Economic Viability of a Virtual ISP

Liang Zheng*, Carlee Joe-Wong†, Jiasi Chen‡, Christopher G. Brinton§, Chee Wei Tan¶, and Mung Chiang*
*Princeton University, †Carnegie Mellon University, ‡University of California, Riverside,
§Zoomi Inc., ¶City University of Hong Kong
*{liangz,chiangm}@princeton.edu, †cjoewong@andrew.cmu.edu, ‡jiasi@cs.ucr.edu,
§christopher.brinton@zoomiinc.com, ¶cheewtan@cityu.edu.hk

***Abstract*—Growing mobile data usage has led to end users paying substantial data costs, while Internet service providers (ISPs) struggle to upgrade their networks to keep up with demand and maintain high quality-of-service (QoS). This problem is particularly severe for smaller ISPs with less capital. Instead of simply upgrading their network infrastructure, ISPs can pool their networks to provide a good QoS and attract more users. Such a *vISP* (virtual ISP), for example, Google's Project Fi, allows users to access any of its partner ISPs' networks. We provide the first systematic analysis of a vISP's economic impact, showing that the vISP provides a viable solution for smaller ISPs attempting to attract more users, but may not maintain a positive profit if users' data demands evolve. To do so, we consider users' decisions of whether to defect from their current ISP to the vISP, as well as ISPs' decisions on whether to partner with the vISP. We derive the vISP's dependence on user behavior and partner ISPs: users with very light or very heavy usage are the most likely to defect, while ISPs with heavy-usage customers can benefit from declining to partner with the vISP. Our analytical results are verified with extensive numerical simulations.**

I. Introduction

Mobile users today are charged high prices for data plans from Internet service providers (ISPs), with an expensive base payment per month for a data quota and steep overage fees above this cap [1]. Most users desire cheaper data plans, but still expect to receive reasonable quality-of-service (QoS) and coverage. Meanwhile, current cellular and WiFi infrastructure are insufficient to support growing user demand [2], making it difficult for ISPs to maintain high QoS. New network technologies (e.g., 5G networks) can increase network capacity, but upgrading cellular networks is a long-term, expensive project.

In some developing countries with a competitive mobile operator market, prepaid or month-to-month data plans allow users to dynamically switch between ISPs' data plans. Another alternative for users to lower data costs is to subscribe to a mobile virtual network operator (MVNO), which resells wireless capacity from an infrastructure-owning ISP, often at lower costs. Given that they restrict to a single network, MVNOs may not meet users' QoS expectations. Thus, to satisfy both cost and QoS concerns, we propose leveraging existing network infrastructure through a cross-carrier data plan in which *users can access multiple ISPs' networks*.

The work in this paper was in part supported by NSF CNS-1347234 and CNS-1456847, the Research Grants Council of Hong Kong under Projects No. RGC 11207615, 7004680 and the Hong Kong Innovation and Technology Fund under Grant ITS/180/16.

A. A Virtual ISP Data Plan

A cross-carrier data plan would allow users to subscribe to a "virtual" ISP (vISP) that combines the resources of multiple partner ISPs. Traffic from vISP users can then be handled by the partner ISP's network with the best QoS, e.g., the highest throughput. While this infrastructure sharing approach is technologically feasible [3], its economic viability remains an open question. Anti-trust regulations can restrict efforts to merge operators [4]. Instead, a *third party* is required to handle this sharing; for instance, in the U.S., Google has introduced a cross-carrier data plan called Project Fi [5] that pools T-Mobile, Sprint, and US Cellular infrastructure.

However, it is unclear whether a third party vISP can earn a positive profit, while satisfying anti-trust regulations.

- ISPs who can maintain a high throughput for their users are less likely to partner with the vISP, and thus become non-partner ISPs. If the vISP charges too much, users of these non-partner ISPs may not wish defect to the vISP.
- Smaller ISPs may join the vISP as partner ISPs to gain some revenue from leasing their capacity. However, if they lose users to the vISP, it will decrease their revenue.
- If the vISP offers a very low price in order to attract users, too many partner ISPs' users may defect, increasing the price charged by the partner ISPs and jeopardizing the vISP's profit. Even if the vISP can make a profit, it may attract too many partner ISPs and users, violating the anti-trust regulations it is supposed to protect.

Figure 1 illustrates the interactions between users, ISPs, and the vISP: users of both partner and non-partner ISPs must decide whether to defect to the vISP, while ISPs must decide whether to partner with the vISP. The viability and impact of a vISP therefore depend on the complex interactions between the decisions of the vISP, partner ISPs, and users. In this work, we study whether, and quantify under what circumstances of user demands, the vISP, partner ISPs, users, and even non-partner ISPs will benefit from the vISP's data plan. Our results show that *while the vISP can make a profit and benefit both users and ISPs in the short term, it may not remain viable in the long term as users' data demands increase*. Rather than cannibalizing the mobile data market, the vISP is better understood as an interim solution for ISPs until they upgrade their networks to accommodate growing user demand.

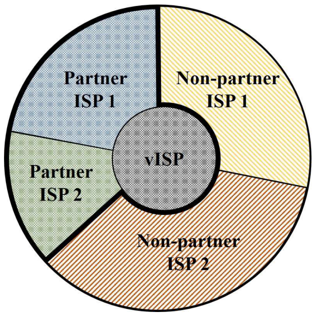


Fig. 1. Example of users on the vISP, partner ISPs, and non-partner ISPs. Sectors represent ISPs' market shares before the vISP enters the market, and users in each sector defect to it. The region inside the bolded line indicates that the vISP's users share the partner ISPs' network infrastructure.

B. Related Work

Some have considered the economic impact of pricing different WiFi access points [6] or joint pricing of different network technologies, e.g., cellular and WiFi, offered by one ISP [7], [8]. Others have gone further in using prices to incentivize users to offload their data onto WiFi networks [9], [10] or allowing users to trade leftover data among each other [11]. Still other works consider inter-ISP pricing in a hierarchical model of transit and local ISPs for wireline networks [12], as well as the tiered pricing often offered by transit ISPs selling capacity to local ISPs [13]. Our work, in contrast, considers a non-hierarchical setting in which a vISP combines the infrastructure of multiple partner ISPs. We focus on the impact of inter-ISP pricing (i.e., the vISP's payment to partner ISPs) on the price that the vISP charges end users, and users' subsequent decisions of whether to defect to the vISP.

Other works have focused on technological aspects of a shared mobile network infrastructure. In [14]–[16], vertical handoff decision algorithms are proposed that consider users' mobility, device switching cost and the quality of connection. The authors in [3], [17] study network switching to maximize user throughput, while [18] proposes a framework for "service" ISPs to use multiple network infrastructures. However, while some works have considered the economics of ISP spectrum sharing agreements [19], existing works generally do not consider the economics of users' decisions of whether to subscribe to a single or shared network.

C. Economic Impact of the vISP

We suppose each vISP user's device can switch between partner ISP networks following policies specified by the vISP.[1] We assume the vISP charges users in proportion to their usage volume, as Google Fi does, while the partner and non-partner ISPs offer a data cap with overage plan. We consider a user population with heterogeneous "natural" usage levels, which we define as the user demands when they are not charged for data. For instance, some users are rarely interested in streaming videos and thus consume little data. We refer to users as "light" or "heavy" depending on their natural usage levels. In our analysis, we answer three major questions:

[1]Google requires its Project Fi users to choose from selected smartphone models, allowing such policies to be implemented on the device.

How many users subscribe to the vISP? (Section II) Users decide whether to defect to the vISP or remain with their current ISP, depending on the achievable throughput and the usage-based price charged by the vISP. Their decisions are not made independently: the number of users on each network influences each user's throughput, leading to a feedback loop. We develop a user model that incorporates the throughput feedback effects on users, and show that users' defection rates for each ISP always reach an equilibrium. While we would expect light users to defect, as they can save money by doing so [20], *we find that heavy users may also defect from partner ISPs if the vISP charges a sufficiently low price.*

Which ISPs should partner with the vISP? (Section III) Given the equilibrium user defection rates, ISPs must decide whether or not to partner with the vISP. We find that *ISPs with lighter users are more likely to partner with the vISP*. These ISPs will experience more user defections, since lighter users (who do not fully utilize their data caps) can save money by switching to the vISP. Partnering with the vISP allows these ISPs to limit the resulting loss of revenue through payments from the vISP. These results cast doubt on the long-term viability of the vISP: increasing mobile data traffic [2], [21] may result in fewer ISPs that are motivated to join the vISP.

When does the vISP make a profit? (Section IV) Given its agreements with partner ISPs, the vISP must decide how much to charge its users so as to obtain a profit, without cannibalizing the market. *We show that the vISP can earn a positive profit if partner ISP users' natural usage is sufficiently light and if the partner ISPs' market share falls below a given upper bound.* The vISP thus aggregates smaller ISPs who might need the vISP in order to attract more users. However, the vISP is unviable if it partners with too many ISPs: intuitively, it then must pay partner ISPs more, resulting in a negative profit and preventing the vISP from cannibalizing the market. Combined with the vISP's dependence on partner ISPs with lighter users, this result suggests that a vISP represents a viable way to benefit users and ISPs when user demand is close to the available network capacity, fulfilling today's need for handling growing user demand.

Then, in Section V, we simulate the behavior of one million users to show that the vISP can make a profit under realistic conditions. We verify Sections II and III's findings on which users defect and which ISPs become partner ISPs, empirically demonstrating the vISP's viability conditions. We conclude in Section VI. All proofs are in the technical report [22].

II. User Decisions

When the vISP joins the mobile data market, users have a choice of defecting to the vISP from their current ISPs. Their decisions affect, and are in turn affected by, the demands and throughputs achieved by other users on each ISP's network. Even those users who do not defect may realize different demands due to other users' defections changing the throughput on their ISPs. We examine these dynamics by first developing a model of user demand in Section II-A, and then showing the implications for their defection decisions in Section II-B.

TABLE I
KEY TERMS AND SYMBOLS

Symbol	Definition
M	Number of ISPs who have their own network infrastructure and offer Internet access for their users.
N	Total number of users in the mobile data market who subscribe to one of the M ISPs.
K	Number of partner ISPs ($1 \leq K < M$).
φ_m	Market share for ISP m, i.e., the percentage of all users who subscribe to ISP m without the vISP in the market.
θ_m	Defection rate for ISP m's users, i.e., the percentage of ISP m's users who defect to the vISP.
z_i	User i's natural usage without considering price effects, i.e., each user's maximum demand if they do not need to pay.
d, η, ρ	Data plan offered by ISPs with a monthly cap d GB data charged at η and overage fee ρ per GB exceeding the cap.
p	Usage-based unit price charged by the vISP.
C_m	ISP m's total network capacity.
$\hat{c}$	Throughput per user of both partner ISPs' and vISP's users.
π_k	Unit price paid by the vISP to partner ISP k.

A. User Demands

Aa a first step, we model user demands for data before and after the vISP enters the market through utility maximization.

