Service Revenue Management in the Presence of Grouping Complementarities

by

Mohammadreza Farahani

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2015

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis studies a revenue management problem faced by providers of differentiated and congested services when customer preferences are subject to externalities. In particular, a customer benefits not only from receiving the offered service, but also from participating in the service with other customers. The provider offers multiple services differentiated by a measurable attribute and customers have idiosyncratic preferences for this attribute. The preference to conform around a particular offered service can nonetheless, cause congestion as customers systematically avoid less popular options. Such congestion increases the marginal cost of serving a customer. For this setting, we address the following questions: When can product differentiation reduce the congestion in the system? If the service provider can differentiate customers, what is the optimal level of differentiation? When is the optimal level of differentiation a Nash equilibrium and when is that Nash equilibrium stable (i.e. not disrupted by a deviation by a small number of customers)? We consider first a simple model where idiosyncratic customer preferences are uniformly distributed on a determined interval and the provider places up to two differentiated services. We then consider more general settings including allowing arbitrary customer distributions, increasing the number of possible differentiated services and associating provider revenue with the differentiation attribute. The model was initially motivated by a problem in the gaming industry but extends to many other types of service providers where customers have similar attributes such as online video games and other entertainment settings.

Acknowledgements

I would like to express my gratitude to:
My supervisor, Dr. Michael Pavlin, who has supported me though out my thesis with his patience and knowledge whilst allowing me the room to work in my own way. I attribute the level of my Masters degree to his encouragement and effort and without him this thesis, too, would not have been completed or written.
My supervisor, Dr. Kumaraswamy Ponnambalam, for his insightful guidance, caring, patience, and providing me with an excellent atmosphere for doing research.
My committee members, Dr. Benny Mantin and Dr. Nasser Lashgarian Azad, for finding time in their busy schedules to help me at my graduate career.

Table of Contents

List of Tables viii

List of Figures ix

1 Introduction 1

1.1 Initial Problem 1

1.2 Concepts 3

1.3 Organization of the Thesis 4

2 Literature Review 5

2.1 Introduction 5

2.2 Customer Externality 5

2.3 Product Differentiation 8

2.4 Revenue Management 9

2.5 Contributions 10

3 A 2-Server Model 11

3.1 General Settings 11

3.1.1 Service Provider's Setting 11
3.1.2 Customers' Setting 13
3.2 A 2-Server Model with Uniform Distribution of Customers 15
3.2.1 Revenue Function 15
3.2.2 Customers' decision model 18
3.2.3 Partition Set Feasibility 23

4 The Generic Model **28**
4.1 Analysis of a Model With an Arbitrary Continuous Distribution 29
4.1.1 The indifference point and partition set relation 29
4.1.2 Stability of a partition set with 2 servers 30
4.2 Stability of a 2-server Model with Uniform distribution of customers 35
4.2.1 Mathematical Model 38
4.3 Analysis of a Multi-server Model 41
4.4 Empirical Study 46

5 Results **49**
5.1 Differentiation Strategies for 2-server model 49
5.2 Differentiation Strategies for multi-server model 53

6 Conclusions **56**
6.1 Conclusions and Summary 56
6.2 Future Research and Recommendations 57

APPENDIX **59**

A Empirical Study Parameters **60**

References **61**

List of Tables

A.1 Game Settings for all Game Types 60

List of Figures

3.1 Position of indifference points with regard to the servers' position in utility line . . . 22

3.2 Max and Min value of indifference point based on β $(N = 5, \alpha = 0.14)$ . . . 24

3.3 Maximum revenue based on β $(\alpha = 0.14,\ P_1 = 2.25,\ P_2 = 2.1)$ . . . 25

3.4 Maximum revenue of the service provider with regard to the size of demand $(\beta, \alpha = 0.2,\ P_1 = 2.25,\ P_2 = 2.1)$ . . . 26

4.1 Revenue contour and stability constraint . . . 36

4.2 Revenue of service provider subjected to different constraints . . . 37

4.3 Relationship between indifference points and partition set with 3 server . . . 46

5.1 Optimal Differentiation with respect to the conformity level,β, $(\alpha = 0.12, N = 2)$ . . . 50

5.2 Optimal differentiation level with respect to the demand size $(N)(\alpha = 0.09, \beta = 0.2)$ . . . 51

5.3 Maximum revenue of the service provider with regard to the size of demand $(\alpha = 0.08, \beta = 0.4)$ . . . 52

5.4 Optimal differentiation level with respect to the congestion penalty intensity $(\alpha)[N = 2, \beta = 0.2]$ . . . 53

5.5 Difference of current policy with the optimal policy for the game type #1 . 54
5.6 Difference of current policy with the optimal policy for the game type #4 . 55

Chapter 1

Introduction

This introductory chapter provides outlines the primary components of the problem, a brief overview of the concepts we employed in this study, and finally offers a breakdown and guideline for how this thesis should be read.

1.1 Initial Problem

For certain types of services, customers receive greater utility when they participate in the service in larger groups. This effect can caused by different reasons. For example, in night clubs customers prefer the crowded club, because they have more "*fun*", or among some similar restaurants they choose the crowded one, because the number of restaurant's customers can imply the perception of better food quality for them.

These phenomena produce "*conformity*" between customers and consequently one of multiple offered services may absorb most of the load resulting in congestion. In many cases, congestion can reduce the marginal revenue of each independent server. This work considers how differentiating services can alleviate such negative effect. For example, an online game company offers a variety of differentiated services ,e.g., different card game. When there are more players at a server, fewer hands can be played resulting in lower revenue.

This state of affairs occurs frequently even with ample capacity because players often prefer to play in larger groups. The managerial challenge is to determine which games should be opened to maximize the revenue. Is it more profitable to open more of the same and highest margin games with the highest rate of revenue or different games with a different rate of revenue? Offering similar services may increase the probability of suffering congestion while offering dissimilar services may do a better job at segmenting the customer base but will require introducing less profitable services. This problem will appear when customers are charged à la carte and select from a menu of offered services. These problems are particularly prevalent in entertainment contexts, such as casino gaming and multiplayer video games.

This study addresses the following questions:

1. When can service differentiation improve system revenues?
2. If the service provider can differentiate (partition) customers, what is the optimal level of differentiation?

We consider first a simple model where idiosyncratic customer preferences are uniformly distributed on the [0,1] interval and the provider offers up to two differentiated services. We then consider more general settings including allowing arbitrary customer distributions, increasing the number of possible differentiated services and associating provider revenue with the differentiation attribute. The model was initially motivated by a problem in the casino gaming industry but extends to many other types of service providers where customers have similar attributes, such as online video games and other entertainment settings.

1.2 Concepts

The main concept that this thesis tackles is *externality*. In the context of economics, externality refers to the manner in which the cost or utility function is influenced not only by the individual player's action, but also by the actions of players outside the individual [10]. The focus of this study is the effect of externality among customers, which means that a customer's optimal strategy is ultimately dependent on the whole demand. Externality can either reinforce or weaken the utility of customers.

Using the game theory perspective to model the system with externalities is an acceptable technique among the researchers [3, 13]. This approach can help us to define stable solutions by adopting the *Nash equilibrium* to our problem. In addition, the concept of externality leads us to employ the *complementarity* optimization model. The complementarity optimization model appears when many optimization problems share a variable. Without the presence of externality, each customer solves an independent optimization model to find the best solution; however, in our case we need to solve all customers' optimization problems together. Therefore it is necessary to employ the complementarity optimization model.

1.3 Organization of the Thesis

We complete next chapter by reviewing the previous works and outlining the significance of the current study. In the next chapter, we build the initial 2-server model which forms the foundation of the extended model. We introduce most of the necessary concepts and lemmas based on this model, and then we expand them as needed. Also, we offer a simple computational methodology to find the exact solution for the 2-server model in an expeditious time.

In Chapter 4, we establish a general structure of the problem to capture a more extensive and comprehensive characteristics of the problem, followed by a return to the initial problem to investigate this findings on it. This generalization is mainly based on two of previous assumptions: a. the maximum number of servers which the service provider can open, and b. the distribution of customers. For both cases, we extend the notions of feasibility and stability, and then we establish a mathematical optimization model to find the optimal solution with presence of these constraints. At the end of the chapter, we introduce an applicable model for an online game service that offers different game types with different settings.

In the last chapter, we review the effect of conformity and differentiation on the solution space and the optimal solution in particular. Moreover, we compare the current policy of service provider with the optimal differentiation policy to determine their advantages and disadvantages in counteracting congestion.

Chapter 2

Literature Review

2.1 Introduction

This section discusses the three essential streams of research studies that we are going to build this thesis from. First, we survey the research studies around the concept of customer externality, i.e., when the customer's decision is subjected to the other customers' action. Then, we explore the findings in the area of product differentiation and customer segmentation, and finally, we investigate the chronicle of the Revenue Management Problem.

2.2 Customer Externality

The first serious discussions about customer externalities in economic literature emerged in 1899 [31]. In his book *"The Theory of the Leisure Class"*, Thorstein Veblen studied how consumption of a good by a set of customers can change the utility of that good for other customers. He used the word *Conspicuous consumption* to capture the demand to acquire luxury goods that is used to exhibit the consumer's wealth. He believed this consumption in the middle class depends on the consumption of the upper class. Veblen was not the first to mention the effect of customer consumption on the demand, but his work became the basis of many other studies in this area.

In 1950, Harvey Leibenstein classified the external effects on utility into three major groups: *Bandwagon effect, Snob effect* and *Veblen effect.* Two of them, *Bandwagon* and *Snob* effect, are the externalities caused by the customer consumption and the other, Veblen effect, caused by price. Leibenstein referred to the Bandwagon effect as a state in which consumption of a product by others reinforces the utility, and the Snob effect as a state in which the increase of demand reduces the utility [19].

Rather than these two externalities, we can also count the network effect as a different type of externality motivated by the customers' consumption or subscription. The network effect describes a situation where the utility of customers increases when the number of consumers of the same or compatible product increases. This effect is commonly used to show the increase of utility in the product which employing them depends on the existence of that specific technology among others. The famous example of network effect is telephone, in which you gain no utility of a telephone device until somebody else owns at least one. In this study we mainly focus on the positive externalities.

The studies about the customer externality continued in three major groups: behavioral studies, economic theory and market welfare [19]. The first group tried to model and examine the effect of externality on individuals. In 1951, S. Asch started a series of experiments to understand the effect of conformity on an individual's decisions. He found a positive bond between the size of a group and the level of conformity [2]. These series of experiments later became well known as the Asch conformity experiments or the Asch Paradigm.

Asch demonstrated the existence of the positive externality, but why does conformity change the demand? The answer to this question varies in different contexts. Bernheim categorized the studies about the causes of conformity into two main groups: a. the *informational cascade,* and b. the *mutually reinforcing.* The first theory suggests that customers follow the act of others because they believe that the customers ahead of them have more information [6]. For this setting, Bernheim referred to the work of Bikhchandani *et. al.* in which they stated that when the individuals observe the act of those ahead of them, this can change their perception of the product and can cause an informational cascade [8]. The next theory he mentioned is the mutually reinforcing externalities which captures the positive effect of grouping on the the customers' utility [6].

Becker also conducted a study about a restaurant in Palo Alto, California that captures the business in that area without having an ex-ante technology. He believed that a customer decision is subjected to the action of players ahead of them, because they believe they may have more information and this can cause a snowball effect which he called the informational cascade [4]. More recently, Grilo *et. al.* offered a spatial duopoly competition model to study the effect of conformity and vanity in the market. They find out by rise of demand when there is conformity among customers, the duopoly market can deform to a monopoly [17].

Aside from these practical studies, the concept of externalities in theoretical economy has been studied for a long time, and in this area we can refer to the work of Sasaki which studied the two-sided matching problems with externalities. He defined the concept of stable matching and compared the Pareto optimal solutions with Nash equilibrium solutions and also offered a strategy to find them [24].

2.3 Product Differentiation

The second stream of studies refers to the idea of product differentiation which commonly has been used in marketing literature. The modern concept of product differentiation was introduced by Edward Chamberlin in his book *Theory of Monopolistic Competition* in 1933 [12]; however, some researchers believe that the primary indications to this subject is commenced by the work of Shaw in 1912. Shaw stated that product differentiation can fit the needs of customers more efficiently [26]. Nonetheless, most researchers accept the fact that the modern usage of product differentiation to adjust the demand curve is mainly credited to Chamberlin's astonishing works. He argues that differentiation can reduce the elasticity of customers to the price. Afterwards, researchers conducted studies on the different methods of product differentiation. For example, P. Danaher studied the cases of Airline and Telecommunication while there are customer heterogeneity [14]. However, product differentiation can be based on any perceivable attribute of product [28], but in the empirical studies, product differentiation is usually associated with the pricing strategies [7].

2.4 Revenue Management

The last stream of literature explores the related Revenue Management Problem studies. The most revenue management studies, the managerial tools that a firm can control –i.e, decision variables– are divided into three main categories: structural decisions, price decisions, and quantity decisions [30]; however, the share of pricing is significantly more than the others. The practice of revenue management is based on the previously overbooking management research studies [20]. In 1970s, after the deregulation Act of 1978 which let the airlines to price their services, a series of novel studies began in this field [30]. In most of the early studies, researchers assumed that the demand for each service is independent of other services [29]. Later on, researchers offered theoretical dynamic pricing models which considered the demand for each product based on the offered price of the product [16]. These studies initiate a new field of research in revenue management with the focus on the customers behavior.

In addition, research studies about the customer choice model, wