

---

# THE STRATEGIC USE OF DOWNLOAD LIMITS BY A MONOPOLY PLATFORM\*

NICHOLAS ECONOMIDES<sup>†</sup> AND BENJAMIN E. HERMALIN<sup>‡</sup>

---

## ABSTRACT

We consider a heretofore unexplored explanation for why platforms, such as Internet service providers and mobile-phone networks, offer plans with download limits: through one of two mechanisms, doing so causes the providers of the content consumer purchase to either reduce their prices or increase their quality. This, in turn, generates greater surplus for consumers. A platform captures this increased surplus by charging consumers higher access fees. Even accounting for congestion externalities, we show that a platform will tend to restrict downloads more than would be welfare-maximizing; indeed, in some instances, by so much that a complete prohibition on such behavior is welfare superior to the practices the platform would pursue. Somewhat paradoxically, we show that a platform will install more bandwidth when allowed to restrict downloads than when prevented from doing so. Other related phenomena are explored.

*Keywords:* two-sided markets, Internet, download limits (caps), congested platforms, network neutrality, price discrimination.

*JEL Classification:* L1, D4, L12, L13, C63, D42, D43.

---

\*The authors thank Joshua Gans, Jim Hosek, Paul Klemperer, Francine Lafontaine, Chengsi Wang, two anonymous referees, and seminar participants at Berkeley-Stanford IOFest, Nuffield College, the University of Mannheim, the Paris School of Economics, and CSEF University of Naples for helpful suggestions. The financial support of the Newhouse Foundation and of the Thomas and Alison Schneider Distinguished Professorship in Finance, as well as the hospitality of Nuffield College, University of Oxford, are gratefully acknowledged.

<sup>†</sup>Stern School of Business • New York University • 44 West Fourth Street • New York, NY 10012-1126 • email: [neconomi@stern.nyu.edu](mailto:neconomi@stern.nyu.edu) • Also NET Institute, [www.NETinst.org](http://www.NETinst.org) and Haas School of Business, University of California, Berkeley.

<sup>‡</sup>University of California • Department of Economics • 530 Evans Hall #3880 • Berkeley, CA 94720-3880 • email: [hermalin@berkeley.edu](mailto:hermalin@berkeley.edu).

CONTENTS

1	INTRODUCTION	1
2	BASELINE MODEL	5
3	PRELIMINARY ANALYSIS	7
4	BENCHMARK: NO CONGESTION	9
5	CONGESTION EXTERNALITIES	14
6	ENDOGENOUS BANDWIDTH	19
7	ADVERTISING-SUPPORTED CONTENT AND SURPLUS EXTRACTION VIA HIGHER QUALITY	22
8	HETEROGENOUS CONSUMERS, CONTENT, AND PRICE DISCRIMINATION	27
9	DISCUSSION	33
	APPENDIX A: PROOFS NOT GIVEN IN TEXT	36
	REFERENCES	39

## 1 INTRODUCTION

Why do platforms, such as residential Internet service providers (ISPs) and mobile telephone companies, impose caps (limits) on how much content their customers can download each month? An immediate answer is that these caps may be part of a second-degree price discrimination scheme via quantity discounts. Another is that caps represent the platform's efforts at alleviating the congestion externality consumers impose on each other: reducing the externality raises consumer welfare, which the platform captures as higher access (hookup) fees. Although both answers are likely part of the story, in this paper we identify an additional motive for caps: they can, through one of two mechanisms, cause content providers to lower their prices, raising consumer surplus and, thus, increasing what the platforms can capture.<sup>1</sup> In essence, unlike a traditional price-discrimination analysis, in which a monopolist uses quantity-based pricing to appropriate surplus from end users, we show that a platform has an incentive to use similar practices to capture surplus from upstream providers.

One of the two mechanisms is the following: if the caps are binding, then the more binding, the more consumers will perceive the digital products they acquire from different content providers as substitutes. This, in turn, will increase the competitive pressures on the content providers, who will respond by lowering their prices. Lower prices mean greater consumer surplus for customers, which the platform can capture via higher access fees.

The basic idea can be readily illustrated. Suppose there is a measure one of end consumers and two content providers. A consumer's utility is

$$u = -H + \sum_{n=1}^2 (v - p_n)\chi_n,$$

where  $\chi_n$  indicates whether a consumer purchases a unit of the  $n$ th content provider's product ( $\chi_n = 1$ ) or not ( $\chi_n = 0$ ),  $p_n$  is the  $n$ th content provider's price, and  $H$  is the hookup fee charged by the platform. Assume the content providers each have a constant marginal cost of 0 and no overhead or fixed costs. Assume the platform has no costs. If we assume the content providers set prices first, the platform sets  $H$  next, with consumers making their purchase decisions last, then it is readily seen that, in equilibrium,  $p_1 = p_2 = v$  and  $H = 0$ .<sup>2</sup>

Now consider an alternative game: for some reason (*e.g.*, a download restriction), each consumer is limited to one unit in total; that is, a consumer can buy

---

<sup>1</sup>Although there is much free content on the Internet, there are many content providers that directly charge consumers (*e.g.*, Netflix, iTunes, Pandora One, Amazon ebooks, many periodicals, and many gaming sites). As we discuss in Section 7, the content providers could also respond by improving the quality of their products, again raising consumer surplus and, thus, the fees the platform can charge. This effect would be of particular relevance when content is provided free to customers and the content providers generate revenue by selling advertising.

<sup>2</sup>An arguably more realistic timing, which is equivalent in terms of the resulting equilibrium, is (i) the platform sets  $H$ ; (ii) consumers decide to connect (subscribe to the network) or not; (iii) the content providers set prices; and (iv) consumers decide what content to buy. There are some nuances, though, to this alternative timing, which we discuss later.

from one content provider or the other, but not both. Otherwise the setting is as just described. The content providers are now effectively in Bertrand competition and the resulting equilibrium exhibits  $p_1 = p_2 = 0$ . Consumer surplus *gross* of the hookup fee is  $v$ , which means the platform can charge consumers a hookup fee of  $v$ .

Comparing the two games, the download restriction harms the content providers, which see their profits go from  $v$  each to zero; benefits the platform, which sees its profit go from 0 to  $v$ ; and reduces welfare (the sum of consumer surplus and profits) from  $2v$  to  $v$ . Because, in either scenario, the consumers' surplus is fully extracted, they are neither better nor worse off.

The second mechanism requires a bit of background. Although, in reality, some platforms' download limits are inviolate, many are loose: they can be exceeded, but doing so subjects the customer to an overage charge (effectively) linear in the amount by which the limit is exceeded.<sup>3</sup> Obviously if the overage rate is high enough, the limit is effectively inviolate; but if the rate is modest, then the "limit" would appear simply to be a kink point in a multi-part tariff.<sup>4</sup> We show that, with homogeneous consumers, any multi-part tariff is equivalent to either a simple two-part tariff or a *binding*-limit (cap) regime in which the platform charges a hookup fee only—see Lemma 1.<sup>5</sup>

Absent a congestion externality and assuming the platform's marginal cost is zero (a reasonable approximation in our context), one would expect, based on standard theory, that a platform's profit-maximizing two-part tariff would have a positive hookup (fixed) fee but a *zero* per-unit fee, given that we assume consumers are homogenous. We find, however, that a platform does better with a positive per-unit fee: the greater that fee, the lower are the prices the content providers charge. This second mechanism, though, has different mechanics than the first: whereas a binding cap causes prices to fall because it turns content providers into (fiercer) competitors, a positive per-unit fee acts like an excise tax on trade between consumers and content providers. Although the statutory incidence of that "tax" falls on the consumers, the actual incidence ends up,

---

<sup>3</sup>According to Higginbotham (2012), at least three American ISPs had firm limits in 2012 (*i.e.*, households that hit the limit could download nothing more). As of this writing, the Canadian ISP Rogers charges \$1 per gigabyte of overage, rounding the overage to the closest whole gigabyte (so, *e.g.*, a 1.49GB overage costs \$1 and a 1.51GB overage costs \$2) and the American mobile-phone company Verizon charges \$15 per gigabyte of overage, to give two examples of platforms with loose limits.

<sup>4</sup>To the best of our knowledge platforms such as ISPs and mobile-phone networks do not charge a positive per-unit fee on units below the cap. In our treatment below, however, we allow for the possibility that they might wish to do so; that is, have one per-unit fee below the cap and a different one above.

<sup>5</sup>In light of Lemma 1, one might ask why platforms bother to state caps when, effectively, utilizing a two-part tariff. We discuss this in detail in Section 9, but, foreshadowing that discussion as well as other results, we note now there are many reasons why platform pricing could have evolved to be quoted in terms of download caps and overage charges. That said, there are some platforms that utilize more straightforward two-part tariffs: the UK mobile-telephone provider Three, for example, offers data plans under which consumers pay 1p per megabyte on every megabyte.

as is well known, divided between the two. Consequently, the sum of consumer surplus and “tax revenue” can exceed consumer surplus absent the tax. Because the platform absorbs consumer surplus via the hookup fee, it is better off setting a positive per-unit fee (a fee above marginal cost) even though its consumers are homogenous.

The reader’s immediate response could be, “okay, these are theoretical possibilities, but is there any real-world evidence?” There is indeed some. For example, Reed Hastings, Netflix’s CEO, arguing against the Canadian ISPs Rogers and Bell Canada’s download caps, said, “It’s an effective way to drive the bill up, that tends to be why caps are used.”<sup>6</sup> Further, some reports suggest that ISPs are interested in imposing caps to enhance revenue because they are blocked from directly charging content providers under network neutrality.<sup>7</sup> Keeping in mind that the statutory incidence of an excise tax is irrelevant to its actual incidence (at least ignoring transaction costs and psychological factors), a platform should be indifferent between directly charging content providers a per-unit fee and charging consumers that fee; that is, as suggested in the popular press, if blocked from directly “taxing” content providers, platforms would begin to “tax” consumers. Moreover, evidence indicates ISPs have sought to impose caps even when congestion is not a significant problem.<sup>8</sup>

In the simple example above, the only reason for the platform to impose a download cap is rent extraction. As hinted, an additional rationale for restrictions arises if there is a congestion externality: by limiting total consumption, the platform could enhance welfare by reducing the congestion externality. Much of the analysis that follows—see, in particular, Sections 4–6—focuses on that issue. We find, in a static setting (fixed bandwidth), that, while it is possible that welfare is greater if the platform is unrestrained in its setting of download caps or positive per-unit fees than if prohibited from using such measures, there are many circumstances in which completely prohibiting such measures is welfare superior to allowing the platform to act freely.

In a dynamic setting (endogenous bandwidth), we find—somewhat paradoxically—that a platform’s incentive to build bandwidth is *greater* when it

---

<sup>6</sup>“Netflix says Internet download caps only in place to drive up bills,” *thespec.com*, March 29, 2011 (accessed August 15, 2013 at URL <http://www.thespec.com/news-story/2202174-netflix-says-internet-download-caps-only-in-place-to-drive-up-bills/>). See too “Netflix Wants Help from U.S. Against Cable Data Caps” by Todd Shields, *Bloomberg*, June 27, 2012 (accessed August 18, 2014 at URL <http://www.bloomberg.com/news/2012-06-27/netflix-wants-help-from-u-s-against-cable-data-caps.html>), which also discusses Netflix’s contention that download caps hurt it.

<sup>7</sup>See, *e.g.*, “If Net Neutrality Is Coming, So Is The End Of All-You-Can-Eat Internet Access,” Dan Frommer, *Business Insider*, December 1, 2010 (accessed online August 15, 2013 at URL <http://www.businessinsider.com/if-net-neutrality-is-coming-so-is-the-end-of-all-you-can-eat-internet-access-2010-12>).

<sup>8</sup>See, *e.g.*, “AT&T puts broadband users on monthly allowance,” Ryan Singel, *Wired*, March 15, 2011 (accessed online August 15, 2013 at URL <http://www.cnn.com/2011/TECH/web/03/15/att.broadband.allowance.wired/index.html>). A relevant quote from this article is “There’s [*sic*] little data to demonstrate whether large ISPs actually are experiencing real issues with congestion.”

can impose a download cap or otherwise restrict consumption than when it cannot (see Proposition 10 in Section 6). Intuitively, the platform can extract more rent from content providers via such measures the more content they sell in equilibrium. By expanding its bandwidth, the platform can expand the amount of content sold. Whether the benefits of greater bandwidth outweigh the static inefficiency caused by those measures is ambiguous, as we illustrate via an example.

The principal focus of our analysis is on the effect of caps and other measures on the content providers' pricing. The logic, however, extends, as we show in Section 7, to situations in which content is provided for free, but the content providers choose quality. The platform's imposition of download caps or positive per-unit fees reduces the content providers's demand, which they will seek to counteract by improving their quality. Improved quality increases consumers' surplus, which the platform can, in turn, capture via higher access charges.

Most of the paper assumes homogeneity across consumers and symmetry amongst content providers. In large part, those assumptions are necessary to have tractable models. In Section 8, we use a simplified version of our base model to explore issues that arise with heterogeneity, such as price discrimination across consumers and the mix of content providers active in equilibrium. In that section, we derive conditions such that a platform would never offer a plan with unlimited downloads; that is, discrimination across consumers will not lead a platform to offer unlimited downloads to any consumer.<sup>9</sup> In other words, unlike standard models of discrimination, in which there is no distortion at the top, we find that *all* consumer types will be sold less than first-best quantity. We further show in that section that our previous welfare results are not dependent on symmetry across the content providers. That the platform fails generically to adopt welfare-maximizing policies given asymmetric content providers arises for two reasons: one, the platform is more concerned about marginal effects than infra-marginal effects, so distortions arise (a result in the spirit of Spence, 1975, and others); and, two, the platform is seeking to extract rents from the content providers, as well as capturing consumer surplus. Hence, while the first reason makes it ambiguous as to whether the platform would set too liberal or too stringent a cap *vis-à-vis* the welfare optimum, the second reason leads it to set too stringent a cap.