1) Before the vISP: Before the vISP enters the market, we consider $N \in \mathbb{Z}_+$ users who subscribe to one of the $M \in \mathbb{Z}_+$ ISPs ($N \gg M$). We suppose that ISP m has a market share of $\varphi_m N$ users ($\varphi_m \in (0,1)$ and $\sum_{m=1}^{M} \varphi_m = 1$). To focus on the impact of the vISP rather than the effects of different ISP data plans, we assume that an ISP charges users η for up to d GB of data per month with overage fee of ρ per GB exceeding this cap ($\eta/d < \rho$). We suppose that each ISP m has a fixed amount of available capacity C_m across all cells to support its users' traffic, and each user is allocated an equal share of the wireless medium $C_m/(\varphi_m N)$.[2] Over the time scale of one month, all users on ISP m's network are assumed to experience similar average throughputs.[3]

Suppose that user i's "natural" usage in a month, with free data usage, is z_i. We take z_i to be finite to account for the fact that there is an intrinsic limit to the amount of data most users wish to consume in a month. Most U.S. consumers, for instance, use less than 3GB of cellular data per month, far below many ISP data caps, indicating that they could have consumed additional data without paying more had they been so inclined [23]. Since ISPs do charge users for their data usage, we let $\tilde{z}_i$ denote their actual data usage over a month, and we model their utility, or satisfaction, from this usage with the standard α-fair utility function $x^{1-\alpha}/(1-\alpha)$ with $\alpha \in [0,1)$. By subtracting each user's payment to the ISP from this utility, each user i's utility from ISP m is then

$$U_i^m(\tilde{z}_i \mid d, \eta, \rho) = c_i^m \frac{\tilde{z}_i^{1-\alpha}}{1-\alpha} - \eta - (\tilde{z}_i - d)^+ \rho, \qquad (1)$$

for $\tilde{z}_i \leq z_i$,[4] where $(\tilde{z}_i - d)^+$ indicates that the user pays no overage for usage under the cap d. The scaling factor c_i^m represents the user's desire for high throughput, which we set to $\frac{C_m}{\varphi_m N}$, to capture the fact that users who experience higher throughputs will likely derive greater utilities from their data usage. By maximizing the utility function in (1), user i's maximum utility and optimal demand from ISP m are:

$$U_i^m(\tilde{z}_i^\star \mid c_i^m, d, \eta, \rho) = \begin{cases} U_i^m(z_i \mid c_i^m, d, \eta, \rho), & \text{if } z_i \leq \left(\frac{c_i^m}{\rho}\right)^{\frac{1}{\alpha}}, \\ U_i^m\left(\left(\frac{c_i^m}{\rho}\right)^{\frac{1}{\alpha}} \Big| \, c_i^m, d, \eta, \rho\right), & \text{otherwise.} \end{cases} \qquad (2)$$

To derive (2), we assume $c_i^m > \rho d^\alpha$, i.e., the throughput is high enough so that users still receive positive marginal utility at their data cap d, unless their natural usage $z_i < d$.

We suppose that the natural usage z of each user on each ISP m is i.i.d. on a heavy-tailed Pareto distribution whose probability density function is $f_m(z) = \frac{\lambda_m \delta_m^{\lambda_m}}{z^{\lambda_m+1}}$ with parameter $\lambda_m > 1$ and a minimum usage δ_m. A smaller λ_m means that this ISP has a higher percentage of heavy users. To ensure that all users receive positive utilities from using data (otherwise they would not subscribe to the ISP), we assume $\delta_m = (\frac{(1-\alpha)\eta}{c_i^m})^{\frac{1}{1-\alpha}}$, where $\delta_m \leq (\eta/\rho)$ due to $c_i^m > \rho d^\alpha$.

2) User Demands with the vISP: We use $\theta_m \in [0,1]$ to denote the fraction of ISP m's users who defect to the vISP, i.e., the *defection rate*.[5] Thus, the total number of vISP users is $\hat{N} = \sum_{m=1}^{M} \theta_m \varphi_m N$. The vISP then connects each of these $\hat{N}$ users to one of its partner ISPs' networks. We assume that partner ISPs are not allowed to prioritize their own users over the vISP's and that there are $K \leq M$ partner ISPs, $k = \{1, 2, \ldots, K\}$, and $M - K$ non-partner ISPs, $m = \{K+1, \ldots, M\}$.

We also suppose that there are $\hat{n}_k$ out of $\hat{N}$ users who are assigned to partner ISP k's network by the vISP. If the vISP always selects the best cellular network among all partner ISPs' networks for its users, eventually, the throughputs of each of the K partner ISPs would be averaged out to equal each other, i.e., $C_k/((1-\theta_k)\varphi_k N + \hat{n}_k) = C_j/((1-\theta_j)\varphi_j N + \hat{n}_j)$, $\forall k, j = 1, 2, \ldots, K$.[6] More formally, we have the following:

Lemma 1: If the vISP always selects the partner ISP network with the best throughput for its users, vISP users' average throughput is given by

$$\hat{c} = \frac{\sum_{k=1}^{K} C_k}{\left(\sum_{k=1}^{K} \varphi_k + \sum_{m=K+1}^{M} \theta_m \varphi_m\right) N}. \qquad (3)$$

This $\hat{c}$ is also the throughput of partner ISPs' users, since vISP and partner ISP users share the same network infrastructure. We term $\hat{c}$ users' *shared throughput*, and assume $c_i^m > \rho d^\alpha$.

[2] We assume that all cells of an ISP have roughly the same capacity, and users access them with uniformly random probability.

[3] We do not explicitly consider users' access to WiFi hotspots, instead assuming that this access does not change with their ISP subscription.

[4] Since users' natural usage is their maximum consumption without being charged, we assume that they consume no more than z_i when actually charged.

[5] We do not consider the case of new smartphone users. Since they must also choose between the vISP, partner, and non-partner ISPs, the discussion in this section is also applicable to new users. We suppose N is large enough that θ can be approximated as continuous on $[0, 1]$.

[6] We suppose that there are sufficiently many vISP users for the vISP to enforce this equality (cf. Section II-B). If there are too few vISP users, it is possible that some partner ISPs may have lower throughputs than others due to handling too much traffic from their own users.

Though for some users $\hat{c} > c_i^m$, i.e., the shared throughput is larger than the throughput before joining the vISP, this may not be the case for all users:

Proposition 1: The shared throughput is lower than the maximum throughput of partner ISPs before the vISP enters the market: $\hat{c} \leq \max_{k=1,\ldots,K} \{C_k/(\varphi_k N)\}$.

Intuitively, some partner ISPs, due to their larger network capacities, would receive more users from the vISP, reducing their average throughput.

We further observe from (3) that $\hat{c}$ is not affected by the number defecting from partner ISPs, only by the number of users defecting from non-partner ISPs. Too many users defecting from non-partner ISPs harm the shared throughput:

Corollary 1: Users' minimum throughput among partner ISPs before the vISP exceeds the shared throughput, i.e., $\hat{c} \leq \min_{k=1,\ldots,K} \left\{\frac{C_k}{\varphi_k N}\right\}$, if the number of users defecting from non-partner ISPs satisfies $\sum_{m=K+1}^{M} \theta_m \varphi_m \geq \left(\max_{k=1,\ldots,K} \left\{\frac{C_k}{\varphi_k N}\right\} \Big/ \min_{k=1,\ldots,K} \left\{\frac{C_k}{\varphi_k N}\right\} - 1\right) \sum_{k=1}^{K} \varphi_k$.

From Lemma 1, we can also find the number of vISP users in partner ISP j's network: $\hat{n}_j = \left(\frac{C_j}{\sum_{k=1}^{K} C_k}\left(\sum_{k=1}^{K} \varphi_k + \sum_{m=K+1}^{M} \theta_m \varphi_m\right) - (1-\theta_j)\varphi_j\right) N$. The vISP needs to pay partner ISP j for the traffic generated by these $\hat{n}_j$ users.

User i's utility from the vISP data plan then consists of the user's usage utility for consuming $\hat{z}_i$ amount of data, and a usage-based payment of p per GB for their usage:

$$\hat{U}_i(\hat{z}_i \,|\, \hat{c}, p) = \hat{c}\frac{\hat{z}_i^{1-\alpha}}{1-\alpha} - \hat{z}_i p, \tag{4}$$

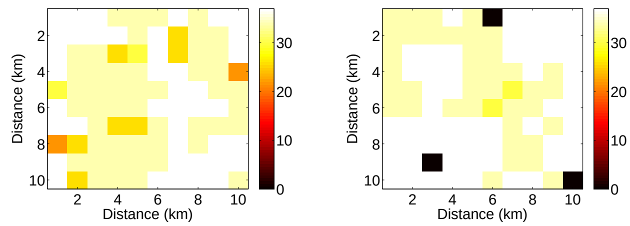
where $\hat{z}_i \leq z_i$, user i's natural usage. We note that in (4), the scale factor for usage utility is replaced with the shared throughput $\hat{c}$. We thus find user i's maximum utility and optimal data demand $\hat{z}_i^\star$ if user i defects to the vISP:

$$\hat{U}_i(\hat{z}_i^\star \,|\, \hat{c}, p) = \begin{cases} \hat{U}_i(z_i \,|\, \hat{c}, p), \text{ if } z_i \leq \left(\frac{\hat{c}}{p}\right)^{\frac{1}{\alpha}}, \\ \hat{U}_i\left(\left(\frac{\hat{c}}{p}\right)^{\frac{1}{\alpha}} \,\Big|\, \hat{c}, p\right), \text{ otherwise}, \end{cases} \tag{5}$$

where we have $\hat{c}/p \geq \hat{c}/\rho \geq d^\alpha$ due to the assumption $p < \rho$ and $\hat{c} > \rho d^\alpha$. Comparing (5) with users' utility without the vISP, (2), we observe that partner or non-partner users consume at most $(c_i^m/\rho)^{\frac{1}{\alpha}}$ amount of data, while vISP users consume at most $(\hat{c}/p)^{\frac{1}{\alpha}}$. Thus, users can realize higher demands for data at the vISP if $\hat{c}/p > c_i^m/\rho$. In the next section, we compare users' utilities with and without the vISP to determine users' defection rates.

B. User Defection Rates

We can now move on to characterize the users who defect to the vISP. We make the following two assumptions:



(a) $\max\{$T-Mobile, Sprint$\}$. (b) AT&T.

Fig. 2. A comparison of network performance (i.e., average rates) of (a) the vISP and (b) a non-partner ISP [24].

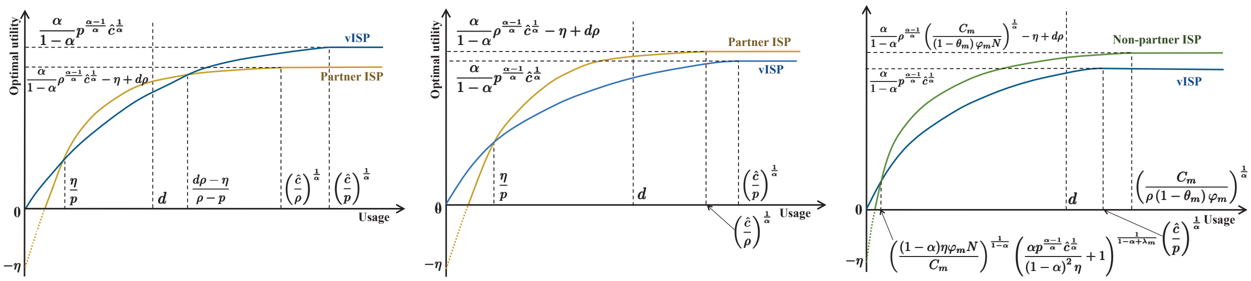
- Partner ISPs originally have lower average throughput than that achieved by the non-partner ISPs: $\frac{C_k}{\varphi_k N} < \frac{C_m}{\varphi_m N}$, $\forall k = 1, \ldots, K$ and $\forall m = K+1, \ldots, M$, resulting in $\hat{c} < \frac{C_m}{\varphi_m N}$. In Figure 2, we use data from public, crowd-sourced databases to estimate cellular signal strength from a 10 km $\times$ 10 km square area in downtown San Francisco [24]. We observe from Figure 2 that AT&T, which is not a Google Fi partner, has average throughput 34.32 Mbps, exceeding the maximum throughput of partner ISPs T-Mobile and Sprint which is 34.07 Mbps.
- The usage-based unit price offered by the vISP is higher than the unit price offered by the data plan with a monthly quota: $\eta/d < p < \rho$. For example, Google Fi offers $p =$\$10/GB, while T-Mobile, Sprint, AT&T, and Verizon offer roughly $\eta/d \approx$\$7/GB, and both Verizon and AT&T have an overage fee $\rho =$\$15/GB.

We suppose that users defect from their current ISPs to the vISP if they can obtain a better utility with the vISP. Figure 3 depicts users' utilities when subscribing to the vISP, partner and non-partner ISPs. Since the vISP employs usage-based pricing, the vISP users' utilities are nonnegative and increase from $z_i = 0$ but are eventually exceeded by both the utilities for non-partner and partner ISPs' users. Since the vISP users and partner users, sharing the same network infrastructure, have the same average throughput, their relative utilities depend heavily on the vISP's price p.

Compounding the difficulty of our analysis is the fact that users's defection decisions are not made independently. As more users defect to the vISP, for instance, the vISP's shared throughput will decrease, potentially driving some users to switch back to their original ISPs. We thus derive users' defection decisions in terms of users' aggregate defection rates θ_m and then analyze the resulting time dynamics of users' defection rates in Section II-B3.

1) Defections from Partner ISPs: As discussed above, the partner ISPs' users have the same average throughput as the vISP users, since they share the same physical infrastructure. Thus, users who do not defect obtain utility $U_i^k(\tilde{z}_i^\star \,|\, \hat{c}, d, \eta, \rho)$. Users defect if they can gain more utility from the vISP, i.e., $U_i^k(\tilde{z}_i^\star \,|\, \hat{c}, d, \eta, \rho) \leq \hat{U}_i(\hat{z}_i^\star \,|\, \hat{c}, p)$. Partner users' decisions then depend entirely on the price p charged by the vISP, making the defection rate a piece-wise function of p.

Proposition 2: Users of partner ISP k defect to the vISP if



(a) Partner ISP when $p \le \rho - (d\rho - \eta)\left(\frac{\hat{c}}{\rho}\right)^{-\frac{1}{\alpha}}$. (b) Partner ISP when $p \ge \rho - (d\rho - \eta)\left(\frac{\hat{c}}{\rho}\right)^{-\frac{1}{\alpha}}$. (c) Non-partner ISP.

Fig. 3. User utility comparison subscribing to the vISP, non-partner and partner ISPs under different price conditions.

and only if

$$\begin{cases} z_i \le \frac{\eta}{p} \text{ or } z_i \ge \frac{d\rho-\eta}{\rho-p} & \text{if } p \le \rho - (d\rho-\eta)\left(\frac{\hat{c}}{\rho}\right)^{-\frac{1}{\alpha}} \\ z_i \le \frac{\eta}{p} & \text{otherwise.} \end{cases} \quad (6)$$

The defection rate for partner ISP k is

$$\theta_k(p) = \begin{cases} 1 - \left(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\right)^{\frac{\lambda_k}{1-\alpha}} \left(\left(\frac{\eta}{p}\right)^{-\lambda_k} - \left(\frac{d\rho-\eta}{\rho-p}\right)^{-\lambda_k}\right), \\ \qquad\qquad \text{if } p \le \rho - (d\rho-\eta)\left(\frac{\hat{c}}{\rho}\right)^{-\frac{1}{\alpha}}, \\ 1 - \left(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\right)^{\frac{\lambda_k}{1-\alpha}} \left(\frac{\eta}{p}\right)^{-\lambda_k}, \text{ otherwise.} \end{cases} \quad (7)$$

In Proposition 2 and the rest of the paper, we suppose that N is sufficiently large that the (expected) number of users for which (6) holds can be approximated by $N\theta_k(p)$. If the vISP charges a relatively high price, i.e., $p \ge \rho - (d\rho-\eta)\left(\frac{\hat{c}}{\rho}\right)^{-\frac{1}{\alpha}}$, only users with natural usage less than η/p will defect (cf. Figure 3(b)); otherwise, users with natural usage more than $\frac{d\rho-\eta}{\rho-p}$ will also defect (cf. Figure 3(a)). Users with a lower natural usage that is well below the partner ISP's data cap can always save money with the vISP compared to the partner ISP, since they can avoid the flat-rate fee for the partner ISP's cap. Those with higher natural usage z_i will need to pay the partner ISP more than the vISP, as long as $d < z_i < \frac{d\rho-\eta}{\rho-p}$. However, if $\frac{d\rho-\eta}{\rho-p} < \left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}$, or equivalently $p \le \rho - (d\rho-\eta)\left(\frac{\hat{c}}{\rho}\right)^{-\frac{1}{\alpha}}$, the vISP users pay less than the partner ISP's for usage above $\frac{d\rho-\eta}{\rho-p}$, inducing heavier users to defect to the vISP.

2) Defections from Non-partner ISPs: We also consider a non-partner ISP m and suppose that a fraction θ_m of the original $\varphi_m N$ non-partner users defect to the vISP, increasing its average throughput to $c_i^m = \frac{C_m}{(1-\theta_m)\varphi_m N}$. Substituting c_i^m into (2), we find that user i's utility from ISP m and the vISP respectively are $U_i^m(\tilde{z}_i^\star \mid \frac{C_m}{(1-\theta_m)\varphi_m N}, d, \eta, \rho)$ and $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$. As with non-partner ISPs, we would expect light users to defect in order to avoid the non-partner ISP's flat data cap fee. Moreover, since non-partner ISPs provide better throughputs than the vISP $\left(\frac{C_m}{(1-\theta_m)\varphi_m N} \ge \frac{C_m}{\varphi_m N} \ge \hat{c}\right)$, heavy users who are sensitive to throughput changes are less likely to defect:

Lemma 2: No non-partner user with $z_i \ge d$ defects.

The vISP is unable to provide higher throughput to attract non-partner users, so it can only attract light users with $z_i < d$, who may pay a higher unit price for their usage with the non-partner ISP than the unit price offered by the vISP.

By comparing users' utilities from the vISP and non-partner ISP m, and recalling that users' natural usage follows a Pareto distribution, we identify the users who would defect and derive the defection rate for non-partner ISP m.

Proposition 3: If the vISP provides sufficient throughputs satisfying $\hat{c} \ge \frac{1}{\alpha \delta_m^{\lambda_m}} \frac{C_m}{\varphi_m N}$, users defect from non-partner ISP m to the vISP if and only if

$$z_i \le \left(\frac{(1-\alpha)\eta\varphi_m N}{C_m}\right)^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{1-\alpha}\left(\frac{p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}}}{(1-\alpha)\eta}\right) + 1\right)^{\frac{1}{1-\alpha+\lambda_m}}. \quad (8)$$

The defection rate for non-partner ISP m is then

$$\theta_m(p) = 1 - \left(\frac{(1-\alpha)\eta}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} + (1-\alpha)\eta}\right)^{\frac{\lambda_m}{1-\alpha+\lambda_m}} \quad (9)$$

Intuitively, as p decreases and the vISP charges users less, more users will defect to the vISP. Mathematically, we see that both $\theta_m(p)$ in (9) and $\theta_k(p)$ in (7) decrease with p.

3) Defection Rate Equilibria: We now show that users' defection decisions converge to a long term equilibrium. The defection conditions derived in Propositions 2 and 3 assume that users make their decisions based on the fixed defection rates θ_m, but these user decisions can change the defection rate, prompting a change in users' defection decisions. We address these dynamics in this section.

From Lemma 1, we note that the shared throughput $\hat{c}$ in (3) depends only on the defection rates $\theta_{k+1}, \ldots, \theta_M$ from non-partner ISPs; it does not depend on the partner ISPs' defection rates. Thus, from (7) and (8), the defection rates θ_m from each partner ISP are completely determined by fixed system parameters and $\hat{c}$, while the defection rates from each non-partner ISP do not depend on the rates for partner ISPs. We therefore focus on the non-partner ISPs' defection rates. For ease of notation, we write the shared throughput as $\hat{c}(\vec{\theta}(t))$ with $\vec{\theta}(t) = [\theta_{k+1}(t), \ldots, \theta_M(t)]^\top$ representing a vector of the non-partner ISPs' defection rates at a given time t.

Given defection rates $\theta_m(t)$ and the shared throughput $\hat{c}(\vec{\theta})$, we define $h_m(\vec{\theta})$ to be the time derivative of $\vec{\theta}$:

$$\frac{\mathrm{d}\vec{\theta}(t)}{\mathrm{d}t} = 1 - \theta_m(t) - \left(\frac{\eta(1-\alpha)}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}(\vec{\theta})^{\frac{1}{\alpha}} + (1-\alpha)\eta}\right)^{\frac{\lambda_m}{1-\alpha+\lambda_m}}, \quad (10)$$

for each non-partner ISP m. The quantity h_m represents the fraction of users who wish to defect, as derived from (8), less the fraction who have already defected, $\theta_m(t)$. Our goal is now to show that the dynamics (10) converge to a long-term equilibrium. Note that if $\theta_m(0) \in [0,1]$ for all m, then each $\theta_m \in [0,1]$ at any time t: the unit cube $[0,1]^{M-K}$ is a positively invariant set for these dynamics. This sanity check ensures that θ_m can always be interpreted as a defection rate.

We observe that these equations form a nonlinear dynamical system with state variables given by $\vec{\theta}$. Proposition 3 gives a set of fixed-point equations that any equilibrium point of (10) must satisfy, namely, (9). We show that there is a unique point satisfying (9), and that (10) always converges to it:

Proposition 4: There exists a unique limit point $\vec{\theta}^\star \in [0,1]^{M-K}$ of (10). Moreover, (10) converges to $\vec{\theta}^\star$.

We can thus take (9) as determining the unique equilibrium defection rates for non-partner ISPs' users. These rates can then be substituted into (7) to determine the partner ISPs' equilibrium defection rates.

III. Impact on Partner and Non-partner ISPs

Given users' defection rates for partner and non-partner ISPs, we now turn to analyzing the vISP's impact on both types of ISPs. In particular, we examine the implications for their revenue, using our results to understand which ISPs are more likely to partner with the vISP.

A. Partner ISP Revenue

Suppose the partner ISP k charges the vISP a usage-based price π_k. After losing $\theta_k \varphi_k N$ users to the vISP, ISP k experiences the following expected change in revenue:

$$\Delta\mathcal{R}_k(\theta_k, p) = \left(\frac{\rho}{\lambda_k - 1}\Big(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\Big)^{\frac{\lambda_k}{1-\alpha}}\left(\Big(\frac{C_k}{\rho\varphi_k N}\Big)^{\frac{1-\lambda_k}{\alpha}} - \chi\right) - \theta_k\eta\right)\varphi_k N, \tag{11}$$

where χ is given by

$$\chi = \begin{cases} \Big(\dfrac{\hat{c}}{\rho}\Big)^{\frac{1-\lambda_k}{\alpha}}, \text{ if } p \geq \rho - (d\rho - \eta)\Big(\dfrac{\hat{c}}{\rho}\Big)^{-\frac{1}{\alpha}}, \\ \Big(\dfrac{dp-\eta}{d\rho - \eta} + \dfrac{1}{\lambda - 1}\Big)\Big(\dfrac{d\rho - \eta}{\rho - p}\Big)^{-\lambda_k + 1}, \text{ otherwise.} \end{cases} \tag{12}$$

These equations are derived in the proof of Proposition 5.

By partnering with the vISP, the partner ISP not only loses some of its own users, but may also decrease its average throughput (cf. Corollary 1) and thus user demands, leading to a decrease in revenue:

Proposition 5: If the shared throughput is less than the average throughput originally offered by partner ISP k, i.e., $\hat{c} \leq \frac{C_k}{\varphi_k N}$, then $\Delta\mathcal{R}_k(\theta_k, p) \leq 0$, i.e., the partner ISP's revenue decreases after sharing its network infrastructure with the vISP.

As discussed in Corollary 1, Proposition 5 is likely to occur if too many users from non-partner ISPs are attracted to vISP. To compensate its revenue loss, a partner ISP should charge the vISP a sufficiently high price for accessing its network to ensure that it does not lose any revenue. The partner ISP thus charges the vISP the minimum amount for which it is incentivized to partner with the vISP. We suppose that the vISP will refuse to pay more than this amount, knowing that the ISP will still partner with it for a lower payment. In what follows, we derive this price, which we denote as π_k, by dividing the partner ISP's loss in revenue by its vISP traffic.

By Lemma 1, $\frac{\hat{n}_k}{\hat{N}}$ of the vISP's expected traffic goes through partner ISP k's network, and the total vISP traffic is:

$$\mathcal{D}(p) = \left(\sum_{k=1}^{K}\varphi_k\int_{\mathcal{Z}_k} z f_k(z)\mathrm{d}z + \sum_{m=K+1}^{M}\varphi_m\int_{\mathcal{Z}_m} z f_m(z)\mathrm{d}z - \mathbb{1}(p)\sum_{k=1}^{K}\varphi_k\int_{(\frac{\hat{c}}{p})^{\frac{1}{\alpha}}}^{\infty}\left(z - \Big(\frac{\hat{c}}{p}\Big)^{\frac{1}{\alpha}}\right) f_k(z)\mathrm{d}z\right)N, \tag{13}$$

where $\mathbb{1}(p)$ is an indicator function that equals 1 if $\frac{d\rho-\eta}{\rho-p} \leq (\frac{\hat{c}}{\rho})^{\frac{1}{\alpha}}$, and 0 otherwise. We use $\mathcal{Z}_k$ to denote the users who defect from ISP k, integrating over their Pareto natural usage distributions. To understand (13), we recall from (5) that a vISP user i does not change his or her data consumption if $z_i \leq (\frac{\hat{c}}{p})^{\frac{1}{\alpha}}$, but otherwise reduces his or her usage to $(\frac{\hat{c}}{p})^{\frac{1}{\alpha}}$. Thus, when $\frac{d\rho-\eta}{\rho-p} \leq (\frac{\hat{c}}{\rho})^{\frac{1}{\alpha}}$, the partner users for whom $z_i \geq \frac{d\rho-\eta}{\rho-p}$ would defect (Proposition 2), but those with $z_i \geq (\frac{\hat{c}}{p})^{\frac{1}{\alpha}}$ would only add $(\frac{\hat{c}}{p})^{\frac{1}{\alpha}}$ amount of traffic each to vISP. The partner ISP k thus sells data to the vISP at a price:

$$\pi_k = \frac{-\Delta\mathcal{R}_k(\theta_k, p)}{\frac{\hat{n}_k}{\hat{N}}\mathcal{D}(p)}. \tag{14}$$

Partner ISPs neither lose nor gain revenue from partnering with the vISP. Non-partner ISPs, however, may lose revenue, driving some ISPs to partner with the vISP.

B. Non-partner ISP Revenue

Although non-partner ISPs lose some users to the vISP, they may experience greater traffic in their networks as their remaining users increase their demands due to higher throughputs. Non-partner ISP m's change in revenue is then:

$$\Delta\mathcal{R}_m(\theta_m, p) = \left(\frac{(\eta(1-\alpha))^{\frac{\lambda_m}{1-\alpha}}}{\lambda_m - 1}\Big(\frac{C_m}{\rho^{1-\alpha}\varphi_m N}\Big)^{\frac{1-\lambda_m-\alpha}{\alpha(1-\alpha)}}\Big(1 - (1-\theta_m)^{\frac{\lambda_m - 1}{\alpha}}\Big) - \theta_m\eta\right)\varphi_m N, \tag{15}$$

where λ_m is the parameter of the Pareto distribution for its users' natural usage. We derive (15) in the proof of Proposition 6:

Proposition 6: If the parameter λ_m of users' natural usage distribution for ISP m satisfies

$$\lambda_m \leq \min\left\{1 + \alpha, \frac{(1-\alpha)\big(\log(d\rho) - \log(\alpha\eta)\big)}{\log(d\rho) - \log\big((1-\alpha)\eta\big)}\right\}, \tag{16}$$

then the non-partner ISP's revenue increases after the vISP enters the market.

Proposition 6 implies a lower bound on the minimum natural usage for ISP m's users:

Corollary 2: If (16) holds for ISP m, the minimum usage of its users' natural usage distribution satisfies $\delta_m \geq (\alpha\eta/\rho)$.

Since a smaller parameter λ_m and a larger minimum usage δ_m for a Pareto distribution indicate a CDF with more moderate increase at the beginning and longer tail at the end, Proposition 6 and Corollary 2 indicate that ISPs with heavier users are more likely to increase their revenue by not partnering with the vISP. Since lighter users are more likely to defect to the vISP (Proposition 3), these ISPs will experience fewer defections and a lower revenue loss, which can be compensated with an increase in demand from heavier users.

These results cast doubt on the long-term viability of the vISP: the increase in data usage predicted in [2], [21] can be modeled as an increase in users' natural usage, as it is driven by an increase in ways to use mobile data, not by the price or throughput of data consumption. Thus, over time we would expect λ_m to decrease and δ_m to increase, resulting in more ISPs with heavier users who can gain more revenue by declining to partner with the vISP. In the next section, we examine the vISP's profit and show that it can remain viable even as fewer ISPs are willing to partner with it.

IV. Optimal vISP Strategy and Its Viability

Building on our analysis of user behavior and ISPs' willingness to partner with the vISP in Sections II and III, we can now derive the vISP's optimal strategy, i.e., the price it charges its users, which we denote as p. Figure 4 summarizes our findings, with the top row of rectangles representing users' defections, and the bottom row representing vISP profit and ISP revenue before and after the vISP joins the market. Intuitively, the vISP can maximize its profit by offering a lower price, thus attracting more users. Yet, as more users defect from partner ISPs, the vISP needs to pay the partner ISPs more to compensate their loss in revenue. Thus, the vISP's goal is to simultaneously attract more users from non-partner ISPs[7] and pay as little to partner ISPs as possible.

The vISP's objective in choosing its price is to maximize its profit, which consists of its income from vISP users, $p\mathcal{D}(p)$, less its payment to partner ISPs. The vISP pays each partner ISP k at the rate π_k found in (14), for a total payment of $\sum_{k=1}^{K} \pi_k \frac{\hat{n}_k}{\hat{N}} \mathcal{D}(p) = -\sum_{k=1}^{K} \Delta\mathcal{R}_k(\theta_k, p)$. The vISP thus derives its price by solving the optimization problem:

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & p\mathcal{D}(p) + \sum_{k=1}^{K} \Delta\mathcal{R}_k(\theta_k, p) \\ \text{subject to} \quad & \frac{\eta}{d} \leq p \leq \rho \end{aligned} \tag{17}$$

[7]We assume that Proposition 6 holds for all non-partner ISPs; otherwise, they would be partner ISPs.

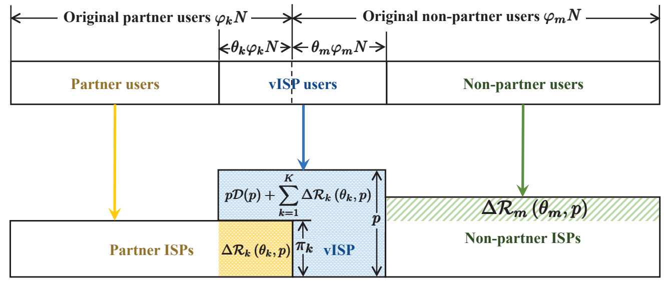


Fig. 4. Market dynamics and payments between users and ISPs. The top rectangles represent the number of users on the vISP and each ISP, while the bottom rectangles represent each ISP's revenue. An arrow from A to B means that party A pays party B for their data traffic. The shaded areas in the bottom rectangles represent the change in revenue when the vISP joins the market.

We call the vISP's business model *viable* if it makes a positive profit, i.e., the optimal value of (17) is larger than zero. We find realistic conditions for the vISP's viability:

Proposition 7: There exists a feasible point for which the objective (17) is positive, if the parameters λ_k of users' natural demand distributions for each partner ISP k satisfy

$$\lambda_k \geq \frac{\eta}{p}\left(1 + \left(\frac{p}{\rho}\right)^{\lambda_k}\right)\left(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\right)^{-\frac{1}{1-\alpha}}, \tag{18}$$

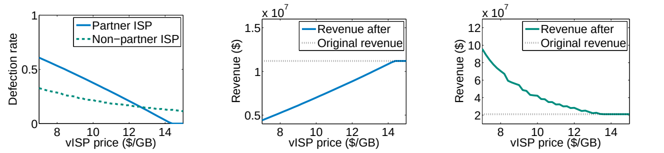
and the total percentage of users for all partner ISPs satisfies

$$\sum_{k=1}^{K} \varphi_k \leq \frac{\frac{\alpha^2}{(1-\alpha)^2} dp}{\frac{1}{1-2\alpha}\frac{d\rho-\eta}{\rho-p}\rho + \frac{\alpha^2}{(1-\alpha)^2} dp}. \tag{19}$$

This finding dovetails with our result for non-partner ISPs in Proposition 6: non-partner ISPs tend to have heavier users, while the vISP is more likely to be viable if its partner ISPs' users have lighter usage distributions with a larger parameter λ_k. Moreover, the vISP can actually jeopardize its profit by partnering with too many ISPs, or with ISPs that have too many users. Thus, the vISP can serve as a way for smaller ISPs with fewer users to work together in order to attract more users, as T-Mobile, Sprint, and US Cellular have done with Google Fi. The limit to the vISP's market share further prevents it from cannibalizing the market, helping to ensure its viability from a regulatory perspective.

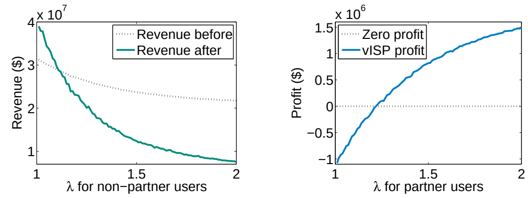
Though Proposition 7 establishes the vISP's short-term viability with a positive profit, the condition (18) may not hold in the long term as usage levels increase. As discussed in Section III, users' natural usage is expected to increase over time, meaning that the λ_k parameters may decrease as more users join the "heavy tail" of the natural usage distribution. Thus, the vISP may eventually be forced out of business; however, it can still benefit the mobile data market in the short-term by allowing partner ISPs to attract more users and allowing some users to increase their utilities.

Although (17) is a nonlinear programming problem, it can be numerically solved by a line search over all possible values of p. As data prices are usually rounded to integral values in practice for ease of users' understanding, searching over the integers in $[\frac{\eta}{d}, \rho]$ would generally suffice. In the next section,



(a) Defection rate. (b) Partner revenue. (c) Non-partner revenue.

Fig. 5. Defection rate and revenue changes for partner and non-partner ISPs ($K = 1$ and $M = 2$) in terms of vISP price. The partner and non-partner ISPs have the market share $\varphi_1 = 0.16$ and $\varphi_2 = 0.30$ with $\lambda_1 = 1.3$ and $\lambda_2 = 1.1$, and their total network capacities are $C_1 = 2.56 \times 10^6$ Mbps and $C_2 = 6.9 \times 10^6$ Mbps respectively.



(a) Non-partner revenue. (b) vISP profit.

Fig. 6. Changes in revenue for the non-partner ISP and profit for the vISP with different values of the parameters for the Pareto-distributed partner and non-partner users' natural usage. We take $p = \$10$/GB as the vISP's price.

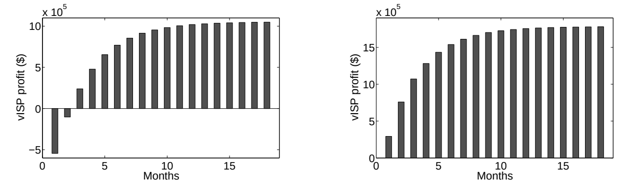
we provide numerical examples of a positive vISP profit and optimal price, as well as further numerical insights into the analysis in this section and Sections II and III.

V. Numerical Evaluation

We now evaluate the market dynamics caused by a vISP on a total of one million users, whose natural usage is randomly generated according to the Pareto distribution parameters of their associated ISPs. We set $\alpha = 0.25$, $\rho = \$15$/GB, $d =$ 10GB, and $\eta = \$15$/GB for all experiments in the section.

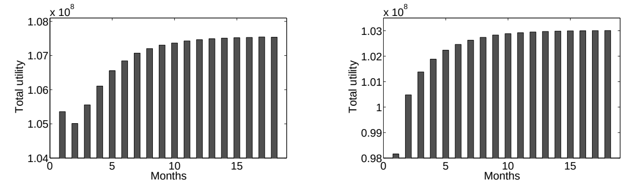
Figure 5 shows users' equilibrium defection rates and ISP revenues in a simple example of one partner ISP and one non-partner ISP. In Figure 5(a), defection rates for both partner and non-partner ISPs decrease with the vISP price: the defection rate for the partner ISP decreases sharply with the vISP price, while the non-partner ISP's defection rate decreases more moderately. We also observe that when the vISP price approaches the overage fee $\rho = \$15$/GB, almost no partner ISP users defect to the vISP data plan: users can no longer save money by defecting, and they experience the same throughput on the vISP and partner ISP. As expected, the partner ISP loses revenue without counting the payment received from the vISP, while the non-partner ISP's revenue in fact increases (Figures 5(b) and 5(c)). Surprisingly, as more light users defect, the non-partner ISP gains more revenue.

To be consistent with Section III, the partner ISP users' natural usage distribution has a larger parameter λ than the non-partner ISP's users. We further elaborate on the relationship between ISPs' partnership decisions and their users' natural usage distributions in Figure 6. In Figure 6(a), we fix $\lambda = 1.3$ for partner users and randomly generate natural usage for non-partner users based on the λ values on the x-axis, while in Figure 6(b), we fix $\lambda = 1.05$ for non-partner users and vary the λ parameter for the partner users. In Figure 6(a), the non-partner ISP's original revenue decreases as λ increases (i.e., there are more light users) as shown by the dotted black



(a) $p = \$8$/GB. (b) $p = \$14$/GB.

Fig. 7. The vISP earns a positive profit over time.



(a) $p = \$8$/GB. (b) $p = \$14$/GB.

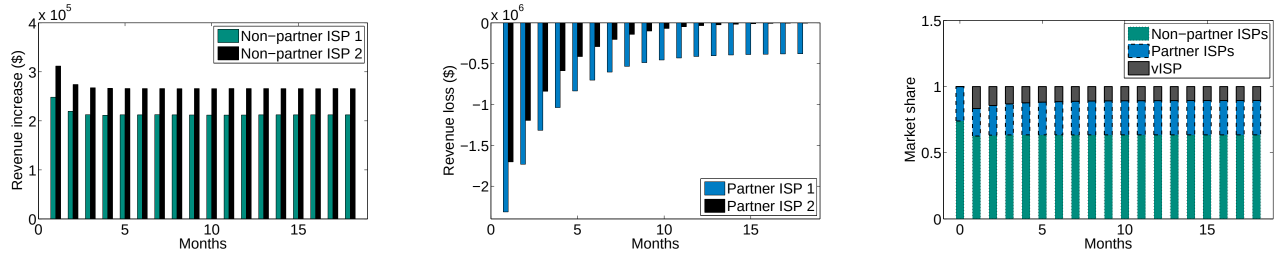
Fig. 8. Total utility of all users in the market increases over time.

curve, and its revenue after more light users defect to the vISP decreases even faster as shown by the blue solid curve. As expected from Proposition 6, the non-partner ISP gains revenue by not partnering with the vISP when λ is small, while ISPs with greater λ values partner with the vISP to avoid revenue loss. The vISP earns more profit with a greater λ for partner users, verifying Proposition 7; if partner ISPs' λ is too small, the vISP has negative profit.

We finally examine the market dynamics, considering two different prices charged by the vISP to their users: $p = \$8$ (i.e., $p \to \eta/d$) and $p = \$14$ (i.e., $p \to \rho$). We consider two partner ISPs (ISP 1 and ISP 2) and two non-partner ISPs (ISP 3 and ISP 4) with market shares $\varphi_1 = 0.12$, $\varphi_2 = 0.14$, $\varphi_3 =$ 0.34, and $\varphi_4 = 0.40$ and network capacities $C_1 = 3.36 \times 10^6$ Mbps, $C_2 = 2.80 \times 10^6$ Mbps, $C_3 = 1.36 \times 10^7$ Mbps, and $C_4 = 1.6 \times 10^7$ Mbps respectively. Since we abstract away from user mobility across cells, these capacities are the *total* network capacity, across all cells. Assuming non-partner ISPs have more heavy users than partner ISPs, we use $\lambda_1 = 1.5$, $\lambda_2 = 1.6$, and $\lambda_3 = \lambda_4 = 1.06$.

We simulate the dynamics of users switching between their original ISP and the vISP over 18 months. Users decide to defect or not at the beginning of each month by estimating their utilities on each ISP. However, they cannot anticipate other users' decisions, so their actual throughputs after defecting may differ from their estimates, possibly leading them to switch back after a month. We suppose that users who would gain utility by switching actually switch ISPs with probability $\sigma = 0.3$, e.g., if some users may not want to be bothered by signing up for a different data plan. We calculate the resulting total user utilities, vISP revenues, partner and non-partner ISP revenue changes, and market share between ISPs over time in Figures 7, 8, and 9 respectively.

As shown in Figure 7(a), the vISP has a negative profit in the first two months since it needs to pay partner ISPs sufficiently to make up for partner ISPs' high revenue loss (cf. Figure 9(b)). Starting from the third month, as some partner users switch back and more non-partner users defect to the vISP (cf. Figure 9(c)), vISP profit gradually increases. In both Figures 7(a) and 7(b), the vISP profit converges to a positive



(a) Revenue increase of non-partner ISPs. (b) Revenue loss of partner ISPs. (c) Market share between ISPs.

Fig. 9. ISPs' revenue and their market share converge over time when $p = \$8$/GB.

value over time when it charges users at either \$8/GB or \$14/GB. Comparing the converged profit values in Figure 7(a) and 7(b), the vISP is viable at both prices but earns more with $p = \$14$/GB. Figure 8(a) also shows the dynamics of the total utility for all users in the market. As the original total utility without the vISP is 8×10^7, users benefit from higher utilities with more data plan options.

We compare the difference of revenues for non-partner and partner ISPs in Figures 9(a) and 9(b) respectively. Non-partner ISPs' revenues increase as derived in Proposition 6, and their revenues are stable over time. Conversely, partner ISPs lose revenue unless they charge the vISP. As the vISP still earns a positive profit after paying partner ISPs, the vISP could motivate more ISPs to partner with it by paying them more. Finally, Figure 9(c) plots the market shares of all ISPs. Although non-partner ISPs initially dominate (as shown by the bar at 0), the vISP helps even out this imbalance.

VI. Conclusion

We examine the economic viability of a third-party virtual ISP and its effects on the mobile data market. Are there conditions under which the vISP is not only viable but also benefits partner ISPs, non-partner ISPs, and users? By investigating users' incentives to defect to the vISP and ISPs' incentives to partner with the vISP, we find that the vISP can make a positive profit if its partner ISPs' market share falls below an upper bound. Lighter users are more inclined to choose the vISP data plan, as they can save money by doing so, but heavy users may also defect if the vISP's prices are low enough. ISPs with more light users are correspondingly more likely to partner with the vISP, as they can lose revenue otherwise, while non-partner ISPs can benefit from their light users' defections. Over time, however, as users' natural usage increases and there are fewer lighter users, fewer ISPs will want to partner with the vISP and fewer users will defect to the vISP, jeopardizing the vISP's profit. Thus, the vISP represents an economically viable interim solution for ISPs to increase user utilities until they can upgrade their network infrastructure to handle growing user demands. If demands continue to outstrip infrastructure growth, there may continue to be a viable place for the vISP in the mobile data market.

References

[1] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, "A survey of smart data pricing: Past proposals, current plans, and future trends," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, p. 15, 2013.

[2] Nokia, "Bell labs consulting report finds human demand for a digital future anywhere can only partially be met by networks in 2020," 2016. online at: https://goo.gl/ytmKLW.

[3] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "Rat selection games in hetnets," in *Proc. of IEEE INFOCOM*, 2013.

[4] M. Isaac, "AT&T drops its T-Mobile merger bid in \$4b fail." Wired, 2011. http://www.wired.com/2011/12/att-tmobile-merger-ends/.

[5] Google, "Project Fi," *online at: https://fi.google.com/*, 2015.

[6] J. Musacchio and J. Walrand, "WiFi access point pricing as a dynamic game," *IEEE/ACM Trans. Netw.*, vol. 14, no. 2, pp. 289–301, 2006.

[7] C. Joe-Wong, S. Sen, and S. Ha, "Offering supplementary network technologies: Adoption behavior and offloading benefits," *IEEE/ACM Trans. Netw.*, vol. 23, no. 2, pp. 355–368, 2015.

[8] S. Ren, J. Park, and M. Van Der Schaar, "Entry and spectrum sharing scheme selection in femtocell communications markets," *IEEE/ACM Trans. Netw.*, vol. 21, no. 1, pp. 218–232, 2013.

[9] J. Lee, Y. Yi, S. Chong, and Y. Jin, "Economics of wifi offloading: Trading delay for cellular capacity," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1540–1554, 2014.

[10] Y. Im, C. Joe-Wong, S. Ha, S. Sen, T. T. Kwon, and M. Chiang, "Amuse: Empowering users for cost-aware offloading with throughput-delay tradeoffs," in *Proc. of IEEE INFOCOM*, 2013.

[11] L. Zheng, C. Joe-Wong, C. W. Tan, S. Ha, and M. Chiang, "Secondary markets for mobile data: Feasibility and benefits of traded data plans," in *Proc. of INFOCOM*, 2015.

[12] S. Shakkottai and R. Srikant, "Economics of network pricing with multiple ISPs," *IEEE/ACM Trans. Netw.*, vol. 14, no. 6, pp. 1233–1245, 2006.

[13] V. Valancius, C. Lumezanu, N. Feamster, R. Johari, and V. V. Vazirani, "How many tiers?: pricing in the internet transit market," in *ACM SIGCOMM Computer Comm. Rev.*, vol. 41, pp. 194–205, 2011.

[14] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J.-P. Makela, R. Pichna, and J. Vallstron, "Handoff in hybrid mobile data networks," *IEEE Pers. Comm.*, vol. 7, no. 2, pp. 34–47, 2000.

[15] G. P. Pollini, "Trends in handover design," *IEEE Commun. Mag.*, vol. 34, no. 3, pp. 82–90, 1996.

[16] C. Sun, E. Stevens-Navarro, and V. W. Wong, "A constrained MDP-based vertical handoff decision algorithm for 4g wireless networks," in *Proc. of IEEE ICC*, 2008.

[17] R. Mahindra, H. Viswanathan, K. Sundaresan, M. Y. Arslan, and S. Rangarajan, "A practical traffic management system for integrated lte-wifi networks," in *Proc. of ACM MobiCom*, 2014.

[18] N. Feamster, L. Gao, and J. Rexford, "How to lease the Internet in your spare time," *ACM SIGCOMM Computer Comm. Rev.*, vol. 37, no. 1, pp. 61–64, 2007.

[19] S. Sengupta and M. Chatterjee, "An economic framework for dynamic spectrum access and service pricing," *IEEE/ACM Trans. Netw.*, vol. 17, no. 4, pp. 1200–1213, 2009.

[20] S. Layton, "Project fi cell phone plans," *online at: https://goo.gl/zAJfLF*, 2016.

[21] Cisco, "The zettabyte era–Trends and analysis," 2016. https://goo.gl/7aWBv2.

[22] Technical Report, "Economic viability of a virtual ISP," *online at: http://scenic.princeton.edu/sdp/papers/vISP.pdf*, 2016.

[23] K. Sheehy, "Best cell phone plans." NerdWallet, 2016. https://www.nerdwallet.com/blog/utilities/cell-phone-plans/.

[24] Open Signal, "Signal strength," *online at: http://opensignal.com/*, 2015.

