# ABSTRACT

Title of thesis:     INTERACTION BETWEEN A SERVICE
PROVIDER AND CONTENT PROVIDERS
AND SOCIAL EFFICIENCY

Ramy ElDelgawy, Master of Science, 2014

Thesis directed by:   Professor Richard J. La
Department of Electrical and Computer Engineering
University of Maryland, College Park

It is unclear how the introduction of "fast lanes" would affect the evolution of the Internet. While this is with no doubt an important issue, we do not aim to address this question. Instead, in this work we study the interaction between an internet service provider and one or more content providers when such fast lanes are allowed. Each entity is considered a player in a game, and it is assumed that these players are both rational and selfish, i.e., they aim to maximize their own gains. Clearly, their objectives are different, and their bargaining powers may vary. We examine how they may reach an agreement and how their interaction affects the resulting quality-of-service on the fast lanes. To get a better understanding of the kind of contract these entities may be willing to sign we look at two cases: (i) a monopoly market that has one internet service provider and one content provider and (ii) one where an extra content provider enters the market.

Our findings show that, in the case of a monopoly market, an efficient contract

can be reached in Stackelberg games, which maximizes the aggregate payoff of both the providers. Similarly, when another content provider enters the market, the entity with the strongest bargaining power chooses a contract that maximizes the aggregate payoff of all providers.

# INTERACTION BETWEEN A SERVICE PROVIDER AND CONTENT PROVIDERS AND SOCIAL EFFICIENCY

by

Ramy ElDelgawy

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2014

Advisory Committee:
Professor Richard J. La, Chair/Advisor
Professor Mark A. Shayman
Professor Michael C. Rotkowitz

*"The God of heaven, he will prosper us; therefore we his servants will arise and build"*

**—Nehemiah 2:20**

# Acknowledgments

It would be hard to appropriately thank all the beloved people who stood by me and helped me through my journey.

First and foremost I'd like to thank my advisor, Professor Richard La for giving me an invaluable opportunity to work with him. He mentored me through a topic that I had very few experience in, he was both kind and patient during that course. His passion to research is admirable and I definitely learned a lot from him professionally and, most importantly, personally.

I would like to thank Professor Michael Rotkowitz and Professor Mark Shayman for serving in my M.Sc. thesis committee. I am especially thankful for Professor Mark Shayman for giving me the opportunity to work with him as a grader for his course. I am also thankful to Professor Robert Newcomb, whom I assisted in teaching two courses.

I would also like to thank my professors back in Alexandria University in Egypt and Professor Kosuke Sato in Osaka University, Japan for without them I would not be having this opportunity of studying at UMD. Finally, I would like to thank Mrs. Melanie Prange for her continuous help during my stay here at UMD.

I owe myself and everything that I am, to my beloved family who provide me – endlessly and unconditionally – with love, energy and prayers. Words cannot express the gratitude I owe them. Thank you for always believing in me even when I did not believe in myself.

I owe a special and a personal thank you message to all the people who en-

# Contents

## Chapter 1:   Introduction

### 1.1   Internet Neutrality

Network neutrality, a term coined by Columbia media law professor Tim Wu in 2003 [18], has been a hot topic for the past decade [11,12,19]. Regulators, legislators and courts have been involved, and wide media and advocacy campaigns occurred. Moreover, scholarly interest from different disciplines emerged to study its effects. The regulatory and judicial contention among network users and access providers has been centered on the Internet service provider's (ISP) existing or potential authorities, which include discriminatory traffic management, i.e., traffic generated by varying Internet applications receives different treatment for the purpose of managing the network. Proponents of net neutrality have advocated for some form of regulatory intervention as they believe that any form of traffic discrimination gives network operators the final say on which applications would succeed or fail. At the other end of the spectrum, some prefer an open market where competition dictates the behavior of network operators and argue that any form of regulation would constrain the future development of broadband.

While the Internet is global, the debate over net neutrality is, however, specific to each country and also to each state. The debate started in the U.S, and has

extended to other countries in Europe and Asia but with distinct differences. On May 15, 2014, the Federal Communications Commission (FCC) in the U.S. decided to consider two options regarding Internet services: first, permit the presence of different broadband lane speeds, which in its essence contradicts net neutrality; and second, reclassify broadband as a telecommunications service, and thus preserving net neutrality.

In early 2014, Comcast and Netflix signed an agreement that allows Netflix to stream their content directly through the Comcast network for an undisclosed fee. The key idea behind this agreement between an Internet service provider (ISP), namely Comcast, and a content provider (CP), namely Netflix, is to bypass the "congested" Internet and to improve the quality-of-service (QoS) experienced by the Netflix streaming users. Soon afterward, another ISP, Verizon, signed a contract with Netflix as well though it is unclear what the terms of this contract are and how similar it is to the one signed with Comcast. The idea of fast lanes could be exemplified in Figure 1.1.

The debate between whether to regulate the ISPs or leave competition take its course and the legal and economic policies to be taken by the regulatory body were triggers to the net neutrality discussions in the academia and among policy makers. In reality though, due to the lack of consistent and coherent messages – and sometimes unrealistic assumptions – found in academia, these findings have not found an impact on the decisions of network operators and regulators. If fast lanes are to be allowed in the future, it is important to understand how the ISPs and other content/service providers would behave and how it would shape the ensuing

*Fig. 1.1:* The express lane on Southbound I-60 is a good example of what is proposed by the FCC. Cars could avoid traffic by going on an express lane in return for a fee.

economic and social benefits of the Internet.

The rest of this chapter discusses the concepts and terminologies of traffic management that would aid us in the rest of this study. In Chapter 2 we review related literature. Chapters 3 and 4 deal with a market where a single ISP interacts with a single CP and two CPs, respectively. We conclude in Chapter 5.

## 1.2  Forms of Traffic Management

Performance or quality-of-service (QoS) is often the key word used in the net neutrality debate. At the network level, the QoS may refer to the congestion level, delay, or delay jitter that network traffic experiences. At the user level, the QoS may be measured by the user's satisfaction level from the service. For example, the user's satisfaction level can be determined by streaming video speed and delays, the clarity of voice calls made over the internet, and file download speed. We use the QoS as one of parameters stipulated in the contract signed between an ISP and a CP, which is required to be provided by the ISP.

The ISP can offer varying QoS using different means. These include packet filtering (i.e., packets from a certain source are not delivered), rate limiting (i.e., outgoing rate of traffic is artificially limited), and differentiation (i.e., certain traffic is given a "fast lane", while other traffic is sent on the "regular lane"). In this thesis, we focus on the differentiation via fast lanes on the basis of contracts signed between an ISP and a CP.

The network is comprised of physical links (wires, cables, and fibers) and the

network devices that connect them (e.g., switches and routers). Increasing network capacity requires investments for (i) replacing existing links with faster links or deploying additional ones, (ii) upgrading network equipment to handle higher link speeds, and/or (iii) installing additional storage devices, e.g., caches and content storage space. Storage devices, such as caches, can be used in a network to reduce the delays experienced by end users by (temporarily) storing popular objects at the edge of the network closer to end users, thereby improving performance. We examine how the ISP's investments in its network are influenced by the relative bargaining positions of involved parties at an equilibrium and the agreed contract duration in different settings we consider in this thesis.

Chapter 2:    Existing literature on net neutrality

The Internet was made available for commercial use in the late 1980s. During that time, the Internet was categorized as an *information service*. However, by the end of the 1990s and early 2000s, new internet devices and services began to emerge, making the Internet more than just an information service. Since then, the U.S. regulators, independent ISPs, CPs and advocate groups have been in a constant battle over the issue of net neutrality.

Loosely speaking, net neutrality refers to the network operation where all data packets in the network are treated equally without discrimination. There are different interpretations of net neutrality; one interpretation is that network operators should not distinguish packets on the basis of their origins. Under this interpretation, the network operator would not be able to charge the content provider a termination fee without violating net neutrality. Another interpretation is that traffic belonging to different applications should not be differentiated or given priority over others.

Different possible interpretations of net neutrality and an increasing call for allowing "fast lanes" sparked academic interests on this topic, which started off as studies of its legal and economical implications by legal and economics scholars. However, later on it also attracted researchers from other disciplines as well.

Despite the growing research interest in net neutrality, there has not been a formal economic model adopted by researchers. In spite of recent progresses, it is still hard to rely on simple theoretical models to draw meaningful concrete conclusions on this dynamic and complex issue. In many of the existing studies, results are often ambiguous and model parameters are unrealistic.

In this thesis, we adopt the second interpretation of net neutrality mentioned earlier. The recent FCC's decision to consider various broadband lane speeds in May 2014. However, we briefly summarize the literature pertaining to either of the aforementioned interpretations.

## 2.1  Discrimination by traffic origin: Termination Fees

The Internet can be viewed as a two-sided market where ISPs provide a platform of interaction between CPs and consumers. Consumers and CPs pay their respective ISPs for access to the internet. However, ISPs do not collect extra fees from CPs for delivering content to the consumers, called *termination fees* [17]. Two central questions arise in this kind of market: What type of pricing mechanism is more likely to be adopted by ISPs to maximize their own profits? And, how does the introduction of this pricing mechanism by ISPs affect the interaction between CPs and consumers and how does it affect the overall social welfare?

The pricing imposed on both sides of the market affects the number of users subscribing to the broadband service on one side and the number of content providers and applications on the other side. In [20], users on the same side of the market

could have different membership values. But, their interaction with users from the other side is the same. In [21], there are no membership values. However, usuage fees varies according to the interactions between agents on different sides of the platform.

Economides and Tåg [22] build on the two-sided market model employed in [20] to examine the effect of departing from net neutrality by allowing the ISPs to charge CPs a fee in addition to the price they charge their subscribers. They study the case of a monopoly and duopoly ISP market and suggest that imposing net neutrality always enhances social welfare when the market is fully covered, i.e., everyone has Internet access, under the condition that the CPs' valuation of an extra consumer connecting to the Internet is greater than the consumers' valuation of an extra CP connecting to the Internet.

Instead of charging a fee, an ISP could also subsidize CPs to increase its value to consumers. In this case, the model discussed above would decrease the total surplus. Lee and Wu [17] argue that one possible explanation for the fact that ISPs, at this day and time, do not charge CPs with termination fees is that it would lead to lower content variety over the Internet which causes consumer discontent. Therefore, ISPs do not find it in their interest to charge these kind of fees.

Musacchio *et al.* [23] study the interaction between multiple CPs and multiple ISPs connecting the CPs to consumers. The CPs collect profits from advertising revenues arising from consumer clicks. Their finding indicates that ISPs overcharge CPs as compared to the social optimum. An ISP charging a termination fee to CPs have a negative effect on the overall market as consumers are less willing to pay

due to reduced availability of Internet content. This effect is more prominent as the number of ISPs in the market increases. They also make these suggestions for the following two cases:

Case 1: In the case that the CP's advertising revenue is low, it might be desirable to offer negative termination fees by the ISP to the CP to increase the CP's incentive to invest in its content;

Case 2: If the number of the ISP subscribers is low or the CP's advertising revenue is high, the ISP charges the CP with positive termination fees, so that the ISP would have an incentive to invest in its own network.

A related problem in which the CPs offer sponsored contents, e.g., advertisements, is studied by Andrews *et al.* [1] where the interaction between CPs and SPs is modelled as a Stackelberg game. The service provider offers a pricing schedule and the CP responds with how much content it wants to sponsor. They suggest that CPs and service providers are better off coordinating together to maximize the total system profit, assuming that the additional profit from cooperation can be split between both entities in an arbitrary manner.

Liang and Wang [15] study a similar problem but with a market that has heterogeneous CPs, which interact with a monopolistic ISP. Their finding indicates that none of the CPs may be willing to sponsor their data if the ISP discriminates the price of sponsored data between CPs according the CPs' net worth. Then, they consider a competition between one small CP and one large (or rich) CP. Their result shows that, on the short run, the smaller CP would benefit from sponsoring its data unlike the larger CP. In the long run, however, the larger CP will gain more

market share if it chooses to sponsor its content.

## 2.2   Discrimination by traffic content: Various lane speeds

As discussed before, the second interpretation for net neutrality implies that the traffic generated by various applications would be treated equally. On the other hand, traffic discrimination would mean that either ISPs would provide CPs with varying QoS for different prices or traffic is managed by the ISP in a certain manner without giving CPs a say on how their traffic is to be handled.