We note that a phenomenon related to our rent-extraction effect of download caps can arise with advertising.<sup>10</sup> Let the platform be a free-to-public media outlet (*e.g.*, a radio station or free newspaper). Consumers (listeners, readers) are one side of the market, merchants the other. Suppose the merchants are in competition and can attract consumers only if they place an ad on the platform. By limiting the number of ads it accepts, the platform lessens the competition among the merchants, thereby raising their profits; which, in turn,

---

<sup>9</sup>In this regard, we note that, until recently, the Canadian ISP Rogers, which offers many plans, only offered plans with download caps. Source: <http://www.rogers.com> (accessed on September 17, 2013 and August 18, 2014).

<sup>10</sup>We thank Chengsi Wang for this observation.

it can capture via higher advertising rates.<sup>11</sup> Indeed, by selling advertising to a single merchant, it creates a monopoly in the relevant product market and it can capture the monopoly profit by setting the advertising fee equal to that profit. Although a related phenomenon, there is a critical difference: in this paper, the platform is seeking to induce the merchants (content providers) to charge *lower* prices, not higher, because the platform benefits by increasing, not decreasing, consumer surplus.

This paper is part of the emerging literature on two-sided markets (see *e.g.*, Roson, 2005, Rochet and Tirole, 2006, and Rysman, 2009, for surveys), specifically the literature on monopoly platform practices.<sup>12</sup> In terms of the model employed, we build most directly on Hermalin and Katz (2007) and Economides and Hermalin (2012), although neither considers download caps. In two related articles, Dai and Jordan (2013a,b) consider download caps in the context of price discrimination by an ISP in a residential market. Unlike us, they don't model price or quality setting by the content providers and, so, do not consider the effect download caps have on those choices. Beyond the Dai and Jordan articles, the academic literature on download caps appears limited and focused on legal and regulatory issues (see, *e.g.*, Van Gorp and Middleton, 2010).

## 2 BASELINE MODEL

Consumers want to engage with some or all of  $N$  content (or application) providers. To reach consumers, the providers' content must go across a monopoly platform. The platform has a capacity (bandwidth) of  $B$ ; that is,  $B$  units (*e.g.*, bytes) can go from the content providers to the consumers per unit of time.<sup>13</sup>

The assumed sequence of play is that the platform moves first, announcing its policies and prices. Second, consumers decide whether to purchase access. Next, the content providers announce their prices. Finally, consumers decide how much content to purchase from each content provider.

This game admits multiple equilibria. One is a degenerate equilibrium in which consumers expect the content providers to set such exorbitant prices that they are deterred from purchasing access. Because consumers don't then acquire access and there are, thus, no consumers to whom to sell, it is a weak best response for the content providers to indeed set exorbitant prices. In what follows, we ignore that equilibrium and focus instead on the equilibrium in which consumers acquire access. Consumers anticipate the prices that will emerge in the subgame that follows if they acquire access. Provided they can attain non-negative surplus, they will acquire access. Because the platform can make money

<sup>11</sup>Dukes and Gal-Or (2003) make a similar point, although in the context of competing media platforms.

<sup>12</sup>A partial list of papers on this topic includes Hermalin and Katz (2007), Choi and Kim (2010), Krämer and Wiewiorra (2012), Cheng et al. (2011), Economides and Tåg (2012), Economides and Hermalin (2012), and Choi et al. (2013). None of these papers, however, consider download caps.

<sup>13</sup>There are some nuanced issues concerning time, which we address in Section 9.

only if the consumers acquire access, it will set its prices and policies so that consumer surplus will be non-negative.

We assume that content providers are limited to linear pricing. Denote content provider  $n$ 's price by  $p_n$ . We further assume that the content providers have zero marginal costs—a reasonable approximation for digital goods—and that they are not in direct competition. Neither of these last two assumptions is critical. Setting marginal cost to zero simplifies the notation and the reader will see that extending the analysis to allow for a positive marginal cost is trivial (at least if it's a constant—presumably true of digital goods). In terms of direct competition, all that is necessary for our basic points to hold is that content providers perceive their demand curves as downward sloping and any download limits *intensify* price competition. The assumption of no direct competition serves to keep the formal analysis tractable.

The platform chooses a download limit (cap),  $L$ ; a hookup (access) fee,  $H$ ; and two per-unit (*e.g.*, per-byte or packet) charges,  $\tau_b$  and  $\tau_o$ . The first per-unit charge applies to all units below the cap, the second to those over the limit (*i.e.*, the second is an overage charge). Hence, a consumer's total payment to the platform if she acquires access and downloads  $x$  total units is

$$H + \tau_b \min\{x, L\} + \tau_o \max\{0, x - L\}.$$

Observe, a binding-cap regime corresponds to  $\tau_o = \infty$ , while a conventional two-part tariff with no cap is equivalent to either  $L = \infty$  (so  $\tau_b$  is the per-unit charge) or  $L = 0$  (so  $\tau_o$  is the per-unit charge). It will turn out, as we show below, that the platform never needs to utilize all four instruments in equilibrium. Consistent with actual practice, we rule out the platform's setting its prices and policies to be contingent on the prices announced by the content providers.

There is a measure one of consumers. Each consumer has quasi-linear utility:

$$U = \sum_{n=1}^N \int_0^{x_n} (\alpha - x\delta(X|B))dx + y, \quad (1)$$

where  $x_n$  is the amount of content acquired from the  $n$ th content provider,  $y$  the numéraire good,  $\alpha$  a positive constant, and  $\delta(X|B)$  relates to the disutility (loss) arising from congestion when  $X$  total content is transmitted.<sup>14</sup> The functional-form assumption on marginal utility—namely that it is  $\alpha - x\delta(X|B)$ —is necessary to have a tractable model.

We assume that a consumer's expenditure on access and content is never so great as to consume all her income; that is, if her income is  $I$ , we assume

$$0 < y = I - H - \sum_{n=1}^N p_n x_n - \tau_b \min\left\{\sum_{n=1}^N x_n, L\right\} - \tau_o \max\left\{0, -L + \sum_{n=1}^N x_n\right\} \quad (2)$$

<sup>14</sup>The utility function in (1) is similar to the ones used by Economides and Hermalin (2012) and Hermalin and Katz (2007); the principal difference being they utilized more general marginal utility functions in which the marginal utility of the  $x$ th unit is  $\mu(x\delta(X|B))$ ,  $\mu(\cdot)$  a decreasing function.

holds always.

For low levels of total platform usage, it is possible that congestion is irrelevant, in which case we set  $\delta(X|B) = 1$ .<sup>15</sup> Otherwise, we assume that congestion and, hence, disutility increases with total usage; that is, for any  $B$ ,  $\delta(\cdot|B)$  is an increasing function. We also assume it is continuous.

Because each consumer is negligible, she rationally does not take into account how her consumption affects congestion. Although one could imagine that content providers recognize their effects on congestion, it seems more plausible that they do not and we limit attention to that case. We note that both assumptions enhance the possibility of our finding that the platform's imposition of download restrictions is welfare improving. Hence, conclusions in the analysis that follows that such restrictions are not welfare improving can be viewed as fairly robust.

### 3 PRELIMINARY ANALYSIS

Assuming a consumer has decided to connect to the platform, she then chooses the amount of content to purchase—a bundle  $(x_1, \dots, x_N)$ —to maximize her utility, expression (1). Substituting for  $y$ , according to (2), her optimization problem can be written as

$$\begin{aligned} \max_{(x_1, \dots, x_N) \in \mathbb{R}_+^N} & \sum_{n=1}^N \int_0^{x_n} (\alpha - x\delta(X|B)) dx \\ & - H - \sum_{n=1}^N p_n x_n - \tau_b \min \left\{ \sum_{n=1}^N x_n, L \right\} - \tau_o \max \left\{ 0, -L + \sum_{n=1}^N x_n \right\}. \quad (3) \end{aligned}$$

Observe that if a given bundle is a solution to (3) for one consumer, it is a solution for all consumers; that is, consumers behave the same.

The ensuing analysis is greatly simplified by the following lemma:

**Lemma 1.** *Consider an equilibrium in which consumers acquire access and the platform's policy and prices are  $L$ ,  $H$ ,  $\tau_b$ , and  $\tau_o$ .*

- (i) *If each consumer's total content consumption in equilibrium does not equal  $L$  (i.e.,  $\sum_{n=1}^N x_n \neq L$ ), then there exists a two-part tariff  $(\tilde{H}, \tau)$  with no cap (i.e.,  $L = \infty$ ) that induces an equilibrium that is equivalent to the original equilibrium in that the content providers set the same prices, consumers make the same access and consumption decisions, and the platform earns the same profit.<sup>16</sup>*
- (ii) *If each consumer's total content consumption in equilibrium equals  $L$  (i.e.,  $\sum_{n=1}^N x_n = L$ ), then there exists a binding-cap regime (i.e.,  $\tilde{\tau}_o = \infty$ ) with*

<sup>15</sup>This normalization is solely for convenience and imposes no loss of generality.

<sup>16</sup>As observed earlier, a two-part tariff is also equivalent to setting a cap of 0 and charging an "overage" fee on all content.

*a hookup fee, but no per-unit fee (i.e.,  $\tilde{\tau}_b = 0$ ), that induces an equilibrium that is equivalent to the original equilibrium in that the content providers set the same prices, consumers make the same access and consumption decisions, and the platform earns the same profit.*

**Proof:** Please see appendix. ■

Intuitively, if a consumer acquires less than  $L$ , it and  $\tau_o$  are effectively irrelevant to her; it is as if she simply faces the two-part tariff  $(H, \tau_b)$ . If she acquires precisely  $L$ , then raising  $\tau_o$  to infinity would be irrelevant to her. Given a binding constraint, her expenditure  $\tau_b L$  is effectively sunk and can be folded into the hookup fee (i.e.,  $\tilde{H} = H + \tau_b L$ ). Given the separability inherent in her utility function, reducing the per-unit fee to zero would not, then, affect her allocation of  $L$  among the content providers' content (i.e., the  $x_n$ s would remain unchanged). Finally, if she acquires more than  $L$ , then  $\tau_b L$  is again a sunk expenditure, which can be folded into the hookup fee. On the margin, the relevant per-unit fee is  $\tau_o$ . Hence, from her perspective, the regime is equivalent to facing a two-part tariff with a hookup fee of  $H + \tau_b L - \tau_o L$  and per-unit fee of  $\tau_o$ . In short, the alternative policies and prices don't change the consumer's behavior provided the content providers don't change their prices; but this, in turn, proves sufficient for the content providers not to wish change their prices: the old prices remain optimal (best responses) in the new regime if they continue to induce the same demand from the consumers.

Given Lemma 1, attention can be restricted, without loss of generality, to binding-cap regimes (with no per-unit fees) and simple two-part tariffs (with no download limits). For the sake of brevity, hereafter the phrase "binding-cap regime" is to be understood to imply no per-unit fee and the phrase "two-part tariff" to imply no download limit. As the proof of Lemma 1 demonstrates, one can also go the other way; that is, construct more elaborate tariffs that are equivalent to either a two-part tariff or a binding-cap regime.

From (3), under a two-part tariff, it is readily seen that a consumer maximizes her utility by consuming

$$x_n = \frac{\alpha - \tau - p_n}{\delta(X|B)}, \quad (4)$$

$n = 1, \dots, N$ . Given a binding-cap regime, it can be shown that a consumer's demand for content provider  $n$ 's product is

$$x_n = \frac{L}{N} + \frac{\sum_{j \neq n} p_j - (N-1)p_n}{N\delta(X|B)} \quad (5)$$

(see the proof of part (ii) of Lemma 1 for details). Given that there is a measure one of consumers, expressions (4) and (5) are content provider  $n$ 's demand, respectively, when the platform utilizes a two-part tariff and when it utilizes a binding-cap regime. Notice that the imposition of a binding download cap effectively turns the previously independent goods into substitute goods (at least from the perspective of the content providers' strategies).

#### 4 BENCHMARK: NO CONGESTION

As a benchmark, we consider the case in which there is no congestion externality; that is,  $\delta(X|B) \equiv 1$ . To help appreciate the platform's rationale for imposing a download cap or a positive per-unit fee, we start by taking the content providers' prices as having been set *exogenously* (but such that (4) and (5) are nonnegative).

Under a two-part tariff, a consumer's surplus from her purchase of content provider  $n$ 's product is

$$\int_0^{\alpha - \tau - p_n} (\alpha - x) dx - (\tau + p_n)(\alpha - \tau - p_n) = \frac{1}{2}(\alpha - \tau - p_n)^2.$$

The platform can extract all consumer surplus by setting the hookup fee equal to the sum over  $n$  of that expression; that is,

$$H = \frac{1}{2} \sum_{n=1}^N (\alpha - \tau - p_n)^2. \quad (6)$$

Its profit is, thus,

$$H + \tau \sum_{n=1}^N x_n = \frac{1}{2} \sum_{n=1}^N (\alpha - \tau - p_n)^2 + \tau \sum_{n=1}^N (\alpha - \tau - p_n). \quad (7)$$

The derivative of (7) with respect to  $\tau$  is  $-N\tau \leq 0$ ; hence, were the content providers' prices set exogenously, the platform would optimally set the per-unit fee to zero. This result is, of course, no more than the standard result that, with homogenous consumers, the optimal two-part tariff has a per-unit fee equal to marginal cost (in this case, 0).