APPENDIX

A. Proof of Lemma 1

Proof: The proof starts from the fact that by sharing capacity with vISP users, the network performance of any two partner ISPs $k, k' = 1, 2, \dots, K$ are the same:

$$\frac{C_k}{(1-\theta_k)\varphi_k N + \hat{n}_k} = \frac{C_{k'}}{(1-\theta_{k'})\varphi_{k'} N + \hat{n}_{k'}}$$

$$\overset{(a)}{\Rightarrow} \quad \big((1-\theta_{k'})\varphi_{k'} N + \hat{n}_{k'}\big) \sum_{k=1}^{K} C_k$$

$$= C_{k'} \sum_{k=1}^{K} (1-\theta_k)\varphi_k N + C_{k'} \sum_{k=1}^{K} \hat{n}_k$$

$$\overset{(b)}{\Rightarrow} \quad \hat{c} = \frac{\sum_{k=1}^{K} C_k}{\sum_{k=1}^{K} (1-\theta_k)\varphi_k N + \hat{N}}$$

$$\overset{(c)}{\Rightarrow} \quad \hat{c} = \frac{\sum_{k=1}^{K} C_k}{\Big(\sum_{k=1}^{K} \varphi_k + \sum_{m=K+1}^{M} \theta_m \varphi_m\Big) N}$$

where (a) is by summing both sides of the equation for all K partner ISPs, (b) is due to $\sum_{k=1}^{K} \hat{n}_k = \hat{N}$, and (c) is due to $\hat{N} = \sum_{m=1}^{M} \theta_m \varphi_m N$. ■

B. Proof of Proposition 1

Proof: Supposing ISP k' provides the highest QoS among all partner ISPs before sharing network infrastructure with the vISP, i.e., $\frac{C_{k'}}{\varphi_{k'} N} = \max_{k=1,\dots,K} \big\{ \frac{C_k}{\varphi_k N} \big\}$, we have $\frac{C_k}{C'_k} \varphi_{k'} \le \varphi_k, \ \forall k = 1, \dots, K$, leading to

$$\sum_{k=1}^{K} \varphi_k \le \sum_{k=1}^{K} \varphi_k + \sum_{m=K+1}^{M} \theta_m \varphi_m$$

$$\Rightarrow \quad \sum_{k=1}^{K} \frac{C_k}{C'_k} \varphi_{k'} \le \varphi_k \le \sum_{k=1}^{K} \varphi_k + \sum_{m=K+1}^{M} \theta_m \varphi_m$$

$$\Rightarrow \quad \frac{\sum_{k=1}^{K} C_k}{\Big(\sum_{k=1}^{K} \varphi_k + \sum_{m=K+1}^{M} \theta_m \varphi_m\Big) N} \le \frac{C_{k'}}{\varphi_{k'} N}.$$

The result can also be proved by the mediant inequality. ■

C. Proof of Corollary 1

Proof: Similar to the proof of Proposition 1, we suppose that before sharing network infrastructure with the vISP, partner ISP k' provides the highest QoS, i.e., $\frac{C_{k'}}{\varphi_{k'} N} = \max_{k=1,\dots,K} \big\{ \frac{C_k}{\varphi_k N} \big\}$, and partner ISP k'' provides the least QoS, i.e., $\frac{C_{k''}}{\varphi_{k''} N} = \min_{k=1,\dots,K} \big\{ \frac{C_k}{\varphi_k N} \big\}$. Thus, we combine $\frac{C_k}{C'_k} \varphi_{k'} \le \varphi_k$ and $\frac{C''_k}{\varphi_{k''}} \varphi_k \ge C_k, \ \forall k = 1, \dots, K$ with $\sum_{m=K+1}^{M} \theta_m \varphi_m \ge \Big(\max_{k=1,\dots,K} \big\{ \frac{C_k}{\varphi_k N} \big\} \big/ \min_{k=1,\dots,K} \big\{ \frac{C_k}{\varphi_k N} \big\} - 1 \Big) \sum_{k=1}^{K} \varphi_k$ to find that $\hat{c} \le \min_{k=1,\dots,K} \big\{ \frac{C_k}{\varphi_k N} \big\}$. ■

D. Proof of Proposition 2

Proof: Since the optimal utilities for both non-partner users and vISP users are piece-wise, we prove the result case by case. When $z_i \le d$, we find $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) \le \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ if $z_i \le \frac{\eta}{p}$. Thus, users with $z_i \le \frac{\eta}{p}$ will defect in any case.

When $d \le z_i \le (\hat{c}/\rho)^{1/\alpha}$, $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) - \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p) = (d\rho - \eta) - (\rho - p) z_i$ is decreasing. When $z_i \ge (\hat{c}/\rho)^{1/\alpha}$, $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) = \frac{\alpha}{1-\alpha} \rho^{1-\frac{1}{\alpha}} \hat{c}^{\frac{1}{\alpha}} - \eta + d\rho$, but $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ keeps increasing until $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p) = \frac{\alpha}{1-\alpha} p^{1-\frac{1}{\alpha}} \hat{c}^{\frac{1}{\alpha}}$. Since $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) - \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p) = (d\rho - \eta) - (\rho - p) z_i$ equals zero at $z_i = \frac{d\rho - \eta}{\rho - p}$ if $\frac{d\rho - \eta}{\rho - p} \le (\hat{c}/\rho)^{1/\alpha}$, we discuss the relationship between $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho)$ and $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ in three different cases below for $z_i \ge d$.

1) $\frac{d\rho - \eta}{\rho - p} \le \big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}$.

In this case, $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) - \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p) \le 0$ also holds for $z_i \ge \frac{d\rho - \eta}{\rho - p}$, so the defection rate is calculated by $\theta_k = 1 - \int_{\frac{\eta}{p}}^{\frac{d\rho - \eta}{\rho - p}} f^k(z) \mathrm{d}z$ where $\delta_k = \big(\frac{(1-\alpha)\eta\varphi_k N}{c_k}\big)^{\frac{1}{1-\alpha}}$. We then obtain the first expression in (7).

2) $\big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}} \le \frac{d\rho - \eta}{\rho - p} \le \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}}$.

Due to $\big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}} \le \frac{d\rho - \eta}{\rho - p} \le \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}}$, $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) - \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p) \ge 0$ always holds for $z_i \le \big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}$, and $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) = \frac{\alpha}{1-\alpha} \rho^{1-\frac{1}{\alpha}} \hat{c}^{\frac{1}{\alpha}} - \eta + d\rho$ would intersect with $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ at some point in $\Big[\big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}, \frac{d\rho - \eta}{\rho - p}\Big]$. To enable the analytical result, we approximate $z_i^{1-\alpha}$ using its Taylor series approximation at $z_i = \big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}$ up to the first order terms. By substituting $z_i^{1-\alpha} \approx \big(\frac{\hat{c}}{\rho}\big)^{\frac{1-\alpha}{\alpha}} + (1-\alpha)\big(\frac{\hat{c}}{\rho}\big)^{-1}\big(z_i - \big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}\big) + \mathcal{O}(z_i^2)$ into $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$, $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) \le \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$. leads to $z_i \ge \frac{d\rho - \eta}{\rho - p}$. Approximately, the second case yields the same result as the first case.

3) $\frac{d\rho - \eta}{\rho - p} \ge \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}}$.

Due to the convexity of the function $g(x) = x^{1-\frac{1}{\alpha}}$, we have $\rho^{1-\frac{1}{\alpha}} \ge p^{1-\frac{1}{\alpha}} + (1-\frac{1}{\alpha}) p^{-\frac{1}{\alpha}}(\rho - p)$. Combining this with $\frac{d\rho - \eta}{\rho - p} \ge \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}}$, we find $\frac{\alpha}{1-\alpha} \rho^{1-\frac{1}{\alpha}} \hat{c}^{\frac{1}{\alpha}} - \eta + d\rho \ge \frac{\alpha}{1-\alpha} p^{1-\frac{1}{\alpha}} \hat{c}^{\frac{1}{\alpha}}$, i.e., $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) \ge \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ for $z_i \ge \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}}$.

Furthermore, $\frac{d\rho - \eta}{\rho - p} \ge \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}} \ge \big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}$ implies that $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) - \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p) \ge 0$ always holds for $z_i \le \big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}$. For $z_i \ge \big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}}$, as $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ still increases while $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho)$ remains the same value, $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) \ge \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ at $z_i = \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}}$ ensures that $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho)$ is also larger than $\hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ in $\big(\frac{\hat{c}}{\rho}\big)^{\frac{1}{\alpha}} \le z_i \le \big(\frac{\hat{c}}{p}\big)^{\frac{1}{\alpha}}$.

Thus, in the third case, $U_i^k(\tilde{z}_i^\star \mid \hat{c}, d, \eta, \rho) \ge \hat{U}_i(\hat{z}_i^\star \mid \hat{c}, p)$ holds for $z_i \ge d$, and only users with $z_i \le \frac{\eta}{p}$ will defect, i.e., $\theta_k = \int_{\delta_k}^{\frac{\eta}{p}} f^k(z) \mathrm{d}z$. We obtain the second expression in (7).

Summarizing the above discussion, we find (7). ■

E. Proof of Lemma 2

Proof: Lemma 2 is equivalent to the statement that if user i defects, then this user must have a natural usage that is less than monthly cap, i.e., $z_i \le d$ for defected non-partner users. As given in (5), the highest possible utility for a vISP user is $\frac{\alpha}{1-\alpha}p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}}$ if this user has $z_i \ge (\frac{\hat{c}}{p})^{\frac{1}{\alpha}}$. Since $U_i^m(\tilde{z}_i^\star \,|\, \frac{C_m}{(1-\theta_m)\varphi_m N}, d, \eta, \rho) \ge U_i^m(\tilde{z}_i^\star = d \,|\, \frac{C_m}{(1-\theta_m)\varphi_m N}, d, \eta, \rho)$ for $z_i \ge d$, we show that $\frac{\alpha}{1-\alpha}p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}}$ is even smaller than the smallest utility $U_i^m(\tilde{z}_i^\star = d \,|\, \frac{C_m}{(1-\theta_m)\varphi_m N}, d, \eta, \rho) = \frac{C_m}{(1-\theta_m)\varphi_m N}\frac{d^{1-\alpha}}{1-\alpha} - \eta$ that a user with a natural usage larger than d can obtain from non-partner ISP m. Before doing so, we consider the function:

$$g(d) = -d^{1-\alpha} + (1-\alpha)(\hat{c}/p)^{-1}d + \alpha(\hat{c}/p)^{\frac{1}{\alpha}-1}$$

that is non-increasing in terms of d due to $\frac{\hat{c}}{p} \ge d^\alpha$. Thus, we find $g(d) \le g\big((\frac{\hat{c}}{p})^{\frac{1}{\alpha}}\big) = 0$. We now derive that

$$\begin{aligned}
& \alpha(\hat{c}/p)^{\frac{1}{\alpha}-1} + (1-\alpha)(\hat{c}/p)^{-1}d \le d^{1-\alpha} \\
\overset{(a)}{\Rightarrow} \quad & \alpha p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} + (1-\alpha)\eta \le \hat{c}d^{1-\alpha} \\
\overset{(b)}{\Rightarrow} \quad & \frac{\alpha}{1-\alpha}p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} \le \frac{C_m}{\varphi_m N}\frac{d^{1-\alpha}}{1-\alpha} - \eta \\
\overset{(c)}{\Rightarrow} \quad & \frac{\alpha}{1-\alpha}p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} \le \frac{C_m}{(1-\theta_m)\varphi_m N}\frac{d^{1-\alpha}}{1-\alpha} - \eta,
\end{aligned}$$

where (a) is due to $\eta/d \le p$, (b) is due to $\frac{C_m}{\varphi_m N} \ge \hat{c}$, and (c) is due to $\theta_m \in [0,1]$. ∎

F. Proof of Proposition 3

Proof: By Lemma 2, only users with $z_i \le d$ would defect to the vISP. Thus, we only need to compare $\hat{U}_i(\hat{z}_i^\star \,|\, \hat{c}, p)$ with $U_i^m(\tilde{z}_i^\star \,|\, c_i^m, d, \eta, \rho) = \frac{C_m}{(1-\theta_m)\varphi_m N}\frac{{z_i}^{1-\alpha}}{1-\alpha} - \eta$ for $z_i < d$. Since $\hat{U}_i(\hat{z}_i^\star \,|\, \hat{c}, p)$ is piecewise, our calculation consists of two step.

First, from $\frac{\alpha}{1-\alpha}p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} \ge \frac{C_m}{(1-\theta_m)\varphi_m N}\frac{{z_i}^{1-\alpha}}{1-\alpha} - \eta$, we obtain:

$$z_i \le \hat{z}^m = \left(\frac{\alpha p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} + (1-\alpha)\eta}{\frac{C_m}{(1-\theta_m)\varphi_m N}}\right)^{\frac{1}{1-\alpha}}, \tag{20}$$

which is combined with the Pareto-distributed user natural demands $\theta_m = 1 - (\frac{\delta_m}{\hat{z}^m})^{\lambda_m}$ and $\delta_m = (\frac{(1-\alpha)\eta\varphi_m N}{C_m})^{\frac{1}{1-\alpha}}$, and leads to (9). Substituting (9) back to (20), we find (8).

