud, Erik Stallman and Michael Weinberg, "Know Your Limits Considering the Role of Data Caps and Usage Based Billing in Internet Access Service," *Public Technical Report*, May 2012.

[11] New America, "Letter to FCC on AT&A Data Caps," May 2011.

[12] Ron Wyden, "Data Cap Integrity Act," *US Senate Bill*, Dec. 2012.

[13] Papak Nabipay, Andrew Odlyzko, Zhi-Li Zhang, "Flat Versus Metered Rates, Bundling, and Bandwidth Hogs," *Proceedings of NetEcon 11: 6th Workshop on the Economics of Networks, Systems, and Computation*, 2011.

[14] Minne, Jacob Joseph Orion, "Data Caps: How ISPs are Stunting the Growth of Online Video Distributors and What Regulators Can Do About it," *Public Technical Report*, May, 2012. Available at http://dx.doi.org/10.2139/ssrn.2049174

[15] David Waterman, Ryland Sherman, Sung Wook Ji, "The Economics of Online Television: Revenue Models, Aggregation, and TV Everywhere," *Telecommunications Policy Research Conference TPRC*, Sep. 2012.

[16] Marshini Chetty, Richard Banks, A.J. Bernheim Brush1, Jonathan Donner and Rebecca E. Grinter, "You're Capped! Understanding the Effects of Bandwidth Caps on Broadband Use in the Home," *ACM Conference on Human Factors in Computing Systems*, May 2012.

[17] Hans U Gerber, Gerard Pafumi, "Utility Functions: From Risk Theory to Finance," *North American Actuarial Journal*, Jul. 1998.

[18] S. Weber and V. Veeraraghavan, "Distributed algorithms for rate-adaptive media streams," *INFORMS Telecommunications Conference*, Dec. 2007.

[19] Frank J. Cesario, "Value of Time in Recreation Benefit Studies," *Land Econmics*, vol. 52, no. 1, pp. 32-41, Feb. 1976.

[20] A. Begen, T. Akgul, and M. Baugher. "Watching video over the web, part 1: Streaming protocols," *IEEE Internet Computing*, vol.15, no. 2, 2011.

[21] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation," *ACM SIGCOMM*, 1998.

[22] Nicholas Economides, "Net Neutrality, Non-Discrimination and Digital Distribution of Content Through the Internet," *Journal of Law and Policy for the Information Society*, vol. 4, no. 2, 2008.

[23] Wei Dai, Scott Jordan, "Modeling ISP Tier Design," *Technical Report*, Feb. 2013.

[24] Besanko and Braeutigam, "Microeconomics," *Textbook*, 2005.

[25] Burstmedia.com, "Online video content & advertising video preferences, habits and actions in Q4 2011," *Technical report*, Oct. 2011, available at http://www.burstmedia.com/pdf/burst_media_online_insights_2011_11.pdf.

[26] United States Census Bureau, "Money income of households, 2009," 2009, *Public report*, available at http://www.census.gov/compendia/statab/2012/tables/12s0690.pdf

[27] CCS Leeds Network Solutions, "UK leased lines pricing," http://www.uk-leased-line.co.uk/prices.html, accessed Apr.30 2012.