Under a binding-cap regime, a consumer's surplus from her purchase of content provider  $n$ 's product is

$$\begin{aligned} & \int_0^{\frac{1}{N}(L + \sum_{j \neq n} p_j - (N-1)p_n)} (\alpha - x) dx - p_n \frac{1}{N} \left( L + \sum_{j \neq n} p_j - (N-1)p_n \right) \\ &= \left( \alpha - p_n - \frac{1}{2N} \left( L + \sum_{j \neq n} p_j - (N-1)p_n \right) \right) \frac{1}{N} \left( L + \sum_{j \neq n} p_j - (N-1)p_n \right). \end{aligned}$$

The hookup fee equals the sum of that expression over  $n$ . Given the hookup fee equals the platform's profit, the optimal  $L$  can be found by differentiating that sum with respect  $L$ .<sup>17</sup> Straightforward algebra reveals the resulting first-order condition is equivalent to

$$L = \sum_{n=1}^N (\alpha - p_n). \quad (8)$$

<sup>17</sup>It is readily seen that program is strictly concave in  $L$ .

Observe that  $\alpha - p_n$  would be the demand for the  $n$ th content provider's product *absent* a download constraint; hence, expression (8) entails that, were the content providers' prices set exogenously (unaffected by the platform's actions), the optimal download "cap" would equal the total amount of content a consumer would demand absent any cap. This makes intuitive sense: restricting consumers only serves to reduce their surplus, which reduces the amount the platform can capture via the hookup fee. To summarize:

**Proposition 1.** *Suppose there is no congestion externality and the platform's policy and pricing do not affect the content providers' pricing. Then the profit-maximizing policy and pricing for the platform to adopt is no download cap (i.e.,  $L = \infty$ ), no per-unit fee (i.e.,  $\tau = 0$ ), and a hookup fee,  $H$ , given by (6) with  $\tau = 0$ .*

Our focus is, however, on situations in which the platform's policies and pricing *do* affect the content providers' pricing. Suppose, first, that the platform sets a two-part tariff. Each content provider sets its price so as to maximize the product of its price and demand, the latter being given by expression (4). The solution to that optimization program is readily seen to be  $p_n = (\alpha - \tau)/2$  (recall we've abstracted from congestion externalities, so  $\delta(X|B) \equiv 1$ ). A consumer's surplus from her consumption of a given content provider's content is

$$\int_0^{\frac{1}{2}(\alpha - \tau)} (\alpha - x) dx - \left( \frac{\alpha - \tau}{2} + \tau \right) \frac{1}{2}(\alpha - \tau) = \frac{(\alpha - \tau)^2}{8}.$$

Consequently, the platform will charge a hookup fee of  $H = N(\alpha - \tau)^2/8$ . Its profit, as a function of  $\tau$ , is

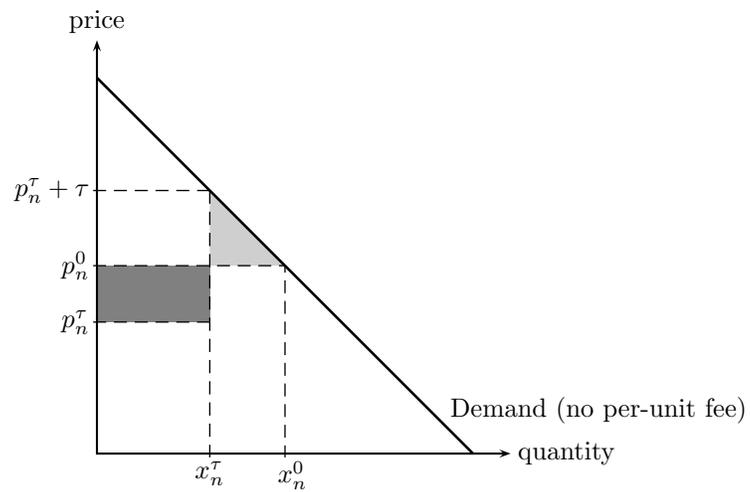
$$N \frac{(\alpha - \tau)^2}{8} + \tau N \frac{\alpha - \tau}{2}.$$

It is readily seen that expression is maximized by setting  $\tau = \alpha/3$ , yielding the platform profit equal to  $N\alpha^2/6$ . To summarize:

**Proposition 2.** *Suppose there is no congestion externality, but the platform's policy and pricing do affect the content providers' pricing. Then the profit-maximizing two-part tariff for the platform has a positive per-unit fee ( $\tau = \alpha/3$ ). The profit-maximizing hookup fee is  $N\alpha^2/18$ .*

The explanation for the difference between this last result and Proposition 1 is Public Finance 101: the per-unit fee acts like an excise tax. The incidence of that "tax" is split between the consumers and the content providers. In particular, it causes the content providers to lower their prices by enough that "tax revenue" plus residual consumer surplus post tax exceeds what consumer surplus was pre tax. Figure 1 illustrates. Because, as that figure shows, setting  $\tau > 0$  reduces the quantity traded (from  $x_n^0$  to  $x_n^\tau$ ), welfare must be lower when the platform is permitted to set  $\tau > 0$  than when it is restricted to  $\tau = 0$ .

**Corollary 1.** *Absent a congestion externality, allowing the platform to charge a per-unit fee in addition to a hookup fee is welfare reducing.*



**Figure 1:** Rationale for a positive per-unit fee: although such a fee reduces consumer surplus, the induced reduction in each content provider's price allows the platform to capture  $\tau \times x_n^\tau$  in direct charges. The platform, thus, loses the area of the light-gray triangle, but gains the area of the dark-gray rectangle, which is greater than the light-gray triangle.

If there is a binding download limit, then, from (5), each content provider prices to maximize

$$p_n \left( \frac{L}{N} + \frac{\sum_{j \neq n} p_j - (N-1)p_n}{N} \right).$$

The corresponding first-order condition is equivalent to

$$L + \sum_{j \neq n} p_j - 2(N-1)p_n = 0,$$

which yields the best-response function

$$p_n = \frac{L + \sum_{j \neq n} p_j}{2(N-1)}.$$

The unique Nash equilibrium is

$$p_1 = \dots = p_N = \frac{L}{N-1}.$$

Hence, from (5), each consumer acquires  $L/N$  from each content provider in equilibrium. Her surplus from doing so is, consequently,

$$\int_0^{L/N} (\alpha - x) dx - \underbrace{\frac{L}{N-1}}_{p_n} \frac{L}{N} = L \frac{\alpha}{N} - \frac{(3N-1)L^2}{2(N-1)N^2}.$$

It follows that the platform will charge a hookup fee of

$$H_{\text{CAP}} = L\alpha - \frac{(3N-1)L^2}{2(N-1)N}. \quad (9)$$

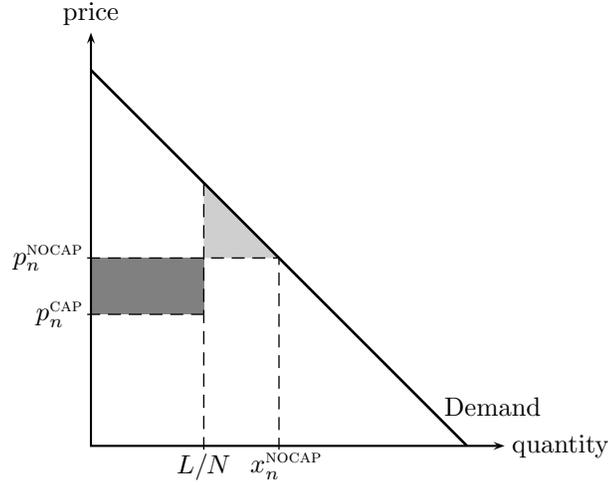
To maximize its profit, the platform will set the download limit to maximize (9); hence,

$$L = \frac{\alpha(N-1)N}{3N-1}. \quad (10)$$

Note this solution is relevant only if the  $L$  given by (10) is less than  $N\alpha/2$ , which is the total content downloaded absent a limit (or if the limit does not bind). Because the  $L$  given by (10) is always less than  $N\alpha/3$ , it follows the limit binds. Expressions (9) and (10) yield the equilibrium access (hookup) fee:

$$H_{\text{CAP}} = \frac{N\alpha^2}{8} \times \frac{4(N-1)}{3N-1}.$$

From the analysis of two-part tariffs without download limits,  $N\alpha^2/8$  is the hookup fee the platform would charge if it chose  $\tau = 0$ . It therefore follows that



**Figure 2:** Rationale for download limits: surplus extraction from a single content provider  $n$ . By effectively inducing the content providers to compete, consumer surplus increases by the area of the dark rectangle less the areas of the light triangle.

**Proposition 3.** *Given the functional forms assumed, relative to charging the profit-maximizing hookup fee absent a download limit and absent a per-unit fee, the platform strictly prefers to impose a download limit if there are four or more content providers; is indifferent between a limit and no limit if there are three content providers; and prefers no limit if there are two or fewer content providers.*

Intuition for Proposition 3 can be gained from Figure 2. The download limit effectively induces competition among the content providers. Hence, their prices fall, which will increase consumer surplus *ceteris paribus* (the dark rectangle in the figure). At the same time, though, the content purchased from each content provider falls, which reduces surplus *ceteris paribus* (the light triangle). However, as the figure suggests—and the algebra confirms—the increase due to competition can outweigh the loss due to reduced consumption. Not surprisingly, the competition effect is greater the more content providers there are, which helps explain why the platform finds a download limit profitable when there are many content providers, but not when there are only a few.

Because the content acquired from each content provider is reduced and there are no congestion externalities, the download limit must reduce welfare, as illustrated in Figure 2.

Which regime—two-part tariff or binding-cap—would the platform prefer? Recall, from Lemma 1, these are the relevant options for the platform. Its

profit under the profit-maximizing binding-cap regime is  $H_{\text{CAP}}$ . Its profit under the profit-maximizing two-part tariff is, recall,  $N\alpha^2/6$ . Because

$$\frac{1}{6} > \frac{N-1}{6N-2}, \quad (11)$$

it follows that the platform does better using a two-part tariff than a binding cap. On the other hand, the righthand side of (11) tends to  $1/6$  as  $N \rightarrow \infty$ ; hence, if there are a large number of content providers, the platform's loss from employing a binding-cap regime rather than a two-part tariff is minor. If there are significant transaction costs in administering a two-part tariff, the platform could prefer the binding-cap regime. To summarize:

**Proposition 4.** *The platform can make a greater profit with a two-part tariff than under a binding-cap regime. However, as the number of content providers gets large, the difference in profits shrinks; in the limit, profits are the same under the two regimes. Consequently, if there are significant transaction costs associated with administering a two-part tariff, then the platform will prefer the binding-cap regime when the number of content providers is large.*

In terms of welfare, observe that content demanded from each content provider under the profit-maximizing two-part tariff is  $\alpha/3$ . From (10), it is

$$\frac{\alpha(N-1)}{3N-1}$$

under the profit-maximizing binding-cap regime; hence, welfare is greater under the former  $((N-1)/(3N-1) < 1/3)$ . The difference, though, shrinks as  $N$  gets large  $((N-1)/(3N-1) \rightarrow 1/3$  as  $N \rightarrow \infty)$ . To summarize:

**Proposition 5.** *Welfare is greater when the platform utilizes the profit-maximizing two-part tariff than when it utilizes the profit-maximizing binding-cap regime. This difference in welfare, however, shrinks to zero as the number of content providers grows large.*

## 5 CONGESTION EXTERNALITIES

Assume, now, that there is a negative congestion externality.

If there is no download limit, then, from (4), it follows that a content provider's profit-maximizing price is  $p_n = (\alpha - \tau)/2$ . Now, however, there is a consistency constraint: total demand must be consistent with the effect the resulting congestion has on demand; that is,

$$X = \sum_{n=1}^N x_n = \sum_{n=1}^N \left( \frac{\alpha - \tau - \overbrace{(\alpha - \tau)/2}^{p_n}}{\delta(X|B)} \right) = \frac{N(\alpha - \tau)}{2\delta(X|B)}. \quad (12)$$

Because (i) the leftmost side of (12) is zero at  $X = 0$ , while the rightmost side is positive; (ii) the leftmost side increases without bound in  $X$  while the rightmost side is non-increasing; and (iii)  $\delta(\cdot|B)$  is continuous, it follows that a unique

solution to (12) exists. Call it  $X(\tau, B)$ . Because  $\delta(\cdot|B)$  is an increasing function, it is readily seen from (12) that  $X(\cdot, B)$  is a decreasing function (*i.e.*, not surprisingly, an increase in the per-unit fee reduces total content consumption).

Each content provider sells  $X(\tau, B)/N$  in content, so consumer surplus per content provider is

$$\begin{aligned} & \int_0^{X(\tau, B)/N} \left( \alpha - x\delta(X(\tau, B)|B) \right) dx - \left( \tau + \frac{\alpha - \tau}{2} \right) \frac{X(\tau, B)}{N} \\ &= \frac{X(\tau, B) \left( (\alpha - \tau)N - X(\tau, B)\delta(X(\tau, B)|B) \right)}{2N^2} = \frac{(\alpha - \tau)X(\tau, B)}{4N}, \end{aligned}$$

where the last equality follows from (12). It follows the hookup fee charged by the platform is  $H = (\alpha - \tau)X(\tau, B)/4$ . The platform's profit, as a function of  $\tau$ , is therefore

$$\frac{(\alpha - \tau)X(\tau, B)}{4} + \tau X(\tau, B) = \frac{1}{4}(\alpha + 3\tau)X(\tau, B). \quad (13)$$

Given the absence of costs, welfare is gross benefit:

$$W = N \int_0^{X(\tau, B)/N} \left( \alpha - x\delta(X(\tau, B)|B) \right) dx = \frac{1}{4}(3\alpha + \tau)X(\tau, B), \quad (14)$$

where the second equality follows from (12).