Next, we prove that with θ_m given in (9), we also have $\hat{c}\frac{{z_i}^{1-\alpha}}{1-\alpha} - z_i p \ge \frac{C_m}{(1-\theta_m)\varphi_m N}\frac{{z_i}^{1-\alpha}}{1-\alpha} - \eta$ for $z_i < (\frac{\hat{c}}{p})^{\frac{1}{\alpha}}$. Due to $\hat{c} \ge \frac{1}{\alpha\delta_m^{\lambda_m}}\frac{C_m}{\varphi_m N}$, we find

$$\alpha p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} + (1-\alpha)\eta \ge \delta_m^{-\lambda_m}\frac{C_m}{\varphi_m N}\Big(\frac{\hat{c}}{p}\Big)^{\frac{1}{\alpha}-1},$$

which leads to

$$\left(\frac{\delta_m^{-\lambda_m}\frac{C_m}{\varphi_m N}}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} + (1-\alpha)\eta}\right)^{\frac{-1+\alpha}{1-\alpha+\lambda_m}} \ge \left(\frac{\hat{c}}{p}\right)^{\frac{1}{\alpha}-1} \tag{21}$$

for $\lambda_m > 1$, i.e., $\frac{-1+\alpha}{1-\alpha+\lambda_m} > -1$. Finally, (21) is equivalent to

$$1 - \theta_m \ge \frac{C_m}{\varphi_m N}\left(\alpha\hat{c} + (1-\alpha)\eta\Big(\frac{\hat{c}}{p}\Big)^{1-\frac{1}{\alpha}}\right)^{-1},$$

leading to $\hat{c}\frac{{z_i}^{1-\alpha}}{1-\alpha} - z_i p \ge \frac{C_m}{(1-\theta_m)\varphi_m N}\frac{{z_i}^{1-\alpha}}{1-\alpha} - \eta$. ∎

G. Proof of Proposition 4

Proof: We first show the existence of a limit point. Taking a linear combination of the dynamics for each m, we conclude that $\sum_{m=K+1}^M \varphi_m\theta_m = \sum_{m=K+1}^M \Big(\varphi_m - \varphi_m\big(\frac{\eta(1-\alpha)}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} + (1-\alpha)\eta}\big)^{\frac{\lambda_m}{1-\alpha+\lambda_m}}\Big)$ at any limit point. Defining $\tau = \sum_{m=K+1}^M \varphi_m\theta_m$, we then have

$$\tau = \sum_{m=K+1}^M \left(\varphi_m - \varphi_m\Big(\frac{\eta(1-\alpha)}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}(\tau)^{\frac{1}{\alpha}} + (1-\alpha)\eta}\Big)^{\frac{\lambda_m}{1-\alpha+\lambda_m}}\right), \tag{22}$$

where we have written $\hat{c}$ in terms of τ instead of θ. We now note that the right-hand side of (22) is monotonically decreasing in τ, while the left-hand side is monotonically increasing. Thus, to show that (22) has a unique solution $\tau^\star$, it suffices to show that the right-hand side is less than $\sum_{m=K+1}^M \varphi_m$ at $\tau = 0$ and larger than 0 at $\tau = \sum_{m=K+1}^M \varphi_m$. Both are true by inspection.

We thus see that at a limit point, $\tau = \tau^\star$. We can thus solve for $\vec{\theta}^\star$ by writing

$$\theta_m^\star = 1 - \left(\frac{\eta(1-\alpha)}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}\,(\tau^\star)^{\frac{1}{\alpha}} + (1-\alpha)\eta}\right)^{\frac{\lambda_m}{1-\alpha+\lambda_m}}. \tag{23}$$

for each non-partner ISP m. It is clear that a unique solution to these equations exists, which determines a unique limit point of (10). To show that (10) converges to this unique limit point, we first show that the Jacobian $\mathrm{d}h/\mathrm{d}\vec{\theta}$ is a negative-definite matrix for any value of $\vec{\theta}$. Using the definition of τ from the proof of Proposition 4, we see that for $m \ne n$,

$$\begin{aligned}
\frac{\partial h_m}{\partial\theta_n} &= -\frac{\partial}{\partial\hat{c}}\left(\left(\frac{\eta(1-\alpha)}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}} + (1-\alpha)\eta}\right)^{\frac{\lambda_m}{1-\alpha+\lambda_m}}\right)\frac{\partial\hat{c}}{\partial\tau}\varphi_n \\
&= \frac{\partial g_m}{\partial\tau}\varphi_n
\end{aligned}$$

where we define

$$g_m(\tau) = -\left(\frac{\eta(1-\alpha)}{\alpha p^{1-\frac{1}{\alpha}}\hat{c}(\tau)^{\frac{1}{\alpha}} + (1-\alpha)\eta}\right)^{\frac{\lambda_m}{1-\alpha+\lambda_m}}.$$

Thus, we find that

$$\frac{\partial h_m}{\partial\theta_n} = \begin{cases} \frac{\partial g_m}{\partial\tau}\varphi_n - 1 & \text{if } m = n \\ \frac{\partial g_m}{\partial\tau}\varphi_n & \text{if } m \ne n \end{cases}$$

and the Jacobian $\mathrm{d}h/\mathrm{d}\vec{\theta}$ can be written as

$$J(\vec{\theta}) = \frac{\partial g}{\partial\tau}\vec{\varphi} - I,$$

where $\vec{\varphi}$ is the horizontal vector concatenating the φ_m for $m = K+1, \ldots, M$ and g is the vertical concatenation of the g_m. It is easy to see that, if μ is an eigenvalue of $J(\vec{\theta})$ for any fixed $\vec{\theta}$, then $1+\mu$ is an eigenvalue of $(\partial g/\partial\tau)\,\vec{\varphi}$. Thus, since this matrix has eigenvalues of 0 and $\vec{\varphi}\,(\partial g/\partial\tau)$, we see that $J(\vec{\theta})$ has eigenvalues of -1 and $\vec{\varphi}\,(\partial g/\partial\tau) - 1$, which are both negative since $\partial g_m/\partial\tau < 0$ and $\varphi_m > 0$ for any m. We have thus shown that $J(\vec{\theta})$ is negative-definite for any $\vec{\theta}$.

We now propose the Lyapunov candidate function

$$L\left(\vec{\theta}(t)\right)=\sum_{m=K+1}^{M} h_m\left(\vec{\theta}\right)^2. \tag{24}$$

It is easy to see that this function is nonnegative on $[0,1]^{M-K}$ and that it is zero if and only if $h_m=0$ for all m (i.e., at a limit point). We now take the time derivative of L to find that

$$\dot{L}=2\sum_{m=K+1}^{M} h_m(\vec{\theta})\left(\frac{\mathrm{d}h_m}{\mathrm{d}\vec{\theta}}f(\vec{\theta})\right)=h(\vec{\theta})^\top J(\vec{\theta})h(\vec{\theta}), \tag{25}$$

which, since $J(\vec{\theta})$ is negative-definite, is negative on $[0,1]^{M-K}$ except at the limit points where $h(\vec{\theta})=0$. Thus, L is a Lyapunov function for (10) on $[0,1]^{M-K}$. LaSalle's invariance principle allows us to conclude that the defection rates $\vec{\theta}$ converge to the largest invariant set S contained in $\left\{\theta|\dot{L}(\theta)=0\right\}$, or equivalently the set of points for which $h=0$. Since we have shown in Proposition 4 that there exists a unique such limit point, (10) converges to this point, $\vec{\theta}^\star$. ■

H. Proof of Proposition 5

Proof: Each partner ISP k's original revenue can be calculated by different types of user usage:

$$\mathcal{R}'_k=\left(\int_{\delta_k}^{d}\eta f_k(z)\mathrm{d}z-\int_{d}^{\left(\frac{C_k}{\rho\varphi_k N}\right)^{\frac{1}{\alpha}}}(\eta+(z-d)\rho)\,f_k(z)\mathrm{d}z\right.$$
$$\left.-\int_{\left(\frac{C_k}{\rho\varphi_k N}\right)^{\frac{1}{\alpha}}}^{\infty}\left(\eta+\left(\left(\frac{C_k}{\rho\varphi_k N}\right)^{\frac{1}{\alpha}}-d\right)\rho\right)f_k(z)\mathrm{d}z\right)\varphi_k N,$$

where users with usage below the cap $\delta_k\le z_i\le d$ pay the monthly fee, users with natural usage $d\le z_i\le\left(\frac{C_k}{\rho\varphi_k N}\right)^{\frac{1}{\alpha}}$ consume the exact amount of their natural usage and pay the monthly fee plus the overage $(z_i-d)\rho$, the rest heavy users maximize their utility and reduce their demands to $\left(\frac{C_k}{\rho\varphi_k N}\right)^{\frac{1}{\alpha}}$.

Since the set of users defecting from the partner ISP, (6), is a piece-wise function, we discuss the two cases that lead to different revenues for the partner ISP after partnering with the vISP. We start with the simpler one when $\frac{d\rho-\eta}{\rho-p}\ge\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}$:

$$\mathcal{R}''_k=\left(\int_{\frac{\eta}{p}}^{d}\eta f_k(z)\mathrm{d}z+\int_{d}^{\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}}(\eta+(z-d)\rho)\,f_k(z)\mathrm{d}z\right.$$
$$\left.+\int_{\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}}^{\infty}\left(\eta+\left(\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}-d\right)\rho\right)f_k(z)\mathrm{d}z\right)\varphi_k N.$$

Substituting $\theta_k=1-\left(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\right)^{\frac{\lambda_k}{1-\alpha}}\left(\frac{\eta}{p}\right)^{-\lambda_k}$ and $\delta_k=\left(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\right)^{\frac{\lambda_k}{1-\alpha}}$ into $\mathcal{R}''_k-\mathcal{R}'_k$ generates the first case in (11).

If $\frac{d\rho-\eta}{\rho-p}\le\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}$, heavy partner users with $z_i\ge\frac{d\rho-\eta}{\rho-p}$ also defect to the vISP and thus no loyal parter user needs to reduce their usage:

$$\mathcal{R}''_k=\left(\int_{\frac{\eta}{p}}^{d}\eta f_k(z)\mathrm{d}z+\int_{d}^{\frac{d\rho-\eta}{\rho-p}}(\eta+(z-d)\rho)\,f_k(z)\mathrm{d}z\right)\varphi_k N.$$

Substituting $\theta_k=1-\left(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\right)^{\frac{\lambda_k}{1-\alpha}}\left(\left(\frac{\eta}{p}\right)^{-\lambda_k}-\left(\frac{d\rho-\eta}{\rho-p}\right)^{-\lambda_k}\right)$ and $\delta_k=\left(\frac{(1-\alpha)\eta\varphi_k N}{C_k}\right)^{\frac{\lambda_k}{1-\alpha}}$ into $\mathcal{R}''_k-\mathcal{R}'_k$ generates the second case in (11).

Combining the above two cases together, we can obtain the result in (11).