### 2.2.1   ISPs offering a range of qualities

Hermalin and Katz [8] analyze a model where a monopolistic SP provides a continuum of QoS. The CPs choose the QoS they would like to receive and pay a corresponding price. They compare this scenario to the case where a single level of QoS is provided by the SP, as proposed by the net neutrality regulation. Not surprisingly, their findings indicate that CPs that would select lower QoS (resp. higher QoS) are almost always harmed because they would be driven out of the market (resp. utilize less efficient, lower quality services), which argue against the net neutrality. On the other hand, CPs in the middle of the market benefit from a single QoS level. They also study the duopoly case and suggest that the same results apply as in the monopoly case. If the ISP is required to provide the highest possible quality, the charge it would require would be unattractive for content providers that prefer lower QoS. Thus, the content available to consumers would be reduced. On

the other hand, not allowing the ISP to collect fees from CPs would force ISPs to provide a single level of QoS that is inefficient.

Cheng *et al.* [5], Choi and Kim [6] and Krämer and Wiewiorra [10] study the effects of network congestion. They use the concepts of queuing theory in operations research to study the effects of traffic discrimination in the market. In Choi and Kim [6], "fast lanes" are sold to the CP with the highest bid, ensuring that its traffic would receive higher QoS over that of other CPs in the market. Their results show that the absence of discrimination leads to a higher social welfare. However, they were not able to confirm similar findings in the long run. Cheng *et al.* [5] consider a similar approach to the problem; a key difference is that they provide the premium lane service to all CPs at the same price. Clearly, if all CPs pay for the priority service, no CP enjoys better QoS over the other CPs. Their results suggest that, contrary to ISPs' claim, net discrimination leads to little incentive for ISPs to invest in their network capacity. In short, net discrimination benefits the ISPs, but hurts the CPs.

In Krämer and Wiewiorra [10], CPs do not compete with each other. But, their advertised content is sensitive to waiting times. Their model also assumes that there is minimum QoS provided to those CPs who choose not to get a priority service. Under these assumptions, their results show that overall welfare is improved in the short run as well as in the long run.

### 2.2.2   Vertical Foreclosure

Many ISPs have their own content to provide which may compete with that of other CPs offering similar service. In that case, ISPs may purposely degrade other CPs' services in order to promote their own service. Chen and Nalebuff [24] study the case with two firms in the market; one provides an essential service and the other offers a non-essential service. Their finding suggests that if the firm providing the essential service decides to enter the other firm's market and compete, it has no incentive to degrade the other firm's service as it collects most of its revenue from its essential product. However, their results seem a bit too optimistic as they are based on a critical assumption that the non-essential good has a low value. Empirical data in real life show a different result. This can be seen in cable operators and the programs that they air, and in mobile companies blocking Viber and Skype services on their 3G network, for example, as in some Middle Eastern and European countries.

## 2.3   Other Practices

Kocsis and de Bijl [25] and Economides and Tåg (2012) [22] express their concern over ISPs striking exclusive deals with CPs. They suspect that these exclusive deals would eventually lead to the monopolization of the market. They argue that a regulatory body should play an important role in such situations.

# Chapter 3:   Interaction Between a Single SP and a Single CP

The problem of interest to us can be viewed as one of task delegation [14]. The task delegation problem is often studied in the context of contract design [3]. One issue that attracted much attention in task delegation is asymmetry of information, which arises when one party has private information that the other party cannot observe. While this could happen even in our problem, here we assume that there is no information asymmetry and we can model the problem as a complete-information game. For an interested reader, we refer to [13] for a study that examines the consequence of information asymmetry in a related setting.

## 3.1   Setup and System Model

We focus on a scenario with only a single (internet) SP, e.g., Comcast or Verizon, and a single CP, e.g., Netflix or Amazon Prime. Subscribers to the SP (resp. the CP) pay a subscription fee $f_{SP}$ (resp. $f_{CP}$). We assume that the subscribers to the CP must also have internet service as well, hence subscribe to the SP.

We assume that, in the absence of any other incentives, the SP only provides certain *minimum* QoS, which we denote by $Q_{\min}$. Typically, the performance of streaming service deteriorates during the peak hours when the Internet is congested.

For this reason, the CP may be interested in convincing the SP to improve the QoS experienced by streaming traffic during congestion.

One possible means of achieving this goal is via a contract; it stipulates the QoS that the SP is required to deliver to the streaming traffic, and the SP collects a payment from the CP in return for the higher QoS enjoyed by the streaming service. Here, we consider a class of contracts given by a pair $(Q, p)$, where $Q$ is the QoS to be provided to the streaming traffic of the CP, and $p$ is the price that the CP will pay per CP subscriber to the SP.[1] Throughout this chapter, we assume that, even when the SP and the CP agree to a contract, the SP continues to provide the minimum QoS $Q_{\min}$ to other non-streaming traffic. Thus, only the streaming traffic of the CP receives (higher) QoS stipulated in the contract.

In order to capture the payoffs of both the SP and the CP, we introduce the following notation. Let $Q$ denote the QoS offered to the streaming service by the SP. We assume that $Q \in \mathcal{Q} := [Q_{\min}, \ Q_{\max}]$ for some maximum QoS $Q_{\max}$ satisfying $Q_{\min} < Q_{\max} < \infty$. Let $N_{CP}(t; Q)$ and $N_{SP}(t; Q)$ be the number of subscribers for CP and SP, respectively, at time $t \in \mathbb{R}_+ := [0, \infty)$ when the streaming service QoS is $Q$. The duration of the contract, if any is signed, is denoted by $M \in (0, \infty)$.

**i. Revenues:** The payoff of the CP and the SP from their revenue over the interval $[0, \ M]$, i.e., the contract duration, when the SP offers the streaming service

---

[1] While we assume that $p$ is a price per CP subscriber for convenience, we can also consider a price per unit bandwidth, assuming that CP subscribers are homogeneous.

QoS $Q$ is given by[2]

$$U_{CP,R}(Q) = \int_0^M N_{CP}(t;Q) \, f_{CP} \, e^{-\tau t} \, dt \quad \text{and} \tag{3.1}$$

$$U_{SP,R}(Q) = \int_0^M N_{SP}(t;Q) \, f_{SP} \, e^{-\tau t} \, dt, \tag{3.2}$$

where $\tau > 0$ is the discount rate that captures the *time value of money*, i.e., a dollar today is worth more than a dollar tomorrow [4].[3]

**ii. SP investment costs:** When the SP wishes to provide a certain level of QoS $Q \in \mathcal{Q}$, it has to invest in the network capacity and/or infrastructure and the required investment depends on (i) the QoS $Q$ and (ii) the increase in streaming traffic. We model this investment using the cost function $c : \mathcal{Q} \times [0, M] \to \mathbb{R}_+$, where $c(Q, t)$ denotes the *additional* investment necessary to serve a new CP subscriber joining at time $t$ with QoS $Q$ (as opposed to QoS $Q_{\min}$).

The investment needed to maintain the minimum QoS $Q_{\min}$ will be made by the SP whether a contract with the CP is signed or not. For this reason, we do not explicitly model this investment and only focus on the additional investment needed to provide the agreed QoS $Q$ in the contract for the streaming traffic. For this reason, we assume that $c(Q_{\min}, t) = 0$ for all $t \in [0, M]$.

The total cost of the SP for this investment over the contract duration is given by

$$C_{SP}(Q) = c(Q, 0) \, N_{CP}(0; Q) + \int_0^M c(Q, t) \, \frac{\partial}{\partial t} N_{CP}(t; Q) \, dt, \tag{3.3}$$

---

[2] To facilitate our analysis, we assume that the subscription fees $f_{CP}$ and $f_{SP}$ do not change during the duration of the contract, and leave the case where the CP can change the subscription fee as a function of QoS $Q$ for a future study.

[3] Throughout the paper, we assume that all fees and costs are adjusted for inflation over time.

where $\partial N_{CP}(t; Q)/\partial t$ is the rate at which the number of CP subscribers increases at time $t$ when QoS $Q$ is provided to its subscribers.

Note that the investment in (3.3) consists of two parts; the first is the up-front investment needed to improve the QoS for existing CP subscribers from $Q_{\min}$ to $Q$. The second accounts for the additional investment required to handle new CP subscribers over the duration of the contract. Since the initial number of CP subscribers $N_{CP}(0; Q)$ does not depend on $Q$, we denote it by $N_{CP}(0)$ when convenient.

**iii. Payment by the CP to the SP:** When the CP and the SP sign a contract $\Gamma = (Q, p)$ with $Q > Q_{\min}$, the CP must pay the SP for the improved QoS provided to the streaming traffic by the SP according to the contract $\Gamma$. The payoff of the CP and the SP from this payment is given by

$$U_{CP,P}(\Gamma) = -\int_0^M p \, N_{CP}(t; Q) \, e^{-\tau t} \, dt \text{ and} \tag{3.4}$$

$$U_{SP,P}(\Gamma) = \int_0^M p \, N_{CP}(t; Q) \, e^{-\tau t} \, dt \tag{3.5}$$

$$= -U_{CP,P}(\Gamma),$$

where $\tau$ is the same discount rate introduced in (3.1) and (3.2) to model the time value of money.

Given a contract $\Gamma = (Q, p)$ the providers agree to, we define their *payoffs* using (3.1) through (3.5) as follows:

$$U_{CP}(\Gamma) = U_{CP,R}(Q) + U_{CP,P}(\Gamma) \text{ and} \tag{3.6}$$

$$U_{SP}(\Gamma) = U_{SP,R}(Q) + U_{SP,P}(\Gamma) - C_{SP}(Q). \tag{3.7}$$

The *aggregate payoff* of the providers, which we denote by $U_{ag}$, is equal to the sum

16

of the payoffs of the providers, i.e.,

$$U_{ag}(\Gamma) = U_{CP}(\Gamma) + U_{SP}(\Gamma)$$

$$= U_{CP,R}(Q) + U_{SP,R}(Q) - C_{SP}(Q). \tag{3.8}$$

Note that the aggregate payoff depends only on the QoS $Q$, but not on the price $p$ because $U_{CP,P}(\Gamma) = -U_{SP,P}(\Gamma)$ and they cancel each other out in the sum. For this reason, with a little abuse of notation, we write the aggregate payoff as $U_{ag}(Q)$ when there is no confusion.

In order to facilitate our analysis, we introduce the following simplifying assumption.

**Assumption 3.1.** (a) The rate at which the number of CP subscribers increases, i.e., $\partial N_{CP}(t;Q)/\partial t$, is given by a *concave* and *strictly increasing* mapping $\zeta : \mathcal{Q} \to \mathbb{R}$. In other words, for fixed QoS $Q$ for streaming service, $\partial N_{CP}(t;Q)/\partial t = \zeta(Q)$.

(b) Similarly, the rate at which the number of SP subscribers increases is given by a *concave* and *strictly increasing* function $\lambda : \mathcal{Q} \to \mathbb{R}$, e.g., $\lambda(Q) = \zeta(Q) + \delta$ for some $\delta \geq 0$.

The interpretation of Assumption 3.1(b) is that, in addition to the new subscribers for streaming service, there may be a separate group of new subscribers to the SP solely for access to the Internet without subscribing to streaming service.

In general, it is expected that the cost of providing a certain level of QoS will diminish over time due to advances in technologies. The following assumption captures this.

**Assumption 3.2.** For each fixed QoS $Q \in \mathcal{Q}$, the cost function $c(Q, t)$ is strictly decreasing in $t$. In particular, we assume that $c(Q, t) = c(Q, 0) \cdot e^{-\eta t}$, where $0 < \tau < \eta$. In addition, the cost $c(Q, 0)$ is *strictly convex* and *strictly increasing* in $Q$.

When there is no confusion, we denote $c(Q, 0)$ simply by $c(Q)$. Here, we assume that $\eta > \tau$ because $\eta$ reflects both (i) the decrease in investment costs due to advances in technologies and (ii) the time value of money.

## 3.2 Stackelberg Game Models

Clearly, the price that the CP will be willing to pay will depend on the QoS delivered by the SP. The terms of the contract will likely be dependent on the bargaining powers of these two parties. Furthermore, they are private entities most likely interested only in their own payoffs.

In this section, we consider the scenarios in which one of the providers is in a much better position to shape the terms of the contract to be offered. We examine two separate cases. In the first (resp. second) case, we assume that the SP (resp. the CP) has more bargaining power. We model each case as a Stackelberg game with the party with stronger bargaining power as a leader and the other as a follower and examine their NEs [7].

First, we note that both the CP and the SP can ensure a *minimum payoff*, which is the payoff they will receive without any contract between them. These minimum payoffs are equal to

$$U_{CP}^{\min} := U_{CP,R}(Q_{\min}) \quad \text{and} \quad U_{SP}^{\min} := U_{SP,R}(Q_{\min}). \tag{3.9}$$

Therefore, a contract $\Gamma$ is *feasible* in the sense that the CP and the SP may agree to it, only if the aggregate payoff is greater than or equal to $U_{CP}^{\min} + U_{SP}^{\min} =: U_{ag}^{\min}$. The reason for this is that if the aggregate payoff is strictly less than $U_{ag}^{\min}$, either the CP or the SP will have to accept a payoff strictly smaller than its minimum payoff. Since a rational player will not agree to a payoff less than its minimum payoff, such a contract will never be signed by rational CP and SP. We denote the set of feasible contracts by $\mathbf{\Gamma}_{fc}$.

In a Stackelberg game, the leader first offers a contract $\Gamma = (Q, p)$ to the follower and then the follower decides whether to accept the contract or not. Therefore, the action space of the leader is the set of all contracts (of the form under consideration)[4] and that of the follower is {accept, reject}. If the follower accepts the contract $\Gamma$, the payoffs of the CP and the SP are given by the payoff vector $(U_{CP}, U_{SP}) = (U_{CP}(\Gamma), U_{SP}(\Gamma))$. Otherwise, their payoff vector is $(U_{CP}, U_{SP}) = (U_{CP}^{\min}, U_{SP}^{\min})$. Throughout the rest of this section and following sections, we assume that the set of feasible contracts is nonempty, i.e., $\mathbf{\Gamma}_{fc} \neq \emptyset$.

Second, at an NE of the Stackelberg game, the leader will be able to extract all the increase in the aggregate payoff and the follower will receive its minimum payoff. The reason is that, assuming that a contract $\Gamma = (Q, p)$ is feasible, the leader can choose a price $p$ such that the payoff to the follower is equal to its minimum payoff. Recall that at an NE, no player can increase its payoff via unilateral deviation, Thus, if the follower receives a payoff larger than its minimum payoff at an NE, the leader

---

[4] In practice, however, the action space of the leader can be assumed to be the set of *feasible* contracts for the reason explained earlier.

can change the price so that it can increase its own payoff, thereby contradicting that it is an NE.

### 3.2.1  Stackelberg Game with the SP as a Leader

In the Stackelberg game in which the SP is the leader, the following procedure can be used to further reduce the set of contracts that will be considered by the SP at an NE. Given the QoS $Q$, compute the maximum price $p_{\max}^{CP}(Q)$ that the CP would be willing to pay by setting $U_{CP}(Q, p_{\max}^{CP}(Q)) = U_{CP}^{\min}$. Clearly, there exists a unique maximum price.

After a little algebra, we obtain

$$p_{\max}^{CP}(Q) = \frac{(A_1 + A_2) f_{CP}(\zeta(Q) - \zeta(Q_{\min}))}{N_{CP}(0)\tau A_1 + \zeta(Q)(A_1 + A_2)}, \tag{3.10}$$

where $A_1 := 1 - e^{-\tau M}$ and $A_2 := -\tau\, M\, e^{-\tau M}$. Note that $A_1 + A_2 = 1 - e^{-\tau M}(1 + \tau M) > 0$ because $\log(1) = 0 > -\tau M + \log(1 + \tau M)$.

The goal of the SP is to maximize its payoff by searching over all feasible contracts of the form $\Gamma = (Q, p_{\max}^{CP}(Q))$, i.e.,

$$\max_{Q \in \mathcal{Q}:(Q,\; p_{\max}^{CP}(Q)) \in \boldsymbol{\Gamma}_{fc}} U_{SP}((Q,\; p_{\max}^{CP}(Q))). \tag{3.11}$$

Suppose that $Q_{SP}^{\star}$ is a solution to (3.11). Then, an NE of the Stackelberg game is given by $((Q_{SP}^{\star}, p_{\max}^{CP}(Q_{SP}^{\star})), \text{accept}).$[5]

---

[5] There may exist other NEs of the Stackelberg game where no contract is signed by the providers.

### 3.2.2 Stackelberg Game with the CP as a Leader

In the second Stackelberg game model, the roles of the providers are reversed; the CP is the leader and the SP is the follower. Similarly as in the first Stackelberg game outlined in the previous subsection, when the CP is the leader, we can use the following procedure to reduce the set of contracts that will be considered by the CP at an NE: for any fixed $Q \in \mathcal{Q}$, calculate the minimum price $p_{\min}^{SP}(Q)$ that the SP would be willing to accept for providing the QoS $Q$ by setting $U_{SP}(Q, p_{\min}^{SP}(Q)) = U_{SP}^{\min}$.

From (3.7) and (3.9), we get

$$p_{\min}^{SP}(Q) = \frac{Num_1(Q) + Num_2(Q)}{\tau \ N_{CP}(0) \ A_1 + \zeta(Q)(A_1 + A_2)}, \tag{3.12}$$

where $A_3 := 1 - e^{-\eta M}$,

$$Num_1(Q) = \tau^2 (N_{CP}(0)(c(Q) - c(Q_{\min})) + (\zeta(Q) - \zeta(Q_{\min}))A_3/\eta)$$

$$Num_2(Q) = -f_{SP}(A_1 + A_2)(\lambda(Q) - \lambda(Q_{\min}))$$

The goal of the CP is to maximize its payoff by searching over all feasible contracts of the form $\Gamma = (Q, p_{\min}^{SP}(Q))$, i.e.,

$$\max_{Q \in \mathcal{Q}:(Q, \ p_{\min}^{SP}(Q)) \in \mathbf{\Gamma}_{fc}} U_{CP}((Q, \ p_{\min}^{SP}(Q))). \tag{3.13}$$

Let $Q_{CP}^{\dagger}$ be a solution to (3.13). An NE of the Stackelberg game is then $((Q_{CP}^{\dagger}, p_{\min}^{SP}(Q_{CP}^{\dagger})),$ accept).

21

### 3.3  Bargaining Problem and Social Efficiency

Another possible scenario that may emerge in practice is that neither the CP nor the SP has much stronger bargaining power than the other. In this case, the interaction between the two providers may be modeled as a two-person bargaining problem [16, 26]. A two-person bargaining problem tells us how two entities (with comparable bargaining power) should cooperate in order to mitigate the well-known inefficiency of Nash equilibria of non-cooperative games.

In this section, we model the problem of contract design/selection as a two-person bargaining problem between the two providers and examine the NBS. In the process, we establish the relation between the NBS and the Pareto optimum that maximizes the aggregate payoff of both providers.

#### 3.3.1  Nash Bargaining Solution

Let us denote the set of *feasible payoff vectors* by $\mathcal{P}$, i.e.,

$$\mathcal{P} = \{(U_{CP}(\Gamma), U_{SP}(\Gamma)) \mid \Gamma \in \mathbf{\Gamma}_{fc}\}.$$

The NBS is a payoff vector $(U_{CP}^{\star}, U_{SP}^{\star})$ for the two providers and is defined via an axiomatic approach [26].

1. (Individual rationality): $(U_{CP}^{\star}, U_{SP}^{\star}) \geq (U_{CP}^{\min}, U_{SP}^{\min})$;

2. (Feasibility): $(U_{CP}^{\star}, U_{SP}^{\star}) \in \mathcal{P}$;

3. (Pareto optimality): if there exists another feasible payoff vector $(U_{CP}', U_{SP}') \geq (U_{CP}^{\star}, U_{SP}^{\star})$, then $(U_{CP}', U_{SP}') = (U_{CP}^{\star}, U_{SP}^{\star})$;

4. (Independence of irrelevant alternatives): Suppose that another bargaining problem has (i) a feasible payoff vector set $\mathcal{P}^* \subset \mathcal{P}$ such that $(U_{CP}^\star, U_{SP}^\star) \in \mathcal{P}^*$ and (ii) the same minimum payoffs. Then, the NBS of the new bargaining problem is $(U_{CP}^\star, U_{SP}^\star)$;

5. (Independence of linear transformations): Consider another bargaining problem in which the payoffs of the two parties are given by linear transformations of those in the original bargaining problem. Then, the NBS of the new bargaining problem is given by the linear transformations of the NBS of the original bargaining problem;

6. (Symmetry): Suppose that $\mathcal{P}$ is symmetric, i.e., if $\mathbf{v} = (v_1, v_2) \in \mathcal{P}$, then $(v_2, v_1) \in \mathcal{P}$, and $U_{CP}^{\min} = U_{SP}^{\min}$. Then, $U_{CP}^\star = U_{SP}^\star$.

Moreover, it is well known that, if there exists another feasible payoff vector $(U_{CP}', U_{SP}') > (U_{CP}^{\min}, U_{SP}^{\min})$, then there exists a unique payoff vector that maximizes the product $(U_{CP} - U_{CP}^{\min}) \times (U_{SP} - U_{SP}^{\min})$ among all feasible payoff vectors, and the unique maximizer is the NBS.

Our problem, however, has a special structure that simplifies the analysis of the NBS. Recall from Section 3.1 that, for fixed QoS $Q$, any feasible contract $\Gamma = (Q, p)$ achieves the same aggregate payoff because the aggregate payoff depends only on the QoS $Q$ of the contract from (3.8). Therefore, the set of feasible payoff vectors achieved by contracts with fixed $Q \in \mathcal{Q}$ (assuming $U_{ag}(Q) > U_{ag}^{\min}$) is given by the line segment between $(U_{ag}(Q) - U_{SP}^{\min}, \ U_{SP}^{\min})$ and $(U_{CP}^{\min}, \ U_{ag}(Q) - U_{CP}^{\min})$. This is illustrated in Fig. 3.1.
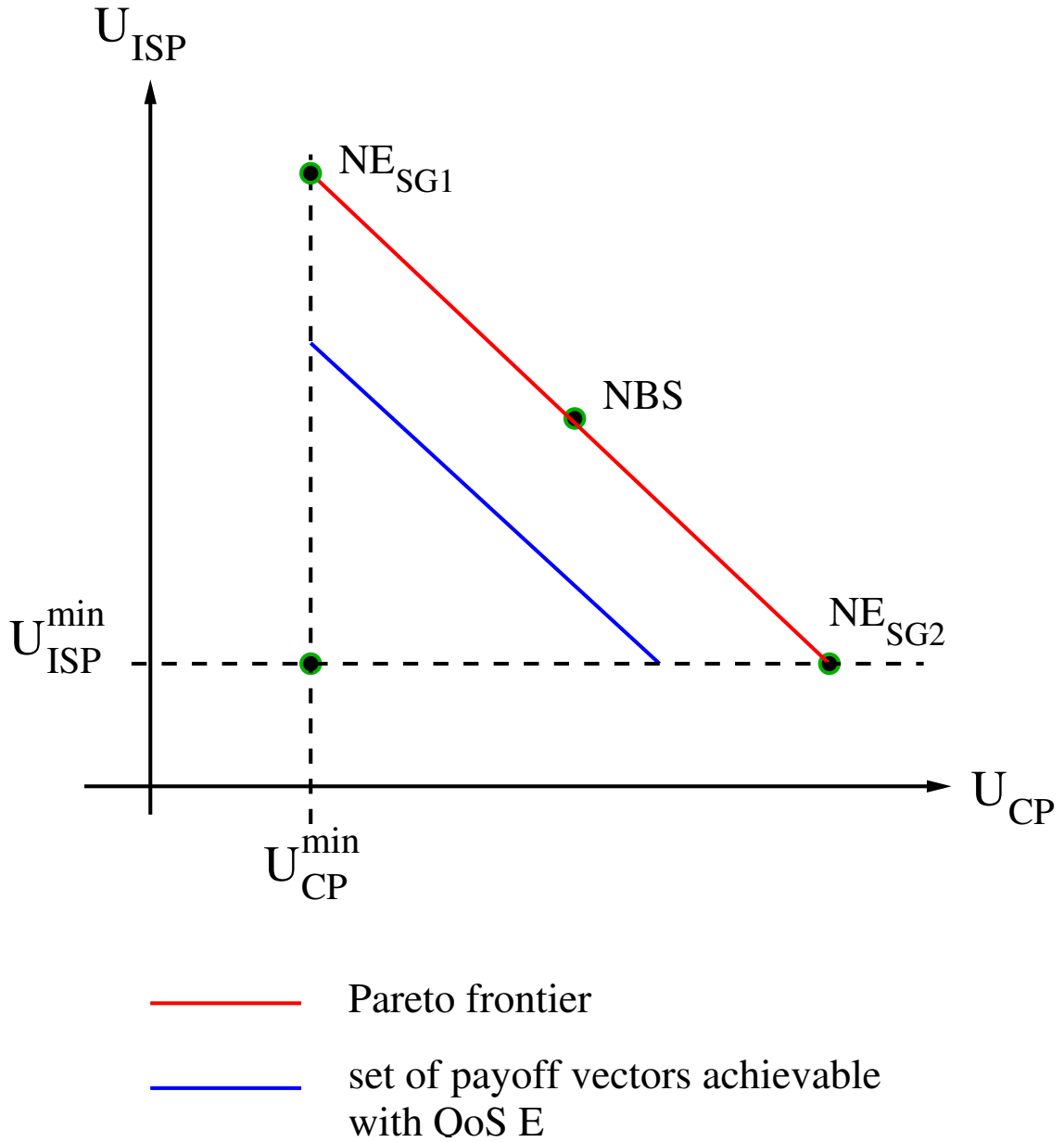
*Fig. 3.1:* Example of a set of achievable payoff vectors with a fixed QoS $Q \in \mathcal{Q}$ and Pareto frontier. $NE_{SG1}$ (resp. $NE_{SG2}$) = Nash equilibrium of the Stackelberg game with the SP (resp. the CP) as the leader, and NBS = Nash bargaining solution.

For the reason explained above, the Pareto frontier is given by the set of feasible payoff vectors that can be achieved with $Q^\star \in \arg\max_{Q \in \mathcal{Q}} U_{ag}(Q)$. Let us denote $\max_{Q \in \mathcal{Q}} U_{ag}(Q)$ by $U_{ag}^{\max}$ and the difference $\Delta U_{ag} := U_{ag}^{\max} - U_{ag}^{\min}$. Then, the NBS is given by

$$NBS = \left(0.5\Delta U_{ag} + U_{CP}^{\min}, \ 0.5\Delta U_{ag} + U_{SP}^{\min}\right)$$
$$= \left(\frac{U_{ag}^{\max} + U_{CP}^{\min} - U_{SP}^{\min}}{2}, \frac{U_{ag}^{\max} + U_{SP}^{\min} - U_{CP}^{\min}}{2}\right).$$

This is shown in Fig. 3.1.

### 3.3.2    Social Optimum

It is clear from the above discussion that the NBS is socially optimal in that it maximizes the aggregate payoff of the two providers. A related interesting question is how efficient the NEs of the Stackelberg games (discussed in the previous section) are. In order to answer this question, we need to first compute $\arg\max_{Q \in \mathcal{Q}} U_{ag}(Q)$.

From (3.1) through (3.3), we obtain

$$U_{ag}(Q) = f_{CP} \left[\frac{N_{CP}(0)\ A_1}{\tau} + \frac{\zeta(Q)(A_1 + A_2)}{\tau^2}\right]$$
$$+ f_{SP} \left[\frac{N_{SP}(0)\ A_1}{\tau} + \frac{\lambda(Q)(A_1 + A_2)}{\tau^2}\right]$$
$$- c(Q)\ N_{CP}(0) - \frac{c(Q)\zeta(Q)A_3}{\eta}.$$

After removing the terms that do not depend on the QoS $Q$, the aggregate payoff maximization problem is equivalent to

$$\max_{Q \in \mathcal{Q}} \quad \left(\frac{1}{\tau^2}((f_{CP} \cdot \zeta(Q) + f_{SP} \cdot \lambda(Q))(A_1 + A_2))\right.$$
$$\left. -c(Q)N_{CP}(0) - \frac{c(Q)\zeta(Q)A_3}{\eta}\right). \tag{3.14}$$

25

From Assumptions 3.1 and 3.2, one can show that there exists a unique optimal solution to the above maximization problem, i.e., $|\arg\max_{Q\in\mathcal{Q}} U_{ag}(Q)| = 1$.

## 3.4   Main Results

One interesting question that arises naturally in the Stackelberg game models discussed in Section 3.2 is which Stackelberg game yields a more efficient NE *when a contract is signed.* Let us denote the NE of the Stackelberg game with the SP (resp. the CP) as the leader by $NE_{SG1}$ (resp. $NE_{SG2}$). The proofs of our main results are provided in the following subsection.

**Theorem 1.** *The QoS chosen at the NEs $NE_{SG1}$ and $NE_{SG2}$ is the same, i.e., $Q^\star_{SP} = Q^\dagger_{CP}$. Hence, the aggregate payoff achieved by both NEs $NE_{SG1}$ and $NE_{SG2}$ is identical.*

A next question of interest to us is where these NEs lie in relation to the Pareto frontier, in particular the NBS. The following theorem answers this question.

**Theorem 2.** *The NBS achieves the maximum aggregate payoff. Furthermore, the aggregate payoff of the NBS is equal to that of $NE_{SG1}$ and $NE_{SG2}$.*

An important implication of this finding is that when one of the providers has a much stronger bargaining position than the other provider and can dictate the terms of the contract, then regardless of which provider is the leader, there is no loss of efficiency due to the selfish nature of the providers in the sense that the aggregate payoff of the providers is maximized. Another way of stating this is that the *price of stability* is one [9].

Before we present the proofs of these two theorems in the subsequent subsection, we first offer some intuition behind them. Recall from Section 3.2 that, in a Stackelberg game, the leader offers a contract that yields the minimum payoff to the follower at the NE $NE_{SG1}$ or $NE_{SG2}$. Hence, in both maximization problems in (3.11) and (3.13), the leader essentially tries to maximize the *aggregate* payoff, knowing that it will be able to keep all of the aggregate payoff except for the minimum payoff of the follower. For this reason, it is in the interest of the leader to maximize the aggregate payoff, thereby achieving a Pareto optimal payoff vector at the NEs.

Another question we are interested in is how the contract duration $M$ affects the QoS at the Pareto optimum that maximizes the aggregate payoff of the two providers. It is not immediately clear from (3.11) and (3.13) how the equilibrium QoS $Q_{SP}^{\star}$ or $Q_{CP}^{\dagger}$ would behave as the contract duration $M$ is varied.

It turns out that there is no general trend that holds in all cases. However, under certain mild conditions, in particular when the contract duration is large in relation to $\tau^{-1}$, increasing contract duration tends to encourage the providers to improve the equilibrium QoS. This is stated formally in the following theorem: Let $Q^{\star}(M)$ denote the QoS at the Pareto frontier with varying contract duration $M$. Recall that the QoS at the NEs of Stackelberg games is identical to the QoS at the Pareto optimum that maximizes the aggregate payoff.

**Theorem 3.** *Suppose that $c$, $\zeta$ and $\lambda$ are twice differentiable. Fix a contract duration $M^*$ and assume that the following hold:*

*c1.* $Q^\star(M^*) \in (Q_{\min}, Q_{\max})$;

*c2.* $\tau \cdot M^* \geq 1$; *and*

*c3.* $c(Q)\frac{d}{dQ}\zeta(Q) + \zeta(Q)\frac{d}{dQ}c(Q)$ *is strictly increasing in* $Q$ *over* $[Q^\star(M^*), Q_{\max}]$.

*Then,* $Q^\star(M)$ *is nondecreasing in* $M$ *over* $[M^*, \infty)$. *In fact, it is strictly increasing when* $Q^\star(M) < Q_{\max}$.

Note that the condition c3 in the theorem is satisfied if $\zeta$ is an affine function by Assumption 3.2.

Theorem 3 tells us that if the contract duration is long relative to the discount rate modeling the time value of money, i.e., $\tau$, then the Pareto optimal QoS, which is the same as the QoS at the NEs of Stackelberg games, increases with the contract duration. Hence, the content service subscribers will likely enjoy better service when the providers sign a long-term contract.

### 3.4.1   Proofs of Main Results

*Proof of Theorem 1:*   Note that in both (3.11) and (3.13), the leader searches over the set of feasible contracts in order to find a feasible contract that (i) yields the minimum payoff to the follower and (ii) maximizes the leader's payoff. Therefore, the maximization problem in (3.11) is equivalent to the following:

$$\max_{Q \in \mathcal{Q}} \left( U_{SP}((Q,\ p_{\max}^{CP}(Q))) + U_{CP}^{\min} \right) \tag{3.15}$$

This is because adding a constant to the objective function does not change the solution. Clearly, the new objective function in (3.15) is the aggregate payoff achieved
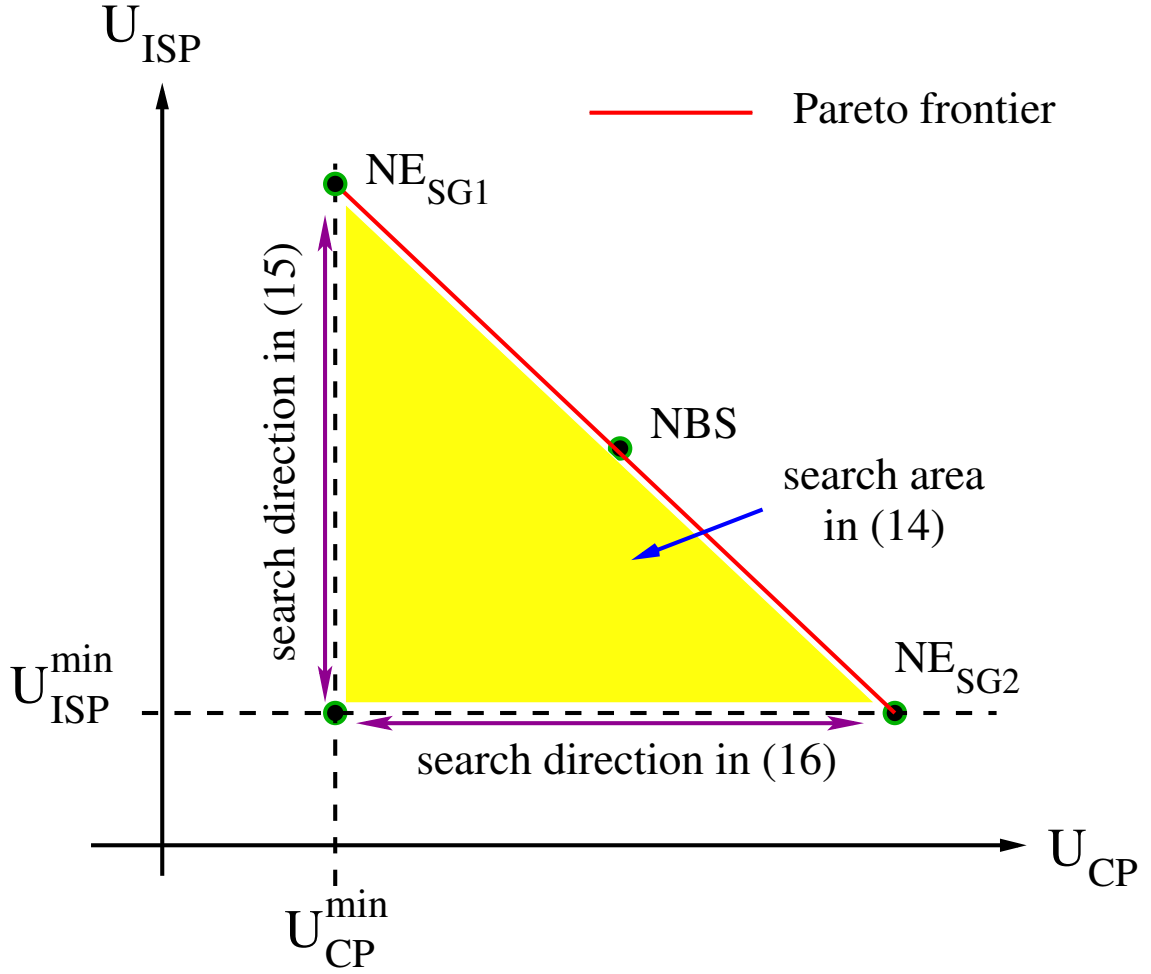
28

*Fig. 3.2:* Search directions and area in maximization problems (3.14), (3.15) and (3.16).

by the contract. Therefore, in order to solve (3.15), the SP simply searches along the dotted vertical line connecting the payoff vector $(U_{CP}^{\min}, U_{SP}^{\min})$ and $NE_{SG1}$ as indicated by the vertical purple arrow in Fig. 3.2.

Similarly, (3.13) is equivalent to

$$\max_{Q \in \mathcal{Q}} \left( U_{CP}((Q,\ p_{\min}^{SP}(Q))) + U_{SP}^{\min} \right). \tag{3.16}$$

By an analogous argument, in order to solve (3.16), the CP looks for a feasible contract that maximizes the aggregate payoff along the horizontal line connecting

the payoff vector $(U_{CP}^{\min}, U_{SP}^{\min})$ and $NE_{SG2}$ as shown by the horizontal purple arrow in Fig. 3.2.

This observation can be made more formal as follows. First, by substituting (3.10) in (3.7) and removing the terms that do not depend on $Q$, we can show that (3.11) is equivalent to the following:

$$\max_{(Q,\ p_{\max}^{CP}(Q)) \in \mathbf{\Gamma}_{fc}} \left( \frac{(f_{CP}\ \zeta(Q) + f_{SP}\ \lambda(Q))(A_1 + A_2)}{\tau^2} \right.$$
$$\left. -c(Q)N_{CP}(0) - \frac{c(Q)\zeta(Q)A_3}{\eta} \right). \tag{3.17}$$

Similarly, plugging (3.12) in (3.6) and deleting the terms that are independent of $Q$, we find the following equivalent maximization problem for (3.13):

$$\max_{(Q,\ p_{\min}^{SP}(Q)) \in \mathbf{\Gamma}_{fc}} \left( \frac{(f_{CP}\ \zeta(Q) + f_{SP}\ \lambda(Q))(A_1 + A_2)}{\tau^2} \right.$$
$$\left. -c(Q)N_{CP}(0) - \frac{c(Q)\zeta(Q)A_3}{\eta} \right). \tag{3.18}$$

First, note that the objective functions in (3.17) and (3.18) are identical. Since both of these optimizations are over the QoS $Q$ (with the price being given by either (3.10) or (3.12)), they yield the same optimal value and optimal solution $Q^\star$. Because the aggregate payoff depends only on the QoS $Q$, this completes the proof of Theorem 1.

*Proof of Theorem 2:* The proof of the theorem follows directly from the proof of Theorem 1 and (3.14). Note that the aggregate payoff maximization problem in (3.14) has the same objective function as (3.17) (or (3.18)). Therefore, the only difference between (3.14) and (3.15) (or (3.16)) is that, rather than limiting the search to the line segment between the payoff vector $(U_{CP}^{\min}, U_{SP}^{\min})$ and $NE_{SG1}$ (or

30

$NE_{SG2}$), (3.14) searches over the set of feasible contracts, i.e., the yellow area in Fig. 3.2, and the maximum aggregate payoff is achieved on the Pareto frontier shown as the red line in Fig. 3.2. Since the NBS, $NE_{SG1}$ and $NE_{SG2}$ all lie on the Pareto frontier, they achieve the same aggregate payoff.

*Proof of Theorem 3:* For notational simplicity, we denote $Q^\star(M^*)$ by $Q^*$. In order to make the dependence on contract duration $M$ explicit, we denote the aggregate payoff $U_{ag}(Q)$ by $U_{ag}(Q; M)$. From (3.14) and condition c1 in the theorem, the first-order necessary condition [2] tells us

$$
\begin{aligned}
\frac{\partial}{\partial Q} & U_{ag}(Q; M^*)\Big|_{Q=Q^*} \\
&= \frac{1}{\tau^2}(f_{CP} \cdot \zeta'(Q^*) + f_{SP} \cdot \lambda'(Q^*))(A_1(M^*) + A_2(M^*)) \\
&\quad - c'(Q^*)N_{CP}(0) \\
&\quad - \frac{1}{\eta}A_3(M^*)(c'(Q^*)\zeta(Q^*) + c(Q^*)\zeta'(Q^*)) \\
&= 0,
\end{aligned}
\tag{3.19}
$$

where $A_1(M) = 1 - e^{-\tau M}$, $A_2(M) = -\tau Me^{-\tau M}$ and $A_3(M) = 1 - e^{-\eta M}$.

Define $G : \mathbb{R}_+ \times (Q_{\min}, Q_{\max}) \to \mathbb{R}$, where

$$
G(M, Q) = \frac{\partial}{\partial Q}U_{ag}(Q; M).
$$

By its definition and from (3.19), we have $G(M^*, Q^*) = 0$. Thus, by the implicit function theorem [27], there exist (i) open sets $\mathcal{Q}^+ \subset \mathcal{Q}$ and $\mathcal{M}^+ \subset \mathbb{R}_+$ such that $M^* \in \mathcal{M}^+$ and $Q^* \in \mathcal{Q}^+$ and (ii) a continuously differentiable function $\phi : \mathcal{M}^+ \to \mathcal{Q}^+$ such that $G(M, \phi(M)) = 0$ for all $M \in \mathcal{M}^+$. Therefore, in order to

prove the theorem, it suffices to prove that $d\phi(M)/dM > 0$ for all $M \in [M^*, \infty)$ such that $Q^\star(M) < Q_{\max}$.

The implicit function theorem also tells us

$$\frac{d}{dM}\phi(M) = -\left(\frac{\partial}{\partial Q}G(M, \phi(M))\right)^{-1}\frac{\partial}{\partial M}G(M, \phi(M)). \tag{3.20}$$

Let us first consider $\frac{\partial}{\partial Q}G(M, \phi(M))$.

$$\begin{aligned}
\frac{\partial}{\partial Q}&G(M, \phi(M)) \\
&= \frac{1}{\tau^2}(f_{CP} \cdot \zeta''(\phi(M)) + f_{SP} \cdot \lambda''(\phi(M)))(A_1 + A_2) \\
&\quad -c''(\phi(M))N_{CP}(0) \tag{3.21} \\
&\quad -\frac{1}{\eta}A_3(c''(\phi(M))\zeta(\phi(M)) + 2c'(\phi(M))\zeta'(\phi(M)) \\
&\qquad +c(\phi(M))\zeta''(\phi(M)))
\end{aligned}$$

The first term in (3.21) is nonpositive since both $\zeta$ and $\lambda$ are assumed concave by Assumption 3.1. The second term is nonpositive because the cost function $c$ is assumed convex. Finally, by condition c3 in the theorem, the last term is strictly negative. Therefore, $\partial G(M, \phi(M))/\partial Q < 0$.

In light of this and (3.20), in order to prove $d\phi(M)/dM > 0$, we only need to prove $\partial G(M, \phi(M))/\partial M > 0$.

$$\frac{\partial}{\partial M}G(M, \phi(M)) = \Xi_1(M)Me^{-\tau M} - \Xi_2(M)e^{-\eta M}, \tag{3.22}$$

where $\Xi_1(M) := f_{CP} \cdot \zeta'(\phi(M)) + f_{SP} \cdot \lambda'(\phi(M))$ and $\Xi_2(M) := c'(\phi(M))\zeta(\phi(M)) + c(\phi(M))\zeta'(\phi(M))$.

Recall that $G(M, \phi(M)) = 0$ for all $M \in \mathcal{M}^+$. Thus, from (3.19)

$$\frac{1}{\tau^2}\Xi_1(M)(A_1(M) + A_2(M)) \geq \frac{1}{\eta}A_3(M)\Xi_2(M).$$

32

This implies that, for all $M \in \mathcal{M}^+$,

$$\frac{\Xi_1(M)}{\Xi_2(M)} \geq \frac{\tau^2 A_3(M)}{\eta(A_1(M) + A_2(M))} > \frac{\tau^2}{\eta}, \tag{3.23}$$

where the second inequality follows from the inequality

$$\frac{A_3(M)}{A_1(M) + A_2(M)} = \frac{1 - e^{-\eta M}}{1 - e^{-\tau M}(1 + \tau M)} > 1$$

for all $M > 0$ because $\tau < \eta$.

We now proceed to complete the proof of the theorem. From (3.22), proving $\partial G(M, \phi(M))/\partial M > 0$ is equivalent to showing

$$\frac{\Xi_1(M)}{\Xi_2(M)} > \frac{\exp(-\eta M)}{M \exp(-\tau M)}.$$

To prove this inequality, using (3.23), we will demonstrate

$$\frac{\Xi_1(M^*)}{\Xi_2(M^*)} > \frac{\tau^2}{\eta} \geq \frac{\exp(-\eta M)}{M \exp(-\tau M)}$$

or, equivalently,

$$\tau^2 M \exp(-\tau M) \geq \eta \exp(-\eta M). \tag{3.24}$$

Rewriting $\eta$ as $\tau(1 + \epsilon)$ with $\epsilon > 0$ and dividing both sides by $\tau \exp(-\eta M)$, (3.24) is equivalent to

$$\tau M \exp(\epsilon \tau M) \geq (1 + \epsilon)$$

For all $M \geq M^*$, from the assumption $\tau M^* \geq 1$,

$$\tau M \exp(\epsilon \tau M) \geq \exp(\epsilon) > (1 + \epsilon).$$

This completes the proof of Theorem 3.

# Chapter 4: Interaction Between a Single SP and Two CPs

In this chapter we extend our study to the case where there are one SP and two competing CPs in the market. Subscribers to the SP (resp. CP1 and CP2) pay a subscription fee $f_{SP}$ (resp. $f_{CP}^1$ and $f_{CP}^2$). As in the previous chapter, we assume that the subscribers to the CPs must have internet service and subscribe to the SP. Moreover, the SP provides at least minimum QoS, $Q_{\min}$, to all traffic.

We assume that the SP offers only a single fast lane with higher QoS. In other words, when the SP signs a contract with both CPs, it offers the same QoS to both CPs. Hence, the class of contracts considered by the SP is given by a triple $(Q, p_{CP}^1, p_{CP}^2)$, where $Q \in (Q_{\min}, Q_{\max}]$ is the common QoS to be provided to the streaming traffic of the CPs that sign a contract with the SP, and $p_{CP}^1$ (resp. $p_{CP}^2$) is the price that CP1 (resp. CP2) will pay per subscriber to the SP. From the earlier assumption, the streaming traffic of a CP that does not sign a contract with the SP continues to receive the minimum QoS $Q_{\min}$.

In order to capture the payoffs of the SP, CP1 and CP2, we introduce the following notation. Suppose $\mathbf{Q} = (Q_1, Q_2)$ and $\mathbf{T} = (T_1, T_2)$. Let $N_{CP}^i(t; \mathbf{Q}, \mathbf{T})$, $i = 1, 2$, and $N_{SP}(t; \mathbf{Q}, \mathbf{T})$ be the number of subscribers for CP$i$ and SP, respectively, at time $t \in \mathbb{R}_+$ when the streaming service QoS is $Q_i$ for CP$i$, $i = 1, 2$, starting

at time $t = T_i$.[1] The duration of the contract, if any is signed, is denoted by $M \in (0, \infty)$ as before. If CP$i$ signs a contract at time $t = 0$, then the other CP can sign a contract only at time $t = m$, where $0 \leq m \leq M$. Hence, $m = 0$ means that both CPs sign the contract at the same time, and $t = M$ implies that the second CP never signs a contract.

**i. Revenues:** Analogous to the previous case, the payoff of the CP$i$, $i = 1, 2$, and the SP from their revenue over the interval $[0, M]$ when the SP provides the streaming service QoS $Q_i$ to CP$i$'s traffic is given by

$$U_{CP,R}^i(\mathbf{Q}, \mathbf{T}) = \int_0^M N_{CP}^i(t; \mathbf{Q}, \mathbf{T}) \; f_{CP}^i \; e^{-\tau t} \; dt \quad \text{and} \tag{4.1}$$

$$U_{SP,R}(\mathbf{Q}, \mathbf{T}) = \int_0^M N_{SP}(t; \mathbf{Q}, \mathbf{T}) \; f_{SP} \; e^{-\tau t} \; dt, \tag{4.2}$$

where $\tau > 0$ is the discount rate that captures the time value of money.