Because optimization programs are invariant with respect to increasing monotonic transformations, choosing  $\tau$  to maximize (13) is equivalent to choosing  $\tau$  to maximize

$$\log(\alpha + 3\tau) + \log(X(\tau, B)); \quad (13')$$

similarly, choosing  $\tau$  to maximize (14) is equivalent to choosing it to maximize

$$\log\left(\alpha + \frac{1}{3}\tau\right) + \log(X(\tau, B)). \quad (14')$$

Observe the objective functions in (13') and (14') are of the form

$$\log(\alpha + \zeta\tau) + \log(X(\tau, B)),$$

the cross-partial derivative of which with respect to  $\zeta$  and  $\tau$  is

$$\frac{\alpha}{(\alpha + \zeta\tau)^2} > 0.$$

Consequently, by the usual comparative statics, the optimal  $\tau$  is increasing in  $\zeta$ ; from which it follows that the  $\tau$  that maximizes (13') and, thus, (13) is greater than the one that maximizes (14') and, thus, (14). This analysis establishes:

**Proposition 6.** *Assume a congestion externality exists. Suppose the platform uses two-part tariffs. A profit-maximizing platform will set a per-unit fee (*i.e.*,  $\tau$ ) that is greater than the per-unit fee that would maximize welfare.*

Consider, now, a binding-cap regime. From (5), each content provider sets its price to maximize

$$p_n \left( \frac{L}{N} + \frac{\sum_{j \neq n} p_j - (N-1)p_n}{N\delta(L|B)} \right).$$

The corresponding first-order condition is equivalent to

$$L\delta(L|B) + \sum_{j \neq n} p_j - 2(N-1)p_n = 0,$$

which yields the best-response function

$$p_n = \frac{L\delta(L|B) + \sum_{j \neq n} p_j}{2(N-1)}.$$

The unique Nash equilibrium is

$$p_1 = \dots = p_N = \frac{L\delta(L|B)}{N-1}. \quad (15)$$

Expression (15) might, at first, seem counterintuitive: equilibrium price is *increasing* in the distaste for congestion *ceteris paribus*. This result can, however, be understood by considering (5). The greater is distaste, the less sensitive a consumer's demand is to relative prices *ceteris paribus*; hence, the lower are the competitive pressures arising from the download limit and, hence, the higher the price the content providers find themselves able to charge.

From (5), each consumer acquires  $L/N$  from each content provider in equilibrium. Her surplus from doing so is, therefore,

$$\int_0^{L/N} (\alpha - x\delta(L|B)) dx - \frac{L\delta(L|B)}{N-1} \frac{L}{N} = L \frac{\alpha}{N} - \frac{(3N-1)\delta(L|B)L^2}{2(N-1)N^2}.$$

Consequently, the platform will charge a hookup fee of

$$H_{\text{CAP}} = L\alpha - \frac{(3N-1)}{2(N-1)N} \delta(L|B)L^2. \quad (16)$$

Welfare is

$$\begin{aligned} W_{\text{CAP}} &= N \int_0^{L/N} (\alpha - x\delta(L|B)) dx \\ &= L\alpha - \left( \frac{(3N-1)}{2(N-1)N} - \frac{2N}{2(N-1)N} \right) \delta(L|B)L^2. \end{aligned} \quad (17)$$

Observe, in terms of  $L$ , that  $H_{\text{CAP}}$  and  $W_{\text{CAP}}$  have a common “benefit-like” term (*i.e.*,  $L\alpha$ ), but the former has a “cost-like” term with a higher margin. Hence, by the usual comparative statics, it must be that the  $L$  the platform would set to maximize its profit is *lower* than the cap that would maximize welfare. To summarize:

**Proposition 7.** *Assume a congestion externality exists. Suppose the platform uses a binding-cap regime. A profit-maximizing platform will set a download limit (cap) that is lower than the limit that would maximize welfare.*

Propositions 6 and 7 establish that even in the face of a congestion externality, the platform will fail to adopt welfare-maximizing policies and prices. This leaves open the question, however, whether—in the face of congestion—allowing the platform to act freely is welfare superior to prohibiting it from imposing a cap or a per-unit fee? In general, the answer is complicated because of the number of forces at work:

1. because neither content providers nor consumers take into account the congestion externality, there will be a tendency to transmit too much content *ceteris paribus*;
2. but, because the content providers exercise market power, too little content would get traded if there were no congestion externality; and,
3. as noted, the platform has incentives to set policies and prices that are not welfare maximizing.

The first two points indicate the theory of the second best is at work absent any actions by the platform: the reduced trade due to the exercise of market power partially offsets the congestion externality (or, conversely, because of the congestion externality, the welfare loss from the exercise of market power is reduced).

An additional question is whether an unconstrained platform would do better to employ a two-part tariff or a binding-cap regime?

Neither question proves tractable utilizing a general loss-from-congestion function. Hence, we assume:

$$\delta(X|B) = \Lambda \left( \frac{X}{B} \right)^\theta, \quad (18)$$

where  $\Lambda > 0$  and  $\theta \geq 0$  are constants. The analysis of Section 4 corresponds to  $\Lambda = 1$  and  $\theta = 0$ .

Consider a two-part tariff. Given (18), the solution to (12) is

$$X(\tau, B) = \left( \frac{N(\alpha - \tau)B^\theta}{2\Lambda} \right)^{\frac{1}{\theta+1}},$$

Recall the platform's profit is  $(\alpha + 3\tau)X(\tau, B)/4$  (see expression (13)); hence, using (18), its profit is

$$\frac{1}{4} \left( \frac{N(\alpha - \tau)B^\theta}{2\Lambda} \right)^{\frac{1}{\theta+1}} (\alpha + 3\tau). \quad (19)$$

Maximizing with respect to  $\tau$  yields

$$\tau^* = \frac{\alpha(2 + 3\theta)}{3(2 + \theta)}.$$

Substituting that back into (19) reveals the platform's profit to be

$$\frac{\theta + 1}{\theta + 2} \left( \frac{2N\alpha^{\theta+2}B^\theta}{3\Lambda(\theta + 2)} \right)^{\frac{1}{\theta+1}}. \quad (20)$$

Welfare given the profit-maximizing per-unit fee (*i.e.*,  $\tau^*$ ) is, from (14) and (18),

$$W_{\tau^*} = \frac{3\theta + 5}{3(\theta + 2)} \left( \frac{2N\alpha^{\theta+2}B^\theta}{3\Lambda(\theta + 2)} \right)^{\frac{1}{\theta+1}}. \quad (21)$$

Welfare should the per-unit fee be fixed at 0 is

$$W_0 = \frac{3}{4} \left( \frac{2N\alpha^{\theta+2}B^\theta}{4\Lambda} \right)^{\frac{1}{\theta+1}}.$$

Observe that two welfare levels can be related:

$$W_{\tau^*} = W_0 \left( \frac{4}{3(\theta + 2)} \right)^{\frac{\theta+2}{\theta+1}} \frac{3\theta + 5}{3}. \quad (22)$$

From (22), it follows that  $W_0 > W_{\tau^*}$  if  $\theta < 6.47$  (approximately) and  $W_0 < W_{\tau^*}$  otherwise. (We consider the economic meaning of this later.)

Suppose the platform uses a binding-cap regime. Maximizing the platform's profit, expression (16), with respect to  $L$ , utilizing (18), yields

$$L = \left( \frac{2\alpha(N - 1)NB^\theta}{\Lambda(3N - 1)(\theta + 2)} \right)^{\frac{1}{\theta+1}}.$$

Substituting yields:

$$H_{\text{CAP}} = \frac{\theta + 1}{\theta + 2} \left( \frac{2N\alpha^{\theta+2}B^\theta}{3\Lambda(\theta + 2)} \right)^{\frac{1}{\theta+1}} \left( \frac{3(N - 1)}{3N - 1} \right)^{\frac{1}{\theta+1}}. \quad (23)$$

Comparing (23), which is the platform's maximum profit under a binding-cap regime to (20), its maximum profit under a two-part tariff, it follows maximum profits under a two-part tariff are greater given  $3(N - 1) < 3N - 1$ . At the same time, because

$$\lim_{N \rightarrow \infty} \frac{3(N - 1)}{3N - 1} = 1,$$

it follows that the difference between (23) and (20) converges to zero as  $N$  gets larger. To summarize:

**Proposition 8.** *Assume a congestion externality and the loss-from-congestion function is given by (18). Then a profit-maximizing platform makes greater profit using a two-part tariff than using a binding-cap regime. However, as the number of content providers gets large, this difference shrinks; in the limit, profits are the same under the two regimes. Hence, if there are significant transaction fees associated with administering a two-part tariff, then the platform would prefer a a binding-cap regime if the number of content providers is sufficiently large.*

Consider welfare under the profit-maximizing binding-cap regime: substituting  $L$  into the statement for welfare, expression (17), it can be shown that

$$W_{\text{CAP}} = W_0 \left( \frac{4}{3(\theta + 2)} \right)^{\frac{\theta+2}{\theta+1}} \left( \frac{3(N-1)}{3N-1} \right)^{\frac{1}{\theta+1}} \left( \frac{3N\theta + 5N - \theta - 1}{3N-1} \right). \quad (24)$$

Comparing (22) and (24), it is readily seen they converge to the same expression as  $N \rightarrow \infty$ . Moreover, differentiating (24) with respect to  $N$  (treating  $W_0$  as fixed) reveals that expression is increasing in  $N$ . Consequently, it must be that  $W_{\text{CAP}} < W_{\tau^*}$ : welfare is greater under the platform's profit-maximizing two-part tariff than its profit-maximizing binding-cap regime.

This comparison of (22) and (24) further means that, unless  $\theta > 6.47$ , welfare is greater if the platform is barred from imposing a download cap of its choosing. On the other hand, if  $\theta$  is great enough, then allowing the platform to choose the cap is preferable from barring its use of a cap (*e.g.*, if  $\theta = 7$  and  $N = 6$ , then  $W_{\text{CAP}} > W_0$ ).

To summarize:

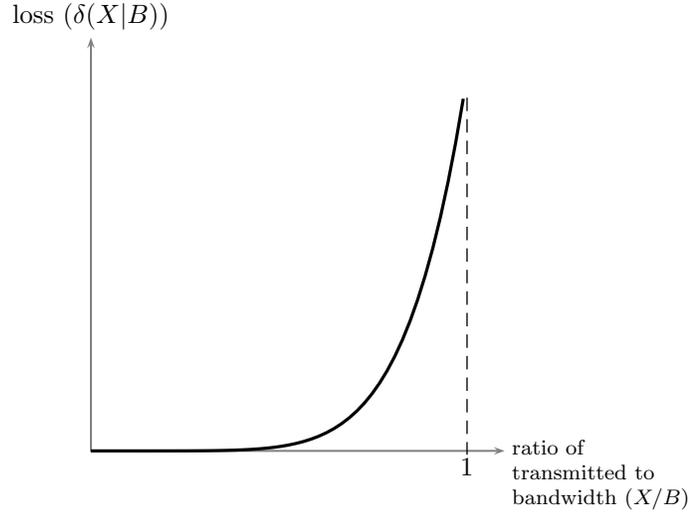
**Proposition 9.** *Assume a congestion externality and the loss-from-congestion function is given by (18). Unless the disutility/loss from congestion function is highly convex (specifically, unless  $\theta \geq 6.47$ ), welfare is greater if the platform is barred from utilizing download caps and per-unit charges of its choosing (*i.e.*, if it is restricted to just a hookup fee and no caps). Welfare under the platform's profit-maximizing two-part tariff exceeds welfare under its profit-maximizing binding-cap regime, but the difference between the two disappears as the number of content providers gets large.*

Figure 3 plots  $z^{6.47}$ , which illustrates how convex it is. It is plausible that such convexity is consistent with actual preferences: presumably marginal disutility/loss is very small at low levels of congestion: going from no freezes in an on-demand video to the occasional freeze due to a small increase in congestion is presumably less costly to people than going from the occasional freeze to constant freezing, as might incur with an increase in congestion at a higher level of congestion. In a sense, Figure 3 can be seen as approximating a backward-L-shaped curve—a “breaking-point” model in which congestion is acceptable to a certain point, but thereafter almost wholly unacceptable.

## 6 ENDOGENOUS BANDWIDTH

The analysis to this point has treated the bandwidth,  $B$ , as fixed. It is possible, over some time horizons at least, that the platform can change the bandwidth. We briefly consider the consequences of endogenizing the bandwidth in this section.<sup>18</sup> Our analysis maintains the previously made functional-form assumptions.

<sup>18</sup>Other articles that have explored platforms' incentives to expand bandwidth include Choi and Kim (2010), Cheng et al. (2011), Economides and Hermalin (2012), and Krämer and Wiewiorra (2012). None of these, however, investigate the relation between bandwidth and



**Figure 3:** A highly convex disutility/loss of congestion function.

If the platform were restricted to charging a hookup fee only (*i.e.*, required to set  $\tau = 0$ ) and barred from imposing a download limit (*i.e.*, required to set  $L = \infty$ ), then its profit is given by (19) evaluated at  $\tau = 0$ . Simplifying that expression, profit is

$$\pi_0^{\text{PLAT}} = \left( \frac{N\alpha^{\theta+2}B^\theta}{2^{2\theta+3}\Lambda} \right)^{\frac{1}{\theta+1}}.$$

Let  $\pi_{\tau^*}^{\text{PLAT}}$  denote its profit under the profit-maximizing two-part tariff (*i.e.*,  $\pi_{\tau^*}^{\text{PLAT}}$  equals the amount in (20)). Let  $\pi_{\text{CAP}}^{\text{PLAT}}$  denote its profit under the profit-maximizing binding-cap regime (*i.e.*,  $\pi_{\text{CAP}}^{\text{PLAT}} = H_{\text{CAP}}$ ).

It is readily verified that, for all three profit measures,

$$\frac{\partial}{\partial B} \pi^{\text{PLAT}} = \frac{\theta}{(\theta+1)B} \pi^{\text{PLAT}}.$$

Hence, utilizing the analysis of the previous section,

$$\frac{\partial \pi_{\tau^*}^{\text{PLAT}}}{\partial B} > \frac{\partial \pi_0^{\text{PLAT}}}{\partial B}, \quad \frac{\partial \pi_{\tau^*}^{\text{PLAT}}}{\partial B} > \frac{\partial \pi_{\text{CAP}}^{\text{PLAT}}}{\partial B}$$

download caps. In terms of the potential *overall* economic value of bandwidth expansion, Greenstein and McDevitt (2011) find that the growth of broadband Internet access increased US GDP by \$8.3 to \$10.6 billion over dial-up. Given the simultaneous change in the nature of content over the period they study (*e.g.*, more on-demand video) and the recongestion effect identified by Economides and Hermalin (2012), it is unclear what relationship those estimates have concerning the contemporaneous reduction in the social costs of congestion (if any reductions occurred).

and, if  $N > 3$ ,<sup>19</sup>

$$\frac{\partial \pi_{\text{CAP}}^{\text{PLAT}}}{\partial B} > \frac{\partial \pi_0^{\text{PLAT}}}{\partial B}.$$

It follows, therefore, that the platform’s marginal return to increasing bandwidth is greatest if it’s permitted to use a two-part tariff of its choosing; moreover, given a sufficient number of content providers, its return to increasing bandwidth is greater if free to impose a download limit of its choosing than if barred from imposing a limit. The following is, thus, a further consequence:

**Proposition 10.** *Assume (i) the platform’s cost of installing bandwidth is everywhere differentiable; and (ii) that, if restricted to no per-unit fee and no download limit, the platform’s choice of bandwidth is an interior solution to the problem of maximizing profit with respect to bandwidth. Then a platform able to choose a two-part tariff of its choosing will install more bandwidth than one barred from doing so. If there are more than three content providers, a platform able to choose a binding-cap regime of its choosing will install more bandwidth than one barred from doing so.*

Proposition 10 may, at first, seem counter-intuitive (especially with respect to a binding cap): one might have expected measures to limit consumption (*e.g.*, a per-unit fee or download cap) to be a substitute for expanding bandwidth, as those measures and expanding bandwidth both address the congestion externality. Although true, there are additional effects. First, the expanded bandwidth is used. This is the re-congestion effect identified by Economides and Hermlin (2012) and familiar to anyone who has seen physical highway expansion fail to end bumper-to-bumper traffic. Greater use increases the potential surplus the platform can capture from the content providers (either indirectly via “taxation”— $\tau > 0$ —or by inducing competition among them— $L < \infty$ ). Moreover, as discussed in conjunction with expression (15), induced competition is fiercer the lower is  $\delta(L|B)$  *ceteris paribus*. Accounting for these other effects, the platform’s incentives to increase bandwidth are greater when unrestricted either in terms of per-unit charges or download caps than when it is barred from using such instruments.

Whether the welfare benefits of greater bandwidth outweigh the potential welfare loss from granting the platform freedom of action is, in general, ambiguous. As an example, suppose that the platform’s cost of installing  $B$  bandwidth is  $kB$ ,  $k$  a positive constant. Continue to assume that  $\delta(X|B)$  is given by (18). Let  $\Lambda = \theta = 1$ . Note, because  $\theta < 6.47$ , static welfare (*i.e.*, for a fixed  $B$ ) is greatest if the platform is barred from setting a positive per-unit fee or imposing a binding download cap. In contrast, in a dynamic setting (*i.e.*, when  $B$  is endogenous), welfare can be greater if those restrictions are lifted. Specifically, continue to suppose the per-unit fee is fixed at zero, but the platform can utilize

<sup>19</sup>Note if  $\theta > 0$ , then  $\pi_{\text{CAP}}^{\text{PLAT}} > \pi_0^{\text{PLAT}}$  provided  $N > 2$ ; and if  $\theta \geq .295$ , then  $\pi_{\text{CAP}}^{\text{PLAT}} > \pi_0^{\text{PLAT}}$  provided  $N \geq 2$ . If  $N > 3$ , then  $\pi_{\text{CAP}}^{\text{PLAT}} > \pi_0^{\text{PLAT}}$  regardless of  $\theta$ .

a binding-cap regime. It can be shown, for the assumed parameter values, that

$$B_{\text{NOCAP}} = \frac{N\alpha^3}{128k^2} \text{ and } B_{\text{CAP}} = \frac{256}{27} \frac{N-1}{3N-1} B_{\text{NOCAP}}.$$

Welfare, including the cost of installing the bandwidth, is

$$W_{\text{NOCAP}} = \frac{5}{128} \frac{N\alpha^3}{k} \text{ and } W_{\text{CAP}} = \frac{256}{135} \frac{(5N-1)(N-1)}{(3N-1)^2} W_{\text{NOCAP}}.$$

It follows that if  $N < 11$ , then welfare is greater if the platform is barred from imposing a download cap, but if  $N > 11$ , then welfare is greater if the platform is permitted to impose a cap of its choosing (if  $N = 11$ , welfare is the same in the two regimes.)

## 7 ADVERTISING-SUPPORTED CONTENT AND SURPLUS EXTRACTION VIA HIGHER QUALITY

Considerable amount of content is provided for free, especially over the Internet, with the content providers' deriving their revenue from advertising. In this section, we briefly consider a model of advertising-supported content in which the content providers potentially respond to the platform's policies and pricing by varying the *quality* of their content.<sup>20</sup> Using a simple model, we demonstrate that the platform continues to derive a surplus-extraction benefit from both two-part tariffs and a binding-cap regime.

A preliminary observation: suppose a content provider provides its content for free, earning greater revenue from advertising the more content consumers obtain. Let  $x_n(\rho, q_n)$  be demand for the  $n$ th content provider's content when the quality of that content is  $q_n \in \mathbb{R}_+$  and consumers face an effective unit price of  $\rho$ . Given the content provider's price is zero,  $\rho = \tau$ , the per-unit component of the platform's two-part tariff. For an increase in price to increase the content provider's return to quality it must be that

$$\frac{\partial^2 x_n(\rho, q)}{\partial \rho \partial q} > 0; \tag{25}$$

that is, the higher the price the greater the marginal increase in demand from improving quality. As is well known from the study of quality provision (see, *e.g.*, Spence, 1975), (25) is a condition met by some demand curves, but not all. On the other hand, were it not met, then, as will become evident, the platform would have no incentive to use either a two-part tariff or a binding-cap regime; it finds those instruments valuable only if content demand satisfies (25).

A tractable analysis again requires functional-form assumptions. Specifically, suppose that the marginal utility a consumer derives from the  $x$ th unit of

---

<sup>20</sup>We thank seminar participants at the Paris School of Economics for encouraging us to explore this extension. An earlier version of this paper considered an alternative model of quality provision, which reached similar conclusions.

content provider  $n$ 's product is

$$(\alpha - x)q_n. \quad (26)$$

Note, solely for convenience, we now assume congestion is not an issue (*i.e.*,  $\delta(X|B) \equiv 1$ ). Having quality enter multiplicatively in (26), ensures that condition (25) will be met.

Assume that all content has a minimum quality level  $\underline{q} > 0$ . By expending  $c\Delta$ ,  $c > 0$  a constant, a content provider can raise its quality to  $q = \underline{q} + \Delta$ . Assume that a content provider's revenue if its demand is  $x_n(\tau, q_n)$  is  $Ax_n(\tau, q_n)$ , where  $A > 0$  is the rate at which content consumed translates into ad revenue (*e.g.*, a unit of content corresponds to a web page viewed and  $A$  equals the rate at which a viewed web page translates into a clicked-through ad times the payment for a click through). A content provider's profit is, thus,

$$Ax_n(\tau, \underline{q} + \Delta) - c\Delta. \quad (27)$$

Under a two-part tariff, a consumer chooses her quantity of the  $n$ th content provider's product so to equate marginal utility,  $(\alpha - x_n)q_n$ , with effective price,  $\tau$ . This yields demand

$$x_n(\tau, q_n) = \alpha - \frac{\tau}{q_n}$$

(observe it satisfies (25)). Note that if  $\tau = 0$ , then demand is no longer a function of quality; that is,  $x_n(0, q_n) = \alpha$ .

From this last observation, it follows that if the platform charges only a hookup fee and sets no download limit, then (27) becomes  $A\alpha - c\Delta$ , which in turn implies each content provider will choose the minimum level of quality,  $\underline{q}$ , in response to such platform policy and pricing. Consumer surplus (and thus the hookup fee) is then

$$H_0 = N\underline{q} \int_0^\alpha (\alpha - x)dx = N\underline{q} \frac{\alpha^2}{2}. \quad (28)$$

If the platform uses a two-part tariff with a positive per-unit fee, then (27) is

$$A \left( \alpha - \frac{\tau}{\underline{q} + \Delta} \right) - c\Delta.$$

A content provider chooses  $\Delta$  to maximize that expression. It is readily seen the profit-maximizing choice is

$$\Delta^* = \max \left\{ 0, \sqrt{\frac{A\tau}{c}} - \underline{q} \right\}. \quad (29)$$

For the moment, assume the second term is the larger, so  $\Delta^* > 0$ . Note this makes  $q_n = \sqrt{A\tau/c}$  for each  $n$ . The platform's corresponding profit (hookup fee plus revenue from per-unit charges) is

$$\pi_\tau^{\text{PLAT}} = N \sqrt{\frac{A\tau}{c}} \int_0^{\alpha - \sqrt{c\tau/A}} (\alpha - x)dx = \frac{N}{2} \sqrt{\frac{\tau}{cA}} (A\alpha^2 - c\tau). \quad (30)$$

The platform's choice of a per-unit fee maximizes (30). It is readily seen the solution to that maximization program is

$$\tau = \frac{A\alpha^2}{3c}. \quad (31)$$

The analysis is consistent provided  $\Delta^* > 0$ . It will be, using (31) to substitute for  $\tau$  in (29), if

$$\frac{A}{c} \frac{\alpha}{\sqrt{3}} > \underline{q}; \quad (32)$$

that is, if the value of advertising is high, or the the cost of quality improvement low, or both. Finally, observe that  $\pi_{\tau}^{\text{PLAT}}$ , evaluated at the profit-maximizing  $\tau$  (*i.e.*, the value in (31)), is

$$\pi_{\tau^*}^{\text{PLAT}} = \frac{NA\alpha^3}{3\sqrt{3}c}. \quad (33)$$

It is readily shown that expression (33) exceeds (28), which is profit given a zero per-unit fee and no download limit, if

$$\frac{A}{c} \frac{2\alpha}{3\sqrt{3}} > \underline{q}. \quad (34)$$

Observe condition (34) implies condition (32). In sum, the preceding analysis establishes:

**Proposition 11.** *Under the assumptions of this section and assuming the value of advertising is sufficiently high, or the cost of quality improvement sufficiently low, or both (specifically, that condition (34) holds), there is a two-part tariff, with a positive per-unit fee, that yields the platform higher profits than simply charging a hookup fee and no per-unit fee. Moreover, content quality is higher if the platform uses such a two-part tariff than if it just charges a hookup fee.*

Intuitively, content providers seek to offset the loss of demand that a positive per-unit charge causes by improving the quality of their content. This raises consumer benefit, despite the reduction in total content consumed. Because a two-part tariff allows the platform to capture that benefit, it follows the platform's profit is greater under a two-part tariff than were it simply to set a hookup fee.

Now consider a binding-cap regime. Maximizing utility subject to the constraint that content not exceed  $L$ , a consumer's first-order conditions are

$$q_n(\alpha - x_n) - \lambda = 0,$$

$n = 1, \dots, N$ , where  $\lambda$  is the Lagrange multiplier on the download constraint (*i.e.*,  $\sum_{n=1}^N x_n = L$ ). Each first-order condition can be solved for quantity demanded:

$$x_n = \alpha - \frac{\lambda}{q_n}.$$

Summing over  $n$ , then solving the download constraint for  $\lambda$ , then substituting that solution back into demand yields

$$x_n = \alpha - \frac{N\alpha - L}{1 + (\underline{q} + \Delta_n) \sum_{j \neq n} (1/q_j)}. \quad (35)$$

Content provider  $n$  chooses  $\Delta_n$  to maximize its profit,

$$Ax_n - c\Delta_n = A \left( \alpha - \frac{N\alpha - L}{1 + (\underline{q} + \Delta_n) \sum_{j \neq n} (1/q_j)} \right) - c\Delta_n. \quad (36)$$

It is readily seen that (36) is strictly concave in  $\Delta_n$  for all  $q_j$ ,  $j \neq n$ . Consequently, content provider  $n$ 's best-response level of quality solves the first-order condition:

$$\frac{A(N\alpha - L)}{\left(1 + (\underline{q} + \Delta_n) \sum_{j \neq n} (1/q_j)\right)^2} \sum_{j \neq n} \frac{1}{q_j} - c \leq 0,$$

with equality holding if  $\Delta_n > 0$ . Positing a symmetric interior equilibrium (*i.e.*,  $\Delta_1 = \dots = \Delta_N \equiv \Delta > 0$ ), the last expression entails

$$\Delta = \frac{A(N\alpha - L)(N - 1)}{cN^2} - \underline{q}. \quad (37)$$

The analysis is consistent provided the fraction in (37) exceeds  $\underline{q}$ . If it does, then that fraction is also the value of  $q_n$ ,  $n = 1, \dots, N$ . Given symmetry, it follows that  $x_1 = \dots = x_N = L/N$  (consider expression (35)). It further follows that the platform can set its hookup fee to be

$$\begin{aligned} H_{\text{CAP}} &= N \frac{A(N\alpha - L)(N - 1)}{cN^2} \int_0^{L/N} (\alpha - x) dx \\ &= \frac{AL(N - 1)(2N\alpha - L)(N\alpha - L)}{2cN^3}. \end{aligned} \quad (38)$$

Maximizing (38) with respect to  $L$  yields

$$L = N\alpha \frac{3 - \sqrt{3}}{3}. \quad (39)$$

As that expression is less than  $N\alpha$ , it follows the download constraint indeed binds. Substituting for  $L$  in (38) using (39), the profit-maximizing hookup (thus, profit) is

$$H_{\text{CAP}} = \frac{NA\alpha^3}{3\sqrt{3}c} \frac{N - 1}{N}. \quad (40)$$

Observe (40) exceeds (28) if

$$\frac{A}{c} \frac{2\alpha}{3\sqrt{3}} \frac{N - 1}{N} > \underline{q}. \quad (41)$$

If that inequality holds, then straightforward algebra reveals that the  $\Delta$  in (37), with  $L$  given by (39), is positive. In sum, the preceding analysis establishes:

**Proposition 12.** *Under the assumptions of this section and assuming the value of advertising is sufficiently high, or the cost of quality improvement sufficiently low, or both (specifically, that condition (41) holds), there is a binding-cap regime that yields the platform higher profits than simply charging a hookup fee. Moreover, content quality is higher if the platform employs such a binding cap than if it just charges a hookup fee.*

As with the analysis of paid content, the imposition of a binding cap effectively puts the content providers in competition with each other. With free content, content providers compete by raising their quality. As with Proposition 11, this increase in quality can raise consumer benefit by more than the loss due to less content acquisition, which means the platform can raise its hookup fee *vis-à-vis* what it would be were there no cap.

Keep in mind that the logic of Proposition 1 still applies: if the platform’s policies and pricing had no effect on the content providers’ behavior, then the platform would do strictly better by eschewing both a positive per-unit fee and a binding download cap; that is, optimally, it would simply set a hookup fee. In other words, at least with homogenous consumers, a platform’s only incentive to deviate from a simple hookup fee when content is free derives from the implications for such deviations on content providers’ quality decisions.

Comparing (33) and (40), the platform’s profit under a two-part tariff exceeds its profit under a binding-cap regime, but the two converge as the number of content providers gets large. In summary:

**Proposition 13.** *Under the assumptions of this section, a profit-maximizing platform makes greater profit using a two-part tariff than using a binding-cap regime. However, as the number of content providers gets large, this difference shrinks; in the limit, profits are the same under the two regimes. Hence, if there are significant transaction fees associated with administering a two-part tariff, then the platform would prefer a a binding-cap regime if the number of content providers is sufficiently large.*

Finally, what are the welfare implications of a platform setting a binding cap or utilizing a two-part tariff? Not surprisingly, the answer is “ambiguous.” For the sake of brevity, we do not pursue a full analysis of welfare here. Rather we limit ourselves to pointing out the relevant issues:

1. as modeled, the content providers do not internalize the benefits consumers get from higher quality and, thus, undersupply quality *vis-à-vis* the welfare-maximizing amount when the platform neither sets a binding cap nor charges a positive per-unit fee;
2. at the same time, in setting its policy and prices, the platform does not internalize the costs the content providers incur to raise quality—there is a danger that they will induce more than the welfare-maximizing level of quality;

3. furthermore, as seen, for any level of quality, a positive per-unit fee or a binding cap lead to less than the welfare-maximizing amounts of content trading; and
4. a complete welfare analysis would have to account for the benefits advertising provides the advertisers.

## 8 HETEROGENOUS CONSUMERS, CONTENT, AND PRICE DISCRIMINATION

To explore issues of price discrimination, we require consumer heterogeneity. Extending the model employed to this point to encompass heterogeneous consumers is not, however, practical. Consequently, in this section, we employ a somewhat simpler model. Although simpler, it does facilitate an analysis of heterogeneous *content* as well as consumers, which offers an additional benefit.

Assume there are two consumer types, 0 and 1. Let  $\phi_i \in (0, 1)$  be the proportion of consumers of type  $i$ ,  $i \in \{0, 1\}$ . Let the gross utility of a type- $i$  consumer be

$$\int_0^N \beta_i v(n, X) \chi_n dn + y,$$

where  $\beta_i \in (0, \infty)$ , with  $\beta_1 > \beta_0$ , and  $\chi_n \in \{0, 1\}$  indicating whether the consumer has acquired a unit of the  $n$ th content provider's content. Note, as in the Introduction, a consumer wants, at most, one unit of any content provider's content. Note, too, we are now assuming a measure  $N$  of content providers. The quantity  $\beta_i v(n, X)$  is the contribution to total consumer utility from a unit of the  $n$ th content provider's content given  $X$  units of total content are being transported. We again assume congestion is detrimental:  $v(n, \cdot)$  is a decreasing function for all  $n$ . We assume a fixed order of preference for the different content; specifically, assume that  $n > n'$  implies  $v(n, X) < v(n', X)$  for all  $X$ . Assume  $v(N, N) \geq 0$  (*i.e.*, even with maximum congestion, a consumer gains utility from her least preferred content). At the same time, we maintain the assumption that limits on congestion can be welfare enhancing: let the function

$$U(M) \equiv \int_0^M v(n, M) dn$$

be concave in  $M$  and assume  $U'(N) < 0$ . Assume that  $v(\cdot, \cdot)$  is continuous in both arguments.

As a preliminary, suppose  $\beta_1 = \beta_0$  (*i.e.*, there is just one type). In this case, there is no loss in setting  $\beta_1 = \beta_0 = 1$ . Absent any download limit or per-unit fee, all content providers will operate and content provider  $n$  sets a price of  $v(n, N)$ . There is no consumer surplus, so  $H_{\text{NOCAP}} = 0$ . Welfare is  $U(N)$ .

Suppose a download limit (cap) of  $L < N$  is imposed while the per-unit fee,  $\tau$ , remains at zero. The following is an equilibrium, as is readily verified:

- All content providers  $n \geq L$  set their prices to 0.

- A content provider  $n$ ,  $n < L$ , sets its price to  $v(n, L) - v(L, L)$ .

Consumer surplus is  $Lv(L, L)$ ; hence,  $H_{\text{CAP}} = Lv(L, L)$ .

Alternatively, suppose no cap, but a per-unit fee of  $\tau > 0$  is set. There are two cases: (i)  $\tau < v(N, N)$  and (ii)  $\tau \in [v(N, N), v(0, 0))$ . In the first case, content provider  $n$  sets its price as  $p_n = V(n, N) - \tau$ . In the second, let  $T$  solve  $\tau = V(T, T)$  (a solution exists because  $V(n, n)$  is continuous in  $n$ ). As is readily verified, the following is an equilibrium of the second case:

- All content providers  $n \geq T$  set their prices to 0.
- A content provider  $n$ ,  $n < T$ , sets its price to  $v(n, T) - \tau = v(n, T) - v(T, T)$ .

In either case, consumer surplus is 0; hence,  $H = 0$ . The platform's profit is  $\tau N$  in the first case and  $\tau T = Tv(T, T)$  in the second. In the first case, the platform does best to set  $\tau = v(N, N)$ . Hence, the first case is a special instance of the second case; that is, platform profit is  $Tv(T, T)$ .

Observe that regardless of whether it has employed a binding cap or a per-unit fee, the platform's profit is an expression of the form  $Mv(M, M)$ . Hence, it is irrelevant whether the platform imposes a binding cap or uses a per-unit fee.

The platform will set  $M$  to maximize profit,  $Mv(M, M)$ . The first-order condition is

$$v(M, M) + M \left( \frac{\partial v(M, M)}{\partial n} + \frac{\partial v(M, M)}{\partial X} \right) = 0. \quad (42)$$

In contrast, welfare maximization entails setting  $M$  to solve

$$v(M, M) + \int_0^M \frac{\partial v(n, M)}{\partial X} dn = 0. \quad (43)$$

Clearly, expressions (42) and (43) are different, from which it follows that the platform's decision will, generically at least, not be welfare maximizing. Further examination of these expressions reveals that a familiar tension exists: the monopolist (the platform) is concerned with marginal values, while welfare depends on infra-marginal values (similar in logic to Spence, 1975). In particular, we have the following result.

**Proposition 14.** *Under the assumptions of this section and assuming homogeneous consumers, if the marginal disutility of congestion is not less for content of lower value than higher value (i.e., if  $\partial(\partial v/\partial X)/\partial n \leq 0$ ), then the platform will set a download cap that is below the welfare-maximizing cap; alternatively, a per-unit fee above the welfare-maximizing level.*

**Proof:** By the intermediate value theorem, the lefthand side of (43) equals

$$v(M, M) + M \frac{\partial v(\hat{n}, M)}{\partial X} \quad (44)$$

for some  $\hat{n} \in (0, M)$ . By assumption, (44) is not less than

$$v(M, M) + M \frac{\partial v(M, M)}{\partial X},$$

which in turn strictly exceeds (42). The result follows.  $\blacksquare$

For example, if  $v(n, X) = K - nX$ ,  $K \geq N^2$ , then the platform will impose a cap of  $\sqrt{K/3}$  (alternatively, a per-unit fee of  $2K/3$ ), while welfare maximization entails a cap of  $\sqrt{2K/3}$  (alternatively, a per-unit fee of  $K/3$ ). Indeed, as long as the platform wishes to limit traffic (*i.e.*, whenever  $N > \sqrt{K/3}$ ), welfare with no cap (alternatively, a per-unit fee below  $K - N^2$ ) exceeds welfare under the platform's profit-maximizing action:

**Proposition 15.** *Under this section's assumptions, with homogenous consumers and assuming  $v(n, X) = K - nX$ ,  $K \in [N^2, 3N^2)$ , welfare is greater with no download cap or a zero per-unit fee than with the cap or per-unit fee the platform would choose.*

**Proof:**

$$U\left(\sqrt{\frac{K}{3}}\right) = \frac{5K^{3/2}}{6\sqrt{3}} < \frac{K^{3/2}}{2} = U(\sqrt{K}).$$

The derivative of  $U(n)$  evaluated at  $n = \sqrt{K/3}$  is  $K/2 > 0$  and  $U(\cdot)$  is a concave function. Hence,  $U(n) > U(\sqrt{K/3})$  for all  $n \in (\sqrt{K/3}, N]$ .  $\blacksquare$

Proposition 14 assumed the disutility from congestion was greater for less desired content than more desired content. It is, of course, possible the opposite is true (*e.g.*, movies on demand could be both highly valued and marginal disutility from congestion high relative to material from online periodicals). Even then, the platform could wish to limit traffic below the welfare-maximizing amount because of the direct effect of inducing the content providers to cut their prices (*i.e.*, reflecting the  $\partial v(M, M)/\partial n$  term in (42)). To illustrate this, suppose

$$v(n, X) = \frac{1}{\sqrt{X}}(K - n), \quad (45)$$

where  $K \geq N$ . Observe  $\partial(\partial v/\partial X)/\partial n > 0$ . It is readily see that

$$U(M) = K\sqrt{M} - \frac{M^{3/2}}{2}.$$

The welfare-maximizing traffic limit (whether imposed by cap or per-unit fee) is, therefore,  $M_W^* = 2K/3$ . Solving (42) yields  $M_{\text{PLAT}}^* = K/3$ .

**Proposition 16.** *Under this section's assumptions, with homogenous consumers and assuming  $v(n, X)$  is given by (45), welfare is greater with no download cap or a zero per-unit fee than with the cap or per-unit fee the platform would choose.*

**Proof:**

$$U\left(\frac{K}{3}\right) = \frac{5K^{3/2}}{6\sqrt{3}} < \frac{K^{3/2}}{2} = U(K).$$

The derivative of  $U(n)$  evaluated at  $n = K/3$  is  $\sqrt{3K}/4 > 0$  and  $U(\cdot)$  is concave. Hence,  $U(n) > U(K/3)$  for all  $n \in (K/3, N]$  (recall  $N \leq K$ ). ■

Return, now, to the assumption of two consumer types. Suppose, initially, that there are no download caps nor any per-unit fees. Content provider  $n$ , if it anticipates both types acquiring access and a total volume of  $X$ , has two candidate monopoly prices:  $\beta_0 v(n, X)$ , which is the maximum price at which it can sell to all consumers; or  $\beta_1 v(n, X)$ , which is the maximum price at which it has any custom. The former is more profitable than the latter if and only if

$$\beta_0 \geq \phi_1 \beta_1. \quad (46)$$

If (46) holds, then the type-0 consumers obtain no surplus; hence, if the platform sets an access fee that induces all to acquire access, it must be that  $H = 0$ : the platform earns no profit. If (46) fails to hold or type-0 consumers don't acquire access, then content provider  $n$ 's price is  $\beta_1 v(n, X)$  and type-1 consumers obtain no surplus (nor would a type-0 consumer who deviated by obtaining access). It again follows that the platform's profit is zero. In short, the situation is analogous to that in the Introduction: a platform that simply sets an access fee can earn zero profit.

**Proposition 17.** *Absent download caps or per-unit fees, the platform earns zero profit given the assumptions of this section.*

Going forward, we limit attention to the case in which (46) holds. Doing so avoids having to calculate mixed-strategy equilibria for certain pricing games played by the content providers.<sup>21</sup>

**Assumption 1.** *The value the low-type consumers (type 0) place on content is not significantly less than the value high-type consumers (type 1) place on content, or the proportion of high-type consumers is sufficiently small, or both; specifically, condition (46) holds.*

Suppose the platform doesn't discriminate, but offers a common package (either with a download cap or a per-unit fee). Because, as the previous analysis demonstrates, the platform can replicate an outcome with a per-unit fee via a download cap (and *vice versa*), there is no loss in limiting attention to a cap,  $L < N$ .

**Proposition 18.** *Maintain the assumptions of this section, including Assumption 1. Suppose the platform offers a common package with hookup fee  $H$  and download cap  $L < N$ . Then the two possible equilibria are*

<sup>21</sup>For instance, if (46) didn't hold, then neither would expression (48) *infra*; in this case, no pure-strategy equilibrium would exist in the pricing subgame should the platform offer a single package with a download cap.

(i)  $L = \operatorname{argmax}_M M\beta_0 v(M, M)$  and  $\pi_{\text{PLAT}} = \max_M M\beta_0 v(M, M)$  if

$$\max_M M\beta_0 v(M, M) \geq \max_M \phi_1 M\beta_1 v(M, \phi M); \quad (47)$$

(ii) and  $L = \operatorname{argmax}_M M\beta_1 v(M, \phi_1 M)$  and  $\pi_{\text{PLAT}} = \phi_1 \times \max_M M\beta_1 v(M, \phi_1 M)$  if (47) does not hold.

**Proof:** Fix  $L$  and suppose the content providers expect all consumers to acquire access. The following is an equilibrium of their pricing subgame:  $p_n = 0$  if  $n > L$  and  $p_n = \beta_0(v(n, L) - v(L, L))$  if  $n \leq L$ . To see this, observe that content provider  $n$ ,  $n > L$ , is clearly playing a (weak) best response. For  $n \leq L$ , there is no gain to lowering price. If it raises price to  $\tilde{p}$ , it will lose type-0 consumers, but retain type-1 consumers provided

$$\beta_1 v(n, L) - \tilde{p}_n \geq \beta_1 v(L, L)$$

(i.e., as long as a type-1 consumer's surplus from buying from content provider  $n$  exceeds her next-best option of buying from an  $n > L$ ). The best such deviation is  $\tilde{p} = \beta_1(v(n, L) - v(L, L))$ , which yields profit

$$\phi_1 \beta_1 (v(n, L) - v(L, L)) \leq \beta_0 (v(n, L) - v(L, L)) \quad (48)$$

(the inequality is due to (46)). Given the expression to the right of the inequality is the content provider's profit if it prices at  $p_n$ , it follows that no content provider  $n$ ,  $n \leq L$ , will wish to deviate by raising its price. Given this equilibrium of the pricing subgame, if the platform will indeed induce access from all consumers, the maximum hookup fee it can charge equals a type-0 consumer's surplus; that is,  $L\beta_0 v(L, L)$ . It follows that  $H = \pi_{\text{PLAT}} = \max_M \beta_0 Mv(M, M)$  and  $L = \operatorname{argmax}_M Mv(M, M)$ .

Alternatively, fix  $L$  and suppose the content providers expect only type-1 consumers to acquire access. The pricing game is analogous to that for homogeneous consumers considered above; in particular, the equilibrium prices are  $p_n = 0$  if  $n > L$  and  $p_n = \beta_1(v(n, \phi_1 L) - v(L, \phi_1 L))$  if  $n \leq L$ . Given this equilibrium of the pricing subgame, if the platform will indeed induce access from type-1 consumers only, the maximum hookup fee it can charge equals such a consumer's surplus; that is,  $L\beta_1 v(L, \phi_1 L)$ . It follows that  $L = \operatorname{argmax}_M Mv(M, \phi_1 M)$  and  $H = \max_M Mv(M, \phi_1 M) = \pi_{\text{PLAT}}/\phi_1$ . For this alternative, the remaining thing to verify is that a type-0 consumer won't deviate by purchasing access. To see she won't, note that any content-purchasing strategy she would pursue a type-1 consumer could also pursue. The type-1's surplus will be greater; hence, given a type-1 consumer is indifferent between acquiring access and not, knowing she will purchase content optimally, it must be that a type-0 consumer would do strictly worse by acquiring access than not.

The platform prefers the first alternative (access for all) if (47) holds and it prefers the second (access for type-1 consumers only) otherwise.  $\blacksquare$

Next, suppose the platform offers two packages, one intended for low-type consumers, the other for high-type consumers. A complete characterization of equilibria is difficult because the content providers will play mixed strategies in the ensuing pricing subgames. Fortunately, we can establish one relevant result, which is conditions exist such that the platform has no incentive to offer a package without restrictions (*i.e.*, an “all-you-can-eat” package). To facilitate a tractable analysis, assume that congestion is no longer an issue; that is, a type- $i$  consumer’s benefit from a unit of content provider  $n$ ’s content is  $\beta_i v(n)$  regardless of the total amount of traffic the platform carries. Continue to assume  $v(\cdot)$  is a decreasing function; however, the rate of decrease could be limited. In particular, assume

$$\max_M Mv(M) > \phi_0 \int_0^N v(n)dn. \quad (49)$$

Observe (49) would hold if  $v(\cdot)$  were constant given  $\phi_0 < 1$ .

Suppose the platform offers a package with a cap,  $L$ , and hookup fee  $H_0$  and another package with no cap and hookup fee  $H_1$ . The usual logic of second-degree price discrimination applies here, namely that, for a pricing scheme to be incentive compatible, the high type must be allocated more than the low type; hence, the high types (type 1) purchase the second package. Because high types are unrestricted in their downloads, content provider  $n$  can always sell to them at price  $\beta_1 v(n)$ ; that is, content provider  $n$  can guarantee itself a profit of  $\phi_1 \beta_1 v(n)$ . It follows that no  $n$  will charge a price less than  $\phi_1 \beta_1 v(n)$ . Consequently, the surplus *left* to the platform and consumers cannot exceed

$$\underbrace{\int_0^N (\phi_0 \beta_0 + \phi_1 \beta_1) v(n) dn}_{\text{max. welfare}} - \underbrace{\int_0^N \phi_1 \beta_1 v(n) dn}_{\text{min. profit of content providers}} = \phi_0 \beta_0 \int_0^N v(n) dn. \quad (50)$$

The platform’s profit under this price-discrimination scheme—call it  $\hat{\pi}$ —cannot exceed surplus, the righthand side of (50). Expressions (49) and (50) imply that  $\hat{\pi}$  is less than the maximum profit the platform could achieve by offering a single package (Proposition 18).<sup>22</sup> It follows that the platform would never offer two packages, one of which has no cap (permits “all you can eat”). In sum:

**Proposition 19.** *Maintain the assumptions of this section, including Assumption 1. In addition assume the value function,  $v(\cdot)$ , is relatively flat, or the proportion of type-0 consumers not too great, or both (specifically, that condition (49) holds). Then the platform will never offer a package with unlimited downloads even if engaged in second-degree price discrimination.*

Note that Proposition 19 does not rule out price discrimination by the platform, it only rules out all-you-can-eat packages.

<sup>22</sup>If congestion is not an issue, then (47) holds given (46).

## 9 DISCUSSION

As with any model, we have had to abstract away from certain questions. We address some of them here, before concluding, both for the sake of completeness and to indicate avenues for future research.

One question is why do platforms seem to avoid “naked” two-part tariffs; that is, why disguise the tariff behind the veil of a permeable limit?<sup>23</sup> We cannot answer the question definitively—and it may well be a topic for future research—but we can speculate. A possible answer is that consumers cannot easily meter the megabytes they download. Hence, when faced with a per-unit fee, there is an element of financial risk. This could either lead to diminished downloads or require the platform to compensate consumers for the risk via a reduced hookup fee, both of which lessen platform profit. A second possibility is that having initially offered plans with only a monthly access fee (as many residential Internet service providers did), platforms find it difficult to impose per-unit fees on every unit of content because that risks alienating consumers; permeable caps could be a more palatable approach from a marketing perspective.<sup>24</sup> A potential explanation for why platforms priced as they did originally is that they lacked sufficient experience to design multipart tariffs (the requisite consumer data had not yet been accumulated). There are likely other explanations that involve marketing and consumer psychology, but it is not our comparative advantage to discuss them and we, thus, leave those to others.

Additionally, because, as shown above, *impermeable* caps do nearly as well as two-part tariffs—at least when the number of content providers is large—it could be that permeable caps are intended to be essentially binding. Their permeability could be to reduce administrative costs (*e.g.*, having to cope with angry customers who have suddenly been cut off mid-movie). In other words, the overage charge is set to deter consumers from exceeding the limit, not to extract revenue, but is finite to limit hassles.

Another set of questions have to do with the time horizon of caps, which are typically set on a monthly basis (in terms of the modeling above, this suggests that the unit of time should be seen as a month). Why not daily or weekly caps? And why should caps apply at all times and not just, for instance, at highly congested (peak) times? As the second question indicates, network congestion is generally not constant throughout the month. Nonetheless, one could justify our model by imagining consumers make monthly usage decisions and the function  $\delta$  reflects some average congestion disutility that they expect given the usage

---

<sup>23</sup>As noted earlier, there are exceptions. For example, the UK mobile-telephone provider Three offers plans that are straightforward two-part tariffs.

<sup>24</sup>There could also be political-economy issues: given that these platforms are regulated, upsetting consumers—voters—could lead to unwanted regulatory responses. This is not necessarily idle speculation: see “Congressman Wants to Ban Download Caps,” Ryan Singal, *Wired*, April 14, 2009 (accessed online September 27, 2014 at URL <http://www.wired.com/2009/04/congressman-to/>). The Representative in question was reacting to a proposed permeable-cap plan by Time Warner Cable with a very low cap, one gigabyte, and a \$2 per GB overage charge.

pattern they've chosen. Alternatively, it could be that consumers are effectively compelled to consume most content at peak hours (*e.g.*, intensive residential Internet usage can occur only in the evenings after work). That is, off-peak usage is necessarily limited and minor (*e.g.*, a quick check of email before heading to work in the morning); so congested periods (peak hours) are effectively all that matter. On the other hand, the question of why platforms don't have limits or prices that vary by time of day remains. An earlier version of this paper contained an analysis, with a very rudimentary model, that found a platform makes greater profits with an "all-the-time" (*e.g.*, monthly) cap than with a time-of-day cap (details available from authors upon request), but we would be the first to admit further exploration is warranted. In addition, there remains the question of why caps do not reset on a daily or weekly basis. The answer may have to do with the typical monthly billing cycle, but given (i) cycle length is endogenous;<sup>25</sup> and (ii) the link between the timing of billing and cap reset not obvious, a monthly billing cycle is, at best, an incomplete answer.

Under a policy regime broadly known as network neutrality, residential Internet service providers (ISPs) cannot charge content providers for so-called "last-mile" delivery of their content to households. As has been noted, given the nature of tax incidence, it is irrelevant if the platform imposes a per-unit fee directly on consumers or content providers. Hence, the analysis above speaks directly to the network neutrality debate; in particular, results such as Corollary 1 and Propositions 6, 9, and 14 are relevant for assessing the welfare effects of relaxing network neutrality. Moreover, the analysis suggests that US ISPs' recent interests in imposing download caps (permeable or not) could be a response to their failure to have network neutrality repealed: as noted, caps are an alternative to directly charging the content providers.

Price discrimination, particularly by ISPs, could and has taken the form of quality distortions; specifically, variations in speed of service. Some mobile-phone networks also offer access with different speeds. The analysis above abstracts from such quality issues. How a platform's ability to engage in second-degree price discrimination via quality distortion (versioning) interacts with its use of caps is, therefore, a topic for further research. We note, in passing, that slow download speeds could be similar to download caps insofar as they cause consumers to see different content providers as substitutes (*e.g.*, one might be willing to wait a long time for content from provider A or B but not the additional time to receive content from both). This too is an issue for future research.

Another quality-related issue is the following: download caps could also affect aspects of content providers' operations in addition to their pricing. For instance, it could cause them to change their compression algorithms, which could affect consumers' perception of the content's quality.<sup>26</sup> Or it could bias

---

<sup>25</sup>In this regard, it is worth noting that some platforms have pay-as-you-go or refillable-bucket plans; that is, payment is not monthly.

<sup>26</sup>According to the article "Netflix says Internet download caps only in place to drive up bills," *thespec.com*, March 29, 2011 (accessed August 15, 2013 at

consumers toward purchasing lower-quality content (*e.g.*, standard-definition video) rather than higher-quality (*e.g.*, high-definition video). In turn, if caps lead to an increase of lower-quality content consumption over higher-content consumption, then this will affect consumer surplus and, thus, the hookup fees platforms can charge. Not only could caps affect the nature of the content offered, but it could also affect the content providers' revenue models: perhaps affecting whether they rely on charging consumers or being advertising supported or the mix of the two. This will also affect the consumer experience, with consequences for what the platforms can charge consumers. All of these extensions are, however, beyond the scope of the current paper and remain work for the future.

Our analysis has been limited to a monopoly platform. Although platform competition in the relevant markets is often limited, it does exist and the consequences of oligopolistic competition among platforms on the use of download caps should be explored. A related point is that our model assumes consumers use a single platform through which to reach content providers. Increasingly, consumers "multi-home," accessing content via multiple devices (*e.g.*, home computer and smartphone) that are connected to different platforms (*e.g.*, a residential ISP and a mobile-phone company). Among the issues worth exploring with multiple platforms is the extent to which the platforms are tempted to free ride on each other: because platform A benefits from the reduction in content prices induced by platform B's download caps, A might be tempted to offer less stringent caps to gain a competitive advantage *vis-à-vis* B. On the other hand, if, as a consequence, A has the lion's share of the consumers, the effect of B's caps on content providers' pricing could be negligible.

Another competitive issue arises when the platform is also a content provider. For instance, a cable TV company could provide broadband Internet and compete directly with Internet-based purveyors of on-demand movies. A download limit could harm such rivals, but because it also makes them fiercer competitors, it could lower the cable company's profits from selling its own on-demand movies.

Finally, future work could seek to quantify the size of the effects considered above empirically. In this regard, comparing the US with Canadian residential Internet markets could be instructive. The former is currently characterized by few download caps, while download caps are prevalent in the latter. Van Gorp and Middleton (2010) report evidence that Canadian access fees tend to exceed those of other OECD countries, which is consistent with the analysis above. At the same time, one would also like to see evidence that content providers charged lower prices in the Canadian market than the US market. Although some anecdotal evidence exists (*e.g.*, accounting for exchange rates, Netflix currently charges less in Canada than the US), we are unaware of any

---

URL <http://www.thespec.com/news-story/2202174-netflix-says-internet-download-caps-only-in-place-to-drive-up-bills/>), "Netflix is responding to the caps on downloading in Canada by crunching down the data used to stream video content ... [rather than] the best available quality, which can sometimes take two gigabytes per hour to stream, ... the new default setting—at a lesser quality—will consume only 300 megabytes an hour."

careful and systematic study of such differences.

To conclude, this paper has considered how a platform, such as a residential ISP or a mobile-telephone company, can profit from the imposition of download caps on its customers even absent motives of price discrimination and congestion alleviation. This is because such caps put downward pressures on content providers' pricing, so they charge consumers less. Because, *ceteris paribus*, consumers would realize greater surplus, the platform can raise its access (hookup) charge, thereby increasing its profit. This downward pressure arises from one of two sources: if the caps are binding, then they serve to induce (greater) competition among the content providers for consumers' limited download capacity; if the caps are permeable, then they act as disguised two-part tariffs and the per-unit portion of the tariff acts as an excise tax that leads the content providers to cut their prices. We also showed a similar effect can exist even with free content: to protect their advertising revenue, the content providers respond to these pressures by raising the quality of their product, thus increasing consumer surplus and, thus, ultimately the platform's hookup fee.

Beyond demonstrating such effects, we have shown that the platform can have incentives to set policies such that, even if there were welfare benefits to be had from congestion alleviation, welfare would be higher were the platform prohibited from using such practices (*i.e.*, required to allow unlimited downloads and set a zero per-unit fee). The implications of these effects for the platform's capacity (bandwidth) decisions were also considered. Somewhat paradoxically, it was shown that allowing the platform to engage in such practices increases its incentives to expand capacity.

Most of the analysis concerned homogenous consumers and symmetrically positioned content providers. We extended the analysis to allow for heterogeneity on both fronts. Although tractability limited us to a less-general model than used for most our analysis, our results indicate that our conclusions are robust to heterogeneity across consumers and content providers.

## APPENDIX A: PROOFS NOT GIVEN IN TEXT

**Proof of Lemma 1:** Suppose, in equilibrium, each consumer's bundle satisfies  $\sum_{n=1}^N x_n < L$ . Recall consumers behave the same, so each consumer's equilibrium utility is

$$\begin{aligned} & \sum_{n=1}^{\infty} \int_0^{x_n} (\alpha - x\delta(X|B)) dx - H - \sum_{n=1}^N (p_n + \tau_b)x_n \\ & \equiv \sum_{n=1}^{\infty} \int_0^{x_n} (\alpha - x\delta(X|B)) dx - H - \sum_{n=1}^N p_n x_n - \tau_b \min \left\{ \sum_{n=1}^N x_n, \infty \right\} \\ & \quad - 0 \times \max \left\{ 0, -\infty + \sum_{n=1}^N x_n \right\}; \quad (51) \end{aligned}$$

hence, holding the content providers' prices fixed, each consumer would make the same decision facing no download limit and the two-part tariff  $\tilde{H} = H$  and  $\tau = \tau_b$  as she does given  $L, H, \tau_b, \tau_o$ . Because  $\sum_{n=1}^N x_n < L$ , the lefthand side of (51) is equivalent to the consumer's objective function under the original policy and prices for bundles in a neighborhood of  $(x_1, \dots, x_N)$ . Hence, the equilibrium bundle must satisfy the local first-order conditions:

$$\alpha - x_n \delta(X|B) - p_n - \tau_b = 0,$$

$n = 1, \dots, N$ . From which we obtain a local demand function for each content provider:<sup>27</sup>

$$x_n = \frac{1}{\delta(X|B)} (\alpha - \tau_b - p_n). \quad (52)$$

Because  $\sum_{n=1}^N x_n < L$ , holding fixed the prices of the content providers other than  $n$ , expression (52) is content provider  $n$ 's demand function locally under the platform's original policy and prices; that is, it is content provider  $n$ 's demand for prices in some interval  $(\underline{p}, \bar{p})$ ,  $\underline{p} < p_n < \bar{p}$ . Because  $p_n$  is content provider  $n$ 's equilibrium price, it must be that

$$p_n \in \operatorname{argmax}_{p \in (\underline{p}, \bar{p})} p \frac{1}{\delta(X|B)} (\alpha - \tau_b - p) \quad (53)$$

(*i.e.*,  $p_n$  must maximize profits locally). As a function of  $p$ , the expression on the righthand side of (53) is strictly concave; hence,  $p_n$  is the unique maximizer. Moreover, because  $p_n$  is superior both to  $p'$  and  $p''$ ,  $p' \in (\underline{p}, p_n)$  and  $p'' \in (p_n, \bar{p})$ , it follows that  $p_n$  maximizes the righthand side of (53) globally. Consequently,  $p_n$  would remain content provider  $n$ 's best response should the platform set no cap and use the two-part tariff  $\tilde{H} = H$  and  $\tau = \tau_b$  (that is, when the lefthand side of (51) is globally the consumer's objective function). This establishes that, if  $\sum_{n=1}^N x_n < L$ , the two-part tariff  $\tilde{H} = H$  and  $\tau = \tau_b$  with no cap induces the same equilibrium pricing by the content providers and the same consumption decisions by the consumers as did the original platform policy and prices. It is readily verified that the platform's profits are also the same.

Suppose, in equilibrium under the original policy and prices, that  $\sum_{n=1}^N x_n > L$ . Each consumer's equilibrium utility is

$$\begin{aligned} & \sum_{n=1}^{\infty} \int_0^{x_n} (\alpha - x \delta(X|B)) dx - \underbrace{(H + \tau_b L - \tau_o L)}_{\tilde{H}} - \sum_{n=1}^N (p_n + \tau_o) x_n \\ & \equiv \sum_{n=1}^{\infty} \int_0^{x_n} (\alpha - x \delta(X|B)) dx - \tilde{H} - \sum_{n=1}^N p_n x_n - \tau_o \min \left\{ \sum_{n=1}^N x_n, \infty \right\} \\ & \quad - 0 \times \max \left\{ 0, -\infty + \sum_{n=1}^N x_n \right\}; \quad (54) \end{aligned}$$

<sup>27</sup>Recall there is a measure one of consumers.

hence, holding the content providers' prices fixed, each consumer would make the same decision facing no download limit and the two-part tariff  $\tilde{H} = H + \tau_b L - \tau_o L$  and  $\tau = \tau_o$  as she does given  $L, H, \tau_b, \tau_o$ . Because  $\sum_{n=1}^N x_n > L$ , the lefthand side of (54) is equivalent to the consumer's objective function under the original policy and prices for bundles in a neighborhood of  $(x_1, \dots, x_N)$ . Hence, the equilibrium bundle must satisfy the local first-order conditions:

$$\alpha - x_n \delta(X|B) - p_n - \tau_o = 0,$$

$n = 1, \dots, N$ . From which we obtain a local demand function for each content provider:

$$x_n = \frac{1}{\delta(X|B)} (\alpha - \tau_o - p_n). \quad (55)$$

Because  $\sum_{n=1}^N x_n > L$ , holding fixed the prices of the content providers other than  $n$ , expression (55) is content provider  $n$ 's demand function locally under the platform's original policy and prices; that is, it is content provider  $n$ 's demand for prices in some interval  $(\underline{p}, \bar{p})$ ,  $\underline{p} < p_n < \bar{p}$ . Identical reasoning to above establishes that  $p_n$  must remain content provider  $n$ 's best response should the platform set no cap and use the two-part tariff  $\tilde{H} = H + \tau_b L - \tau_o L$  and  $\tau = \tau_o$  (that is, when the lefthand side of (54) is globally the consumer's objective function). This establishes that, if  $\sum_{n=1}^N x_n > L$ , the two-part tariff  $\tilde{H} = H + \tau_b L - \tau_o L$  and  $\tau = \tau_o$  with no cap induces the same equilibrium pricing by the content providers and the same consumption decisions by the consumers as did the original platform policy and prices. It is readily verified that the platform's profits are also the same. Part (i) of the lemma is proved.

Turning to part (ii): in this case, consumers view the cap as a binding constraint. Raising  $\tau_o$  to infinity doesn't change this. So a consumer's objective is equivalent to

$$\max_{(x_1, \dots, x_N)} \sum_{n=1}^N \int_0^{x_n} (\alpha - x \delta(X|B)) dx - \underbrace{(H + \tau_b L)}_{\tilde{H}} - \sum_{n=1}^N p_n x_n$$

subject to

$$\sum_{n=1}^N x_n = L. \quad (56)$$

The first-order conditions are

$$\alpha - x_n \delta(X|B) - p_n - \lambda = 0, \quad (57)$$

$n = 1, \dots, N$ , where  $\lambda$  is the Lagrange multiplier on (56). Solving the system (56) and (57) yields

$$x_n = \frac{L}{N} + \frac{\sum_{j \neq n} p_j - (N-1)p_n}{N \delta(X|B)}, \quad (58)$$

$n = 1, \dots, N$ . Observe that (58) is content provider  $n$ 's demand, which, although dependent on the prices of the *other* content providers, is independent of the hookup fee,  $H$ , and the per-unit fee,  $\tau_b$ . Hence, the pricing subgame among the content providers is the same under the platform's original policy and prices and under the alternative of a binding-cap regime ( $\tilde{\tau}_o = \infty$ ) with no per-unit fee (*i.e.*,  $\tilde{\tau}_b = 0$ ) and a hookup fee of  $\tilde{H} = H + \tau_b L$ . Equilibrium content prices are, thus, unchanged and, correspondingly, so to are the consumers' purchase decisions. The platform's profit remains  $H + \tau_b L$ . Part (ii) is proved. ■

## REFERENCES

- Cheng, H. Kenneth, Subhajyoti Bandyopadhyay, and Hong Guo**, "The Debate on Net Neutrality: A Policy Perspective," *Information Systems Research*, March 2011, *22* (1), 1–27.
- Choi, Jay Pil and Byung-Cheol Kim**, "Net Neutrality and Investment Incentives," *RAND Journal of Economics*, Autumn 2010, *41* (3), 446–471.
- , **Doh-Shin Jeon, and Byung-Cheol Kim**, "Net Neutrality, Business Models, and Internet Interconnection," 2013. Working Paper, Georgia Tech.
- Dai, Wei and Scott Jordan**, "Design and Impact of Data Caps," in "IEEE Global Communications Conference (Globecom)" Atlanta, GA December 2013.
- and – , "How do ISP Data Caps Affect Subscribers?," in "Research Conference on Communication, Information and Internet Policy" Arlington, VA September 2013.
- Dukes, Anthony and Esther Gal-Or**, "Negotiations and Exclusivity Contracts for Advertising," *Marketing Science*, May 2003, *22* (2), 222–245.
- Economides, Nicholas and Benjamin E. Hermalin**, "The Economics of Network Neutrality," *RAND Journal of Economics*, Winter 2012, *41* (4), 602–629.
- and **Joacim Tåg**, "Net Neutrality on the Internet: A Two-sided Market Analysis," *Information Economics and Policy*, 2012, *24*, 91–104.
- Greenstein, Shane and Ryan C. McDevitt**, "The Broadband Bonus: Estimating Broadband Internet Value," *Telecommunications Policy*, August 2011, *35* (7), 617–632.
- Hermalin, Benjamin E. and Michael L. Katz**, "The Economics of Product-Line Restrictions with an Application to the Network Neutrality Debate," *Information Economics and Policy*, 2007, *19*, 215–248.
- Higginbotham, Stacey**, "Which ISPs are Capping Your Broadband, and Why?," 2012. <http://gigaom.com/2012/10/01/data-caps-chart/>.

- 
- Krämer, Jan and Lukas Wiewiorra**, “Network Neutrality and Congestion-Sensitive Content Providers: Implications for Service Innovation, Broadband Investment, and Regulation,” *Information Systems Research*, December 2012, *23* (4), 1303–1321.
- Rochet, Jean-Charles and Jean Tirole**, “Two-Sided Markets: A Progress Report,” *RAND Journal of Economics*, Autumn 2006, *37* (3), 645–667.
- Roson, Roberto**, “Two-Sided Markets: A Tentative Survey,” *Review of Network Economics*, June 2005, *4* (2), 142–160.
- Rysman, Marc**, “The Economics of Two-Sided Markets,” *Journal of Economic Perspectives*, 2009, *23*, 125–144.
- Spence, A. Michael**, “Monopoly, Quality, and Regulation,” *Bell Journal of Economics*, Autumn 1975, *6* (2), 417–429.
- Van Gorp, Annemijn F. and Catherine A. Middleton**, “The Impact of Facilities and Service-Based Competition on Internet Services Provision in the Canadian Broadband Market,” *Telematics and Informatics*, 2010, *27*, 217–230.