When $\hat{c}\le\frac{C_k}{\varphi_k N}$, it is straightforward to see that $\left(\frac{C_k}{\rho\varphi_k N}\right)^{\frac{1-\lambda_k}{\alpha}}\le\left(\frac{\hat{c}}{\rho}\right)^{\frac{1-\lambda_k}{\alpha}}$ for $\lambda_k>1$. Thus, $\Delta\mathcal{R}_k(\theta_k,p)$ is negative for the case $\frac{d\rho-\eta}{\rho-p}\ge\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}$. On the other hand, if $\frac{d\rho-\eta}{\rho-p}\le\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}$, then $\left(\frac{\hat{c}}{\rho}\right)^{\frac{1-\lambda_k}{\alpha}}\le\left(\frac{d\rho-\eta}{\rho-p}\right)^{1-\lambda_k}$. The facts of $\rho>p$ and $d\rho>\eta$ lead to $\left(\frac{dp-\eta}{d\rho-\eta}+\frac{1}{\lambda_k-1}\right)\ge 1$. Thus, $\Delta\mathcal{R}_k(\theta_k,p)$ is negative in this case as well. ■

I. Proof of Proposition 6

Proof: Similar to the calculation of (11) but with a single case of user defection, the result in (15) is calculated by

$$\Delta\mathcal{R}_m(\theta_m,p)=\left(\int_{\hat{z}^m}^{d}\eta f_m(z)\mathrm{d}z\right.$$
$$+\int_{d}^{\left(\frac{C_m}{(1-\theta_m)\rho\varphi_m N}\right)^{\frac{1}{\alpha}}}(\eta+(z-d)\rho)\,f_m(z)\mathrm{d}z$$
$$+\int_{\left(\frac{C_m}{(1-\theta_m)\rho\varphi_m N}\right)^{\frac{1}{\alpha}}}^{\infty}\left(\eta+\left(\left(\frac{C_m}{(1-\theta_m)\rho\varphi_m N}\right)^{\frac{1}{\alpha}}-d\right)\rho\right)f_m(z)\mathrm{d}z$$
$$-\int_{\delta_m}^{d}\eta f_m(z)\mathrm{d}z$$
$$-\int_{d}^{\left(\frac{C_m}{\rho\varphi_m N}\right)^{\frac{1}{\alpha}}}(\eta+(z-d)\rho)\,f_m(z)\mathrm{d}z$$
$$\left.-\int_{\left(\frac{C_m}{\rho\varphi_m N}\right)^{\frac{1}{\alpha}}}^{\infty}\left(\eta+\left(\left(\frac{C_m}{\rho\varphi_m N}\right)^{\frac{1}{\alpha}}-d\right)\rho\right)f_m(z)\mathrm{d}z\right)\varphi_m N,$$

with the minimum usage of all ISP m's users, δ_m, substituted by $\delta_m=\left(\frac{(1-\alpha)\eta\varphi_m N}{C_m}\right)^{\frac{1}{1-\alpha}}$ and $\hat{z}^m=\left(\frac{(1-\alpha)\eta\varphi_m N}{C_m}\right)^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{1-\alpha}\left(\frac{p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}}}{(1-\alpha)\eta}\right)+1\right)^{\frac{1}{1-\alpha+\lambda_m}}$ following the condition derived in (8).

We then show that the condition in (16) leads to a nonnegative $\Delta\mathcal{R}_m$ by transforming it to:

$$\frac{1}{\alpha\eta}\left((1-\alpha)\eta\right)^{\frac{\lambda_m}{1-\alpha}}\rho^{\frac{\lambda_m+\alpha-1}{\alpha}}\le\rho^{\frac{\lambda_m+\alpha-1}{\alpha(1-\alpha)}}d^{\frac{\lambda_m+\alpha-1}{1-\alpha}}$$
$$\Rightarrow\quad\frac{1}{\alpha\eta}\left((1-\alpha)\eta\right)^{\frac{\lambda_m}{1-\alpha}}\rho^{\frac{\lambda_m+\alpha-1}{\alpha}}\left(\frac{C_m}{\varphi_m N}\right)^{\frac{1-\alpha-\lambda_m}{\alpha(1-\alpha)}}\ge 1 \tag{26}$$

due to $\frac{C_m}{\varphi_m N}\ge\rho d^{\alpha}$. By taking the first-order and second-order derivatives of $\Delta\mathcal{R}_m(\theta_m,p)$ with respect to θ_m, we find

$$\frac{\partial\Delta\mathcal{R}_m}{\partial\theta_m}\propto\frac{1}{\alpha}\left((1-\alpha)\eta\right)^{\frac{\lambda_m}{1-\alpha}}\rho^{\frac{\lambda_m+\alpha-1}{\alpha}}\left(\frac{C_m}{\varphi_m N}\right)^{\frac{1-\alpha-\lambda_m}{\alpha(1-\alpha)}}$$
$$\times(1-\theta_m)^{\frac{\lambda_m-1-\alpha}{\alpha}},$$

and

$$\frac{\partial^2\Delta\mathcal{R}_m}{\partial\theta_m^2}\propto\frac{\alpha+1-\lambda_m}{\alpha^2}\left((1-\alpha)\eta\right)^{\frac{\lambda_m}{1-\alpha}}\rho^{\frac{\lambda_m+\alpha-1}{\alpha}}$$
$$\times\left(\frac{C_m}{\varphi_m N}\right)^{\frac{1-\alpha-\lambda_m}{\alpha(1-\alpha)}}(1-\theta_m)^{\frac{\lambda_m-1-2\alpha}{\alpha}}.$$

Thus, $\lambda\le(\alpha+1)$ ensures the convexity of $\Delta\mathcal{R}_m$ in terms of θ_m, and $\lambda_m\le\frac{(1-\alpha)(\log(d\rho)-\log(\alpha\eta))}{\log(d\rho)-\log((1-\alpha)\eta)}$ (or (26)) ensures that $\Delta\mathcal{R}_m$ has a critical point satisfying $\theta_m^\star \le 0$. Due to $\Delta\mathcal{R}_m|_{\theta_m=0} = 0$, we conclude that $\Delta\mathcal{R}_m$ increases and is nonnegative in $\theta_m \in [0,1]$. ■

J. Proof of Corollary 2

Proof: We prove that $\delta_m \ge (\alpha\eta/\rho)$ leads to the same inequality in (26). Due to $\lambda_m > 1$, we find that

$$\delta_m^{\lambda_m} \ge \alpha\eta\rho^{\frac{1-\lambda_m-\alpha}{\alpha}}\left(\frac{C_m}{\varphi_m N}\right)^{\frac{\lambda_m-1}{\alpha}}$$

Substituting $\delta_m = \left(\frac{(1-\alpha)\eta\varphi_m N}{C_m}\right)^{\frac{1}{1-\alpha}}$ into the above inequality results in an inequality that is equivalent to (26). ■

K. Proof of Proposition 7

Proof: To prove the positivity of the optimal value for (17), we only need to find a feasible point that makes the objective positive. Thus, we exam the case when $p \to \rho$, i.e., $\frac{d\rho-\eta}{\rho-p} \ge \left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}$. Also, in this case, since only partner users with $z_i \le \frac{\eta}{p}$ will defect to vISP, the profit for the vISP can then be calculated by

$$\begin{aligned}
&p\mathcal{D}(p) + \sum_{k=1}^{K}\Delta\mathcal{R}_k(\theta_k, p)\\
&= p\Bigg(\sum_{k=1}^{K}\frac{\lambda_k}{\lambda_k-1}\delta_k^{\lambda_k}\left(\delta_k^{-\lambda_k+1} - \left(\frac{\eta}{p}\right)^{-\lambda_k+1}\right)\varphi_k N\\
&\quad + \sum_{m=K+1}^{M}\frac{\lambda_m}{\lambda_m-1}\delta_m^{\lambda_m}\left(\delta_m^{-\lambda_m+1} - \hat{z}_m^{-\lambda_m+1}\right)\varphi_m N\Bigg)\\
&\quad + \sum_{k=1}^{K}\Bigg(\eta\left(\left(\frac{\delta_k}{\eta/p}\right)^{\lambda_k} - 1\right)\\
&\quad + \frac{\rho}{\lambda_k-1}\delta_k^{\lambda_k}\left(\left(\frac{C_m}{\rho\varphi_k N}\right)^{\frac{1-\lambda_k}{\alpha}} - \left(\frac{\hat{c}}{\rho}\right)^{\frac{1-\lambda_k}{\alpha}}\right)\Bigg)\varphi_k N,
\end{aligned} \tag{27}$$

where $\hat{z}^m = \left(\frac{(1-\alpha)\eta\varphi_m N}{C_m}\right)^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{1-\alpha}\left(\frac{p^{1-\frac{1}{\alpha}}\hat{c}^{\frac{1}{\alpha}}}{(1-\alpha)\eta}\right) + 1\right)^{\frac{1}{1-\alpha+\lambda_m}}$ follows the result derived in (8).

We then rewrite (27) as

$$\mathcal{D}(p) + \sum_{k=1}^{K}\Delta\mathcal{R}_k(\theta_k, p) = \sum_{k=1}^{K} g_k(p)\varphi_k N + \phi(p)N,$$

with

$$g_k(p) = \frac{\lambda_k}{\lambda_k-1}p\delta_k - \frac{1}{\lambda_k-1}\eta\left(\frac{\delta_k}{\eta/p}\right)^{\lambda_k} - \eta,$$

and

$$\begin{aligned}
\phi(p) = &\sum_{m=K+1}^{M}\frac{p\lambda_m}{\lambda_m-1}\delta_m^{\lambda_m}\left(\delta_m^{-\lambda_m+1} - \hat{z}_m^{-\lambda_m+1}\right)\varphi_m\\
&+ \sum_{k=1}^{K}\frac{\rho}{\lambda_k-1}\delta_k^{\lambda_k}\left(\left(\frac{C_m}{\rho\varphi_k N}\right)^{\frac{1-\lambda_k}{\alpha}} - \left(\frac{\hat{c}}{\rho}\right)^{\frac{1-\lambda_k}{\alpha}}\right)\varphi_k.
\end{aligned}$$

We find $g_k(p) \ge 0$ due to the condition in (18) and $\delta_k \le (\eta/\rho)$. We then prove that $\phi(p)$ is also larger than 0. First, due to the convexity of $x^{1-\lambda}$ for $\lambda > 1$, we find

$$\begin{aligned}
\phi(p) \ge &\sum_{m=K+1}^{M} p\lambda_m\delta_m^{\lambda_m}\hat{z}_m^{-\lambda_m}(\hat{z}_m - \delta_m)\varphi_m\\
&+ \sum_{k=1}^{K}\frac{1}{\alpha}\rho\delta_k^{\lambda_k}\left(\frac{\hat{c}}{\rho}\right)^{\frac{1-\lambda_k}{\alpha}-1}\left(\frac{\hat{c}}{\rho} - \frac{C_k}{\rho\varphi_k N}\right)\varphi_k
\end{aligned} \tag{28}$$

If $\frac{\hat{c}}{\rho} \ge \frac{C_m}{\rho\varphi_k N}$, (28) holds; otherwise, due to $\frac{\delta_m}{\hat{z}_m} \le \frac{\delta_k}{(\hat{c}/\rho)^{1/\alpha}} < 1$, we need to prove a necessary condition for (28) that

$$\sum_{m=K+1}^{M} p\lambda_m(\hat{z}_m - \delta_m)\varphi_m \ge \sum_{k=1}^{K}\rho\frac{1}{\alpha}\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}-1}\left(\frac{C_k}{\rho\varphi_k N} - \frac{\hat{c}}{\rho}\right)\varphi_k \tag{29}$$

for $\frac{\hat{c}}{\rho} \le \frac{C_m}{\rho\varphi_k N}$. Starting from combining the condition in (19) with $\sum_{m=K+1}^{M}\varphi_m = 1 - \sum_{k=1}^{K}\varphi_k$, and $\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}} \le \frac{d\rho-\eta}{\rho-p}$ we find

$$\frac{\alpha^2}{(1-\alpha)^2}dp\sum_{m=K+1}^{M}\varphi_m \ge \frac{2}{1-2\alpha}\rho\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}}\sum_{k=1}^{K}\varphi_k. \tag{30}$$

Then, due to $\delta_m \ge (\alpha\eta/\rho)$ derived in Corollary 2, $\left(\frac{\hat{c}}{\rho}\right)^{\frac{1}{\alpha}} \ge d$, $\alpha \in [0,1)$ and $\lambda_m > 1$, the left-hand side of (30) is smaller than the left-hand side of (29). Furthermore, the right-hand side of (29) is maximized when $\frac{\hat{c}}{\rho} = (1-2\alpha)\frac{C_k}{\rho\varphi_k N}$, so the right-hand side of (30) is larger than the right-hand side of (29). Thus, under the condition in (19), (30) leads to (29) as well as (28).

Finally, we conclude that under the conditions in Proposition 7, the objective in (17) can be positive in its feasible set. ■