**ii. SP investment costs:** The investment necessary for expanding network capacity and improving network equipments to offer QoS $Q \in \mathcal{Q}$ to a CP subscriber depends on the QoS $Q$. We model this investment using the cost function $c$ : $\mathcal{Q} \times [0, M] \rightarrow \mathbb{R}_+$, where $c(Q, t)$ tells us the (additional) investment the ISP needs to make at time $t$ in order to serve a new CP subscriber joining at time $t$ with the QoS $Q$. We assume that this investment does not depend which CP the subscriber belongs to.

Recall that the ISP is assumed to make necessary investments to maintain at least the minimum QoS $Q_{\min}$ for all traffic, which we do not explicitly model. In other words, $c(Q_{\min}, t) = 0$ for all $t \in [0, M]$.

---

[1] An implicit assumption is that, prior to $T_i$, the traffic from CP$i$ receives QoS $Q_{\min}$.

The total cost of the SP for this investment over the contract duration is then given by

$$C_{SP}(\mathbf{Q}, \mathbf{T}) = \sum_{i=1}^{2} \Big( c(Q_i, T_i) N_{CP}^i(T_i; \mathbf{Q}, \mathbf{T})$$

$$+ \int_{T_i}^{M} c(Q_i, t) \, \frac{\partial}{\partial t} N_{CP}^i(t; \mathbf{Q}, \mathbf{T}) \, dt \Big), \qquad (4.3)$$

where $\partial N_{CP}^i(t; \mathbf{Q}, \mathbf{T})/\partial t$ is the rate at which the number of CP$i$ subscribers increases at time $t$ when QoS $Q_i$, $i = 1, 2$, is provided to subscribers of CP$i$, starting at time $T_i$. Since the number of CP$i$'s subscribers, i.e., $N_{CP}^i(0; \mathbf{Q}, \mathbf{T})$, does not depend on $Q_i$, $i = 1, 2$, we denote it by $N_{CP}^i(0)$ for notational ease.

As in the previous case with a single CP, the investment in (4.3) consists of two parts – the up-front investment needed for existing subscribers at time $t = T_i$, i.e., when the contract takes effect, and the investment for newly arriving subscribers over the contract duration.

**iii. Payment by the CP$i$ to the SP:** When a CP and the SP sign a contract $\Gamma = (Q, p)$ with $Q \in (Q_{\min}, Q_{\max}]$, starting at time $t = T$, the CP must pay the SP according to the contract $\Gamma$. Suppose that CP1 signs the contract at $t = 0$ and CP2 signs the contract at $t = m$.[2] We denote the contract signed between the SP and CP$i$ by $\Gamma_i = (Q_i, p_i)$, $i = 1, 2$, the time at which contract $\Gamma_i$ takes effect by $T_i$, and $(\Gamma_1, \Gamma_2)$ by $\Gamma_{1,2}$.

When CP$i$ does not sign a contract with the SP, we assume either $(Q_i, p_i) = (Q_{\min}, 0)$ or $T_i = M$. Thus, we can interpret $(Q_{\min}, 0)$ as an *implicit* contract between the SP and a CP in the absence of an *explicit* contract with $Q > Q_{\min}$. Moreover, in

---

[2] We will study the effect of who signs first shortly.

the rest of thip chapter, we impose the assumption that the SP offers common QoS $Q$ to both CPs if contracts are signed. Hence, when both the CPs sign a contract with the SP, then $Q_1 = Q_2 \in (Q_{\min}, Q_{\max}]$.

The payoff of the CP$i$ and the SP from this payment from the CP to the SP is given by

$$U^i_{CP,P}(\Gamma_{1,2}, \mathbf{T}) = -\int_{T_i}^M p_i \, N^i_{CP}(t; \mathbf{Q}, \mathbf{T}) \, e^{-\tau t} \, dt \quad \text{and} \tag{4.4}$$

$$U_{SP,P}(\Gamma_{1,2}, \mathbf{T}) = \sum_{i=1}^2 \left( \int_{T_i}^M p_i \, N^i_{CP}(t; \mathbf{Q}, \mathbf{T}) \, e^{-\tau t} \, dt \right) \tag{4.5}$$

$$= -U^1_{CP,P}(\Gamma_{1,2}, \mathbf{T}) - U^2_{CP,P}(\Gamma_{1,2}, \mathbf{T}),$$

where $\tau$ is the same discount rate introduced in (4.1) and (4.2) to model the time value of money.

From (4.1) through (4.5), the payoffs of the providers are defined to be

$$U^i_{CP}(\Gamma_{1,2}, \mathbf{T}) = U^i_{CP,R}(\mathbf{Q}, \mathbf{T}) + U^i_{CP,P}(\Gamma_{1,2}, \mathbf{T}), \quad i = 1, 2, \tag{4.6}$$

$$U_{SP}(\Gamma_{1,2}, \mathbf{T}) = U_{SP,R}(\mathbf{Q}, \mathbf{T}) + U_{SP,P}(\Gamma_{1,2}, \mathbf{T}) - C_{SP}(\mathbf{Q}, \mathbf{T}). \tag{4.7}$$

The *aggregate payoff* of the providers, denoted by $U_{ag}$, is given by

$$U_{ag}(\Gamma_{1,2}, \mathbf{T}) = U_{SP}(\Gamma_{1,2}, \mathbf{T}) + \sum_{i=1}^2 U^i_{CP}(\Gamma_{1,2}, \mathbf{T})$$

$$= \sum_{i=1}^2 U^i_{CP,R}(\mathbf{Q}, \mathbf{T}) + U_{SP,R}(\mathbf{Q}, \mathbf{T}) - C_{SP}(\mathbf{Q}, \mathbf{T}). \tag{4.8}$$

Obviously, the aggregate payoff depends only on $(Q_i, T_i)$, $i = 1, 2$, but not on the prices $p_i$, $i = 1, 2$, because $\sum_{i=1}^2 U^i_{CP,P}(\Gamma_{1,2}, \mathbf{T})$ and $U_{SP,P}(\Gamma_{1,2}, \mathbf{T})$ cancel each other out. Therefore, we write the aggregate payoff as $U_{ag}(\mathbf{Q}, \mathbf{T})$ when there is no confusion.

To make progress, we introduce the following simplifying assumption.

**Assumption 4.1.** (a) The rate at which *new* content service users join CP$i$, $i = 1, 2$, is given by a mapping $\zeta_i : \mathcal{Q} \to \mathbb{R}$ for all $t \in [0, M]$. In other words, when CP$i$ is served with QoS $Q_i$, $i = 1, 2$, at time $t \in [0, M]$, new content service subscribers join CP$i$ at the rate $\zeta_i(Q_i)$. In addition, $\zeta_i$ is *strictly increasing* and *concave*. (b) The rate at which the number of SP subscribers increases is given by a *concave* function $\lambda : \mathcal{Q} \times \mathcal{Q} \to \mathbb{R}$. Moreover, $\lambda$ is *strictly increasing* in each argument, e.g., $\lambda(\mathbf{Q}) = \zeta_1(\mathbf{Q}) + \zeta_2(\mathbf{Q}) + \delta$ for some $\delta \geq 0$ and functions $\zeta_i$, $i = 1, 2$, satisfying Assumptions 4.1(a).

Note that we assume that the rate at which new content service subscribers joint a CP depends only on the CP's QoS, but is independent of that of the other CP. While we make this assumption to facilitate our analysis, it may not hold in practice.

In addition to the arrivals of new content service subscribers, we also model so-called *churn rates* when subscribers change their CP from one to the other based on, for instance, different QoS experienced by the CPs. Suppose that one of the CPs, say CP1, signs a contract with the SP and CP2 does not. Then, some of CP2's subscribers may switch to CP1. We model this switching of subscribers using a mapping $\theta : \mathcal{Q} \to \mathbb{R}_+$ as follows; the rate at which CP2's subscribers leave CP2 for CP1 at time $t$ is given by $\theta(Q_1)$ times the number of CP2 subscribers at time $t$.

**Assumption 4.2.** The mapping $\theta : \mathcal{Q} \to \mathbb{R}_+$ is *strictly increasing* and *concave* with $\theta(Q_{\min}) = 0$ and $\theta(Q_{\max}) = \theta_{\max}$ .

When both the CPs have a contract with the common QoS or neither has a contract with the SP, we assume that there is no churning.

As argued in Chapter 3, the cost of providing a certain level of QoS is expected to decline over time owing to advances in technologies which we capture via the following assumption.

**Assumption 4.3.** For each fixed QoS $Q \in \mathcal{Q}$, the cost function $c(Q, t)$ is strictly decreasing in $t$. In particular, we assume that $c(Q, t) = c(Q, 0) \cdot e^{-\eta t}$, where $0 < \tau < \eta$. In addition, the cost $c(Q, 0)$ is *strictly convex* and *strictly increasing* in $Q$.

For notational simplicity, we denote $c(Q, 0)$ by $c(Q)$ in the remainder of this chapter.

## 4.1 Stackelberg Game Model

The terms of the contract between the CPs and the SP will likely depend on the bargaining powers of these parties. Unlike in the previous chapter, however, we argue that the SP will be in a much stronger position to dictate the terms of the contract to be offered to the CPs using a well known power index, namely the *Shapley value*. The Shapley value, introduced by Shapley [26], is an axiomatic approach to determining the importance or bargaining power of each player in an $n$-player cooperative game.

### 4.1.1    The Shapley value

Let $\mathcal{N} := \{1, 2, \ldots, n\}$ denote the set of $n$ players. Suppose that a vector function $\phi(v) = (\phi_1(v), \ldots, \phi_n(v))$ defined on the set of characteristic functions of $n$-person games satisfies the following axioms [26]:

1. (Efficiency) Suppose that a coalition $S'$ satisfies that, for any coalition $S$, the equality $v(S) = v(S \cap S')$ holds. Then, $\sum_{i \in S'} \phi_i(v) = v(S')$;

2. (Symmetry) If $\pi$ is a permutation of the set $\mathcal{N}$ and, for any coalition $S$, the equality $v'(\pi S) = v(S)$ holds, then $\phi_{\pi i}(v') = \phi_i(v)$;

3. (Linearity) $\phi_i(v + u) = \phi_i(v) + \phi_i(u)$.

These axioms uniquely determine the mapping $\phi$, which is given by [26]

$$\phi_i(v) = \sum_{S \subset \mathcal{N}: i \in S} \frac{(|S|-1)!\,(n-|S|)!}{n!} [v(S) - v(S \setminus \{i\})], \ i \in \mathcal{N}. \qquad (4.9)$$

In our cooperative game $v$, there are three players – the SP, CP1 and CP2. Hence, possible coalitions are {CP1}, {CP2}, {SP}, {CP1, CP2}, {CP1, SP}, {CP2, SP} and {CP1, CP2, SP}. Note that the characteristic function $v : 2^{\mathcal{N}} \to \mathbb{R}$ assigns to each coalition the largest aggregate payoff the coalition can achieve. First, note that, as in Chapter 3, all three players can ensure a *minimum payoff*, which is the payoff they will receive without any contract between them. These minimum payoffs are equal to

$$U_{CP}^{1,\min} := U_{CP,R}^1(Q_{\min}, Q_{\max}, M, 0), \ U_{CP}^{2,\min} = U_{CP,R}^2(Q_{\max}, Q_{\min}, 0, M), \qquad (4.10)$$

$$\text{and } U_{SP}^{\min} := U_{SP,R}(Q_{\min}, Q_{\min}, M, M). \qquad (4.11)$$

A contract $\Gamma_{1,2}$ and effective times $T_i, i = 1, 2$, are *feasible* in the sense that the players – CP1, CP2 and SP – may agree to it, only if the aggregate payoff is greater than or equal to $U_{CP}^{1,\min} + U_{CP}^{2,\min} + U_{SP}^{\min} =: U_{ag}^{\min}$ for the same reason explained in Chapter 3. We denote the set of the pairs of feasible contracts and effective times by $\Xi_{feas}$.

We assume that there exist feasible contracts that increase the aggregate payoff. Put differently, for each a coalition $S$ with SP $\in S$, the SP can find $\xi \in \Xi_{feas}$ that yields the aggregate payoff strictly larger than the sum of the minimum payoffs of all members in the coalition. This implies that, for any coalition $S$ including the SP, $v(S) > \sum_{p \in S} v(\{p\})$. Since any contract requires the SP, a coalition without the SP cannot increase the aggregate payoff.

Using this observation, we can show that the SP has the largest Shapley value as follows. From (4.9), we obtain

$$\phi_{CP1}(v) = \frac{1}{6}\Big(2v(\{CP1\}) + 2(v(\{SP, CP1, CP2\}) - v(\{SP, CP2\}))$$
$$+ (v(\{CP1, CP2\}) - v(\{CP2\})) + (v(\{SP, CP1\}) - v(\{SP\}))\Big),$$
$$\phi_{CP2}(v) = \frac{1}{6}\Big(2v(\{CP2\}) + 2(v(\{SP, CP1, CP2\}) - v(\{SP, CP1\}))$$
$$+ (v(\{CP1, CP2\}) - v(\{CP1\})) + (v(\{SP, CP2\}) - v(\{SP\}))\Big),$$

and

$$\phi_{SP}(v) = \frac{1}{6}\Big(2v(\{SP\}) + 2(v(\{SP, CP1, CP2\}) - v(\{CP1, CP2\}))$$
$$+ (v(\{SP, CP1\}) - v(\{CP1\})) + (v(\{SP, CP2\}) - v(\{CP2\}))\Big).$$

Subtracting $\phi_{CP1}(v) - v(\{CP1\})$ from $\phi_{SP}(v) - v(\{SP\})$ yields

$$
(\phi_{SP}(v) - v(\{SP\})) - (\phi_{CP1}(v) - v(\{CP1\}))
$$

$$
= 0.5\Big((v(\{SP, CP2\}) - v(\{SP\})) - (v(\{CP1, CP2\}) - v(\{CP1\}))\Big)
$$

$$
\geq 0.5\Big(v(\{SP, CP2\}) - v(\{SP\}) - v(\{CP2\})\Big)
$$

$$
> 0,
$$

where the last strict inequality is the consequence of the assumption that, for any coalition $S$ that includes the SP, $v(S) > \sum_{p \in S} v(\{p\})$. Similar calculation gives us $(\phi_{SP}(v) - v(\{SP\})) - (\phi_{CP2}(v) - v(\{CP2\})) > 0$. This finding implies that the SP will possess stronger bargaining power than the CPs.

### 4.1.2   Stackelberg Game with the SP as a Leader

In this section, we consider a Stackelberg game where the SP is the leader and the two CPs are followers and examine its NEs. Recall from the previous chapter that, at an NE of the Stackelberg game, the SP will be able to enjoy all the increase in the aggregate payoff and both CPs will receive their minimum payoffs $U_{CP}^{i,\min}$, $i = 1, 2$. This is because the SP can choose prices $p_i$, $i = 1, 2$, such that the payoffs to the CPs are equal to their minimum payoffs in (4.10). We denote by $\Xi_{feas}^{con}$ the set of the pairs of feasible contracts $\Gamma_{1,2}$ and effective times $\mathbf{T}$, which satisfy (i) $Q_1 = Q_2$, (ii) $T_i = 0$ for some $i$, and (iii) $U_{CP}^i(\Gamma_{1,2}, \mathbf{T}) \geq U_{CP}^{i,min}$, $i = 1, 2$.

In the Stackelberg game, the SP first offers a contract $\Gamma = (Q, p_i)$ at time $t = 0$ to CP$i$, and CP$i$ decides whether to accept the contract or not (with no delay). Then, the SP offers a contract $\Gamma = (Q, p_j)$ at time $t = m$ to CP$j$, and CP$j$

decides if it will accept the contract (with no delay). Therefore, an action of the SP consists of (i) the offered QoS $Q$, (ii) prices charged to the CPs, (iii) time at which the second CP is approached, i.e., $m$, and (iv) the first CP to be offered a contract. Hence, $a_{SP} = (\Gamma_{1,2}, \mathbf{T})$, where $\Gamma_{1,2} = (\Gamma_1, \Gamma_2)$ and $\Gamma_i = (Q_i, p_i)$, $i = 1, 2$. We require $Q_1 = Q_2$ and interpret $T_i$ to be the time at which the SP offers a contract to CP$i$[3]. The action space of the CPs is $\{accept, reject\}$, i.e., $a_{CP}^i \in \{accept, reject\}$. We denote the action profile chosen by the players by $\mathbf{a} = (a_{CP}^1, a_{CP}^2, a_{SP})$. With a little abuse of notation, we denote the payoff vector by $(U_{CP}^1(\mathbf{a}), U_{CP}^2(\mathbf{a}), U_{SP}(\mathbf{a}))$.

In addition, the following procedure can be used to further reduce the set of contracts that will be considered by the SP at an NE: Suppose that the SP offers a contract to CP1 first and then to CP2. The other case can be handled in an analogous manner. Given the QoS $Q$, compute the maximum price $p_{\max}^{CP1}(Q)$ that CP1 would be willing to pay at $t = 0$ by setting $U_{CP}^1(\Gamma_{1,2}, 0, m) = U_{CP}^{1,min}$ and the maximum price $p_{\max}^{CP2}(Q)$ that CP2 would be willing to pay at $t = m$ by setting $U_{CP}^2(\Gamma_{1,2}, 0, m) = U_{CP}^{2,min}$, where $\Gamma_{1,2} = ((Q, p_{\max}^{CP1}(Q)), (Q, p_{\max}^{CP2}(Q)))$. One can show that there exists a unique maximum price that each CP is willing to pay and these prices do not depend on each other for any fixed $Q$.

Let $\mathbf{Q} = (Q, Q)$ and $\mathbf{T} = (0, m)$. After some algebra, we obtain

$$
\begin{aligned}
p_{\max}^{CP1}(Q) = \frac{f_{CP}^1}{\int_0^M N_{CP}^1(t; \mathbf{Q}, \mathbf{T}) \, e^{-\tau t} \, dt} \bigg[ & A_1 \zeta_1(Q) + A_2 \zeta_2(Q_{\min}) \\
& + \left( N_{CP}^2(0) - \frac{\zeta_2(Q_{\min})}{\theta} \right) \left( \frac{A_3 \theta - A_4 \tau e^{-\tau M}}{\theta + \tau} \right) \\
& + \left( N_{CP}^1(0) - \frac{\zeta_1(Q_{\min})}{\theta_{max}} \right) \left( \frac{A_5 \theta_{max} - A_6 \tau e^{-\tau M}}{\theta_{max} + \tau} \right) \bigg],
\end{aligned}
\tag{4.12}
$$

---

[3] We assume that the contract takes effect immediately if the CP accepts it

43

$$p_{\max}^{CP2}(Q) = \frac{f_{CP}^2}{\int_m^M N_{CP}^2(t; \mathbf{Q}, \mathbf{T}) \; e^{-\tau t} \; dt} \left[ A_7\zeta_2(Q) + \frac{\zeta_2(Q_{\min})}{\theta} \left( \frac{A_3\theta - A_4\tau e^{-\tau M}}{\theta + \tau} \right) \right.$$

$$+ N_{CP}^2(0) \left( \frac{A_8\theta e^{-\theta m} + A_9\tau}{\theta + \tau} \right) - N_{CP}^2(0) \left( \frac{A_{10}}{\theta_{max} + \tau} \right)$$

$$\left. - \frac{\zeta_2(Q_{\min})}{\theta_{max}} \left( \frac{A_5\theta_{max} - A_6\tau e^{-\tau M}}{\theta_{max} + \tau} \right) \right], \tag{4.13}$$

where

$$A_1 := \frac{1 - e^{-\tau M} - \tau M e^{-\tau M}}{\tau^2}, \quad A_2 := \frac{1 - e^{-\tau m} - \tau m e^{-\tau M}}{\tau^2}, \quad A_3 := \frac{1 - e^{-\tau M} - e^{-(\tau+\theta)m} + e^{-(\tau M + \theta m)}}{\tau},$$

$$A_4 := \frac{1 - e^{-\theta m}}{\tau}, \quad A_5 := \frac{1 - e^{-\tau M}}{\tau}, \quad A_6 := \frac{1 - e^{-\theta_{max} M}}{\tau}, \quad A_7 := \frac{e^{-\tau m} - e^{-\tau M} - \tau(M-m)e^{-\tau M}}{\tau^2},$$

$$A_8 := \frac{e^{-\tau m} - e^{-\tau M}}{\tau}, \quad A_9 := \frac{1 - e^{-(\tau M + \theta m)}}{\tau}, \quad \text{and} \quad A_{10} := 1 - e^{(\theta_{max} + \tau)M}.$$

The goal of the SP is to maximize its payoff by searching over all feasible contracts that ensures the payoffs of at least $U_{CP}^{i,min}$, $i = 1, 2$.

$$\max_{\xi = (\Gamma_{1,2}, \mathbf{T}) \in \Xi_{feas}^{con}} U_{SP}(\Gamma_1 = (Q, p_{\max}^{CP1}(Q)), \Gamma_2 = (Q, p_{\max}^{CP2}(Q)), \mathbf{T}) \tag{4.14}$$

The NE of interest to us is of the form $(a_{CP1}^\star, a_{CP2}^\star, a_{SP}^\star)$, where $a_{CP1}^\star = a_{CP2}^\star = accept$, and $a_{SP}^\star = ((Q^\star, p_{\max}^{CP1}(Q^\star), (Q^\star, p_{\max}^{CP2}(Q^\star), (0, m^\star))$ or $a_{SP}^\star = ((Q^\star, p_{\max}^{CP1}(Q^\star), (Q^\star, p_{\max}^{CP2}(Q^\star), (m^\star, 0))$.

## 4.2 Social Efficiency

Consider a scenario where a social player dictates the actions to be adopted by all the players. The aim of the social player is to maximize the aggregate payoff of all players. Recall that the aggregate payoff depends only on the QoS $Q$ provided by the SP and the times at which the QoS $Q$ will be provided to the CPs.

$$\max_{(Q, \mathbf{T}): Q \in \mathcal{Q}} U_{ag}(\mathbf{Q} = (Q, Q), \mathbf{T}) \tag{4.15}$$

$$\text{subject to} \quad T_i = 0 \text{ for some } i$$

We are interested in determining how efficient the NEs of the Stackelberg game, which has the form mentioned at the end of the previous section, are in relation to the social optimum.

## 4.3 Main Results

One interesting question that arises naturally in the Stackelberg game models discussed in Section 4.1 is whether the Stackelberg game yields an efficient NE *when both contracts are signed* or not. Let us denote the NE of the Stackelberg game with the SP as the leader by $NE_{SG}$

**Theorem 4.** *Both QoS $Q^\star$ and time $m^\star$ chosen at the NE are identical to those chosen by the social player.*

As stated in the previous chapter, an important implication of this finding is that when one of the providers, which is the SP in this case, has a much stronger bargaining position than the CPs and can dictate the terms of the contract, there is no loss of efficiency due to the selfish nature of the SP in the sense that the aggregate payoff of the providers is maximized. Hence, the price of stability is one.

Analogously to Theorem 2, the intuition behind this theorem is that the SP, as a leader, offers a contract that yields the minimum payoff for both CPs. Therefore, it is in the SP's own interest to maximize the aggregate payoff, thereby achieving a Pareto optimal payoff vector at the NEs.

Another question we are interested in is which CP the SP would offer a contract to first. A general rule that holds for all cases is given in the following section. But, for the special case where (a) $f_{CP}^1 = f_{CP}^2$ and (b) $\zeta_1$ and $\zeta_2$ are identical, the following holds:

**Theorem 5.** *Suppose that $f_{CP1} = f_{CP2}$ and $\zeta_1(Q) = \zeta_2(Q)$ for all $Q \in \mathcal{Q}$. Then, the ISP is indifferent as to which CP to approach first.*

For the general case, the answer to the above question depends on many parameters the SP has to take into consideration.

### 4.3.1   Proofs of Main Results

*Proof of Theorem 4:*   The proof follows directly from Theorems 1 and 2 in the previous chapter. The difference between the aggregate utility and the SP's payoff is constant, namely $U_{CP}^{1,\min} + U_{CP}^{2,\min}$ (under the contracts offered by the SP). Since the presence of a constant does not affect the optimization problem, the resulting optimal $Q$ and $m$ coincide at the social optimum and the NE of Stackelberg game.

*Proof of Theorem 5:*   Substituting (4.12) and (4.13) in (4.7) for $p_{\max}^{CP1}$ and $p_{\max}^{CP2}$, respectively, the payoff of the SP when approaching CP1 at $t = 0$ and CP2 at $t = m$ is given by

$$U_{SP}((Q, p_{\max}^{CP1}(Q)), (Q, p_{\max}^{CP2}(Q)), 0, m)$$

$$= f_{SP}\Big[\lambda(Q_{\min}, Q)A_{11} + \lambda(Q, Q)A_{12} - m(\lambda(Q, Q) - \lambda(Q_{\min}, Q))A_8 + N_{SP}(0)A_5\Big]$$

$$+ f_{CP}^1\Big[A_1\zeta_1(Q) + A_2\zeta_2(Q_{\min}) + \Big(N_{CP}^2(0) - \frac{\zeta_2(Q_{\min})}{\theta(Q)}\Big)\Big(\frac{A_3\theta(Q) - A_4\tau e^{-\tau M}}{\theta(Q) + \tau}\Big)$$

$$+\left(N_{CP}^1(0) - \frac{\zeta_1(Q_{\min})}{\theta_{\max}}\right)\left(\frac{A_5\theta_{\max} - A_6\tau e^{-\tau M}}{\theta_{\max} + \tau}\right)\Bigg]$$

$$+f_{CP}^2\left[A_7\zeta_2(Q) + \frac{\zeta_2(Q_{\min})}{\theta(Q)}\left(\frac{A_3\theta(Q) - A_4\tau e^{-\tau M}}{\theta(Q) + \tau}\right)\right.$$

$$+N_{CP}^2(0)\left(\frac{A_8\theta(Q)e^{-\theta(Q)m} + A_9\tau}{\theta(Q) + \tau}\right) - N_{CP}^2(0)\left(\frac{A_{10}}{\theta_{\max} + \tau}\right)$$

$$\left.-\frac{\zeta_2(Q_{\min})}{\theta_{\max}}\left(\frac{A_5\theta_{\max} - A_6\tau e^{-\tau M}}{\theta_{\max} + \tau}\right)\right]$$

$$-\frac{c(Q)}{\eta}(1 - e^{-\eta m})\zeta_1(Q) - \frac{c(Q)}{\eta}(e^{-\eta m} - e^{-\eta M})(\zeta_1(Q) + \zeta_2(Q)),$$

where $A_{11} := \frac{1 - e^{-\tau m} - \tau m e^{-\tau m}}{\tau^2}$ and $A_{12} := \frac{e^{-\tau m} - e^{-\tau M} + \tau m e^{-\tau m} - \tau M e^{-\tau M}}{\tau^2}$.

Due to the symmetry, when the SP approaches CP2 first at $t = 0$ and then CP1 at $t = m$, the payoff of the SP is very similar to the payoff given above except for the following differences: (1) the indices 1 and 2 are reversed, and (ii) $\lambda(Q_{\min}, Q)$ becomes $\lambda(Q, Q_{\min})$. It is reasonable to assume that $\lambda(Q_{\min}, Q) = \lambda(Q, Q_{\min})$.

The SP compares the payoffs it would receive, i.e., $U_{SP}(\ (Q, p_{\max}^{CP1}(Q)), (Q, p_{\max}^{CP2}(Q)), 0, m)$ and $U_{SP}(\ (Q, \tilde{p}_{\max}^{CP1}(Q)), (Q, \tilde{p}_{\max}^{CP2}(Q)), m, 0)$, where $\tilde{p}_{\max}^{CPi}(Q)$, $i = 1, 2$, are the maximum prices the SP would charge if it approached CP2 first, and offers a contract to CP1 first if the first is greater than the latter. After deleting the common terms, this is equivalent to the following inequality.

$$f_{CP}^1\left[A_1\zeta_1(Q) + A_2\zeta_2(Q_{\min}) + \left(N_{CP}^2(0) - \frac{\zeta_2(Q_{\min})}{\theta}\right)\left(\frac{A_3\theta - A_4\tau e^{-\tau M}}{\theta + \tau}\right)\right.$$

$$\left.+\left(N_{CP}^1(0) - \frac{\zeta_1(Q_{\min})}{\theta_{max}}\right)\left(\frac{A_5\theta_{max} - A_6\tau e^{-\tau M}}{\theta_{max} + \tau}\right)\right]$$

$$+f_{CP}^2\left[A_7\zeta_2(Q) + \frac{\zeta_2(Q_{\min})}{\theta}\left(\frac{A_3\theta - A_4\tau e^{-\tau M}}{\theta + \tau}\right) + N_{CP}^2(0)\left(\frac{A_8\theta e^{-\theta m} + A_9\tau}{\theta + \tau}\right)\right.$$

$$\left.-N_{CP}^2(0)\left(\frac{A_{10}}{\theta_{max} + \tau}\right) - \frac{\zeta_2(Q_{\min})}{\theta_{max}}\left(\frac{A_5\theta_{max} - A_6\tau e^{-\tau M}}{\theta_{max} + \tau}\right)\right]$$

$$-\frac{c(Q)}{\eta}(1 - e^{-\eta m})\zeta_1(Q)$$

$$> f_{CP}^2\left[A_1\zeta_2(Q) + A_2\zeta_1(Q_{\min}) + \left(N_{CP}^1(0) - \frac{\zeta_1(Q_{\min})}{\theta}\right)\left(\frac{A_3\theta - A_4\tau e^{-\tau M}}{\theta + \tau}\right)\right.$$

$$+ \left( N_{CP}^2(0) - \frac{\zeta_2(Q_{\min})}{\theta_{max}} \right) \left( \frac{A_5\theta_{max} - A_6\tau e^{-\tau M}}{\theta_{max} + \tau} \right) \Big]$$

$$+ f_{CP}^1 \Big[ A_7\zeta_1(Q) + \frac{\zeta_1(Q_{\min})}{\theta} \left( \frac{A_3\theta - A_4\tau e^{-\tau M}}{\theta + \tau} \right) + N_{CP}^1(0) \left( \frac{A_8\theta e^{-\theta m} + A_9\tau}{\theta + \tau} \right)$$

$$- N_{CP}^1(0) \left( \frac{A_{10}}{\theta_{max} + \tau} \right) - \frac{\zeta_1(Q_{\min})}{\theta_{max}} \left( \frac{A_5\theta_{max} - A_6\tau e^{-\tau M}}{\theta_{max} + \tau} \right) \Big]$$

$$- \frac{c(Q)}{\eta} (1 - e^{-\eta m}) \zeta_2(Q).$$

This gives us the general rule for the SP to decide which CP to approach first. When $f_{CP}^1 = f_{CP}^2$ and $\zeta_1(Q) = \zeta_2(Q)$ for all $Q \in \mathcal{Q}$, all the terms that are independent of $N_{CP}^1(0)$ and $N_{CP}^2(0)$ will cancel each other out. After a little algebra, we would be left with

$$(N_{CP}^1(0) + N_{CP}^2(0)) \left[ \frac{A_5\theta_{max} - A_6\tau e^{-\tau M}}{\theta_{\max} + \tau} \right]$$

on both sides. Therefore, the ISP receives the identical payoff whether it approaches CP1 first or CP2 first. This proves Theorem 5.

Chapter 5:   Conclusions and Future Work

We studied both the interactions between a single SP and a single CP, and a single SP and two CPs and showed that the aggregate payoff and QoS at the NEs of Stackelberg games do not degrade due to selfish nature of these providers when one of them has a much stronger bargaining position and their interaction can be modeled as a Stackelberg game.  Furthermore, we demonstrated that the Pareto optimal QoS and the NE QoS are the same.  Finally, we demonstrated that the Pareto optimal QoS improves as the contract duration increases under mild conditions for the monopoly market case.  For the market with two content providers we have shown a strategy by which the SP determines which CP to approach first in order to increase its payoff.

There are multiple ways to improve our model and make it more realistic. In our study, we have assumed that none of the CPs increase their subscription fee after it signs a contract with the SP, i.e., CP subscribers receive a higher QoS for the same price. This led to the other CP being obligated to sign a contract as well with the SP. However, it is unclear what would happen if the CP that signed a contract would increase its subscription fee.  Some of the users might find the subscription fee too high for its budget and so would turn away to the CP with the lower QoS

(and lower subscription fee). Therefore, there might be a NE where one of the CPs decides to not sign a contract.

Another improvement would be that each CP agrees on a different QoS with the SP. In our model, the SP provides only one QoS to both CPs. This might not be optimal as one of the CPs may need a higher QoS than the other CP. It is not directly clear to say which CP is to be approached first.

# Bibliography

[1] M. Andrews, U. Ozen, M.I. Reiman and Q. Wang, "Economic models of sponsored content in wireless networks with uncertain demand," Proc. of IEEE INFOCOM Computer Communications Workshops, pp. 345-350, Apr. 2013.

[2] D.P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1995.

[3] P. Bolton and M. Dewatripont, *Contract Theory*, The MIT Press, Cambridge (MA), 2005.

[4] E.F. Brigham and P.R. Daves, *Intermediate Financial Management*, 10th edition, Cengage Learning, 2009.

[5] H. K. Cheng, S. Bandyopadhyay and H. Guo, "The debate on net neutrality: a policy perspective," *Information Systems Research*, 22(1):60-82, Mar. 2010.

[6] J. P. Choi and B.-C. Kim, "Net neutrality and investment incentives," *RAND Journal of Economics*, 41(3), Mar. 2010.

[7] D. Fudenberg and D.K. Levine, *The Theory of Learning in Games*, The MIT Press, 1998.

[8] B. E. Hermalin and M. L. Katz, "The economics of product-line restrictions with an application to the network neutrality debate," *Information Economics and Policy*, Elsevier, 19(2):215-248, Jun. 2007.

[9] E. Koutsoupias and C.H. Papadimitriou, "Worst-case equilibria," Proc. of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS), pp. 404-413, 1999.

[10] J. Krämer and L. Wiewiorra, "Network neutrality and congestion sensitive content providers: implications for content variety, broadband investment, and regulation," *Information Systems Research*, 23(4):1303-1321, May. 2012.

[11] M. A. Lemley and L. Lessig, "The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era," *UCLA Law Review*, Vol. 48, p. 925, 2001.

[12] J. B. Speta, "Handicapping the Race for the Last Mile?: A Critique of Open Access Rules for Broadband Platforms," *Yale Journal On Regulation*, 2000.

[13] R. J. La and J. Mo, "Interaction of service providers in task delegation under simple payment rules," Proc. of IEEE Conference on Decision and Control (CDC), Dec. 2009.

[14] J.-J. Laffont and D. Martimort, *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press, Princeton (NJ), 2002.

[15] Z. Liang and D. Wang, "Sponsoring content: motivation and pitfalls for content service providers," Proc. of IEEE INFOCOM Computer Communications Workshops, pp. 577-582, Apr. 2014.

[16] J.F. Nash, "The bargaining problem," *Econometrica*, 18:155-162, 1950.

[17] R. S. Lee and T. Wu, "Subsidizing Creativity through Network Design: Zero-Pricing and Net Neutrality," *Journal of Economic Perspectives*, Vol. 23(3), pp. 61-76, 2009.

[18] T. Wu, "Network Neutrality, Broadband Discrimination," *Journal of Telecommunications and High Technology Law*, Vol. 2, p. 141, 2003.

[19] C. S. Yoo, "Would Mandating Network Neutrality Help or Hurt Broadband Competition? A Comment on the End-to-End Debate," *Journal of Telecommunications and High Technology Law*, Vol. 3, 2004.

[20] M. Armstrong," Competition in Two-Sided Markets," *RAND Journal of Economics*, V ol. 37(3), pp. 668-691, 2006.

[21] J-C. Rochet and J. Tirole, "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, Vol. 1(4), pp. 990-1029, 2003.

[22] N. Economides and J. Tåg, "Network Neutrality on the Internet: A Two-Sided Market Analysis," *Information Economics and Policy*, Vol. 24, 2012.

[23] J. Musacchio, G. Schwartz and J. Walrand "A T wo-Sided Market Analysis of Provider Investment Incentives with an Application to the Net-Neutrality Issue," *Review of Network Economics*, Vol. 8(1), pp. 22-39, 2009.

[24] M. K. Chen and B. Nalebuff, "One-Way Essential Complements," Cowles Foundation Discussion Paper No. 1588, Yale School of Management, 2006.

[25] V. Kocsis and P. W. J. de Bijl, "Network Neutrality and the Nature of Competition between Network Operators," *International Economics and Economic Policy*, Vol. 4(2), pp. 159-184.

[26] G. Owen, *Game Theory*, 3rd edition, Academic Press, 1995.

[27] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, 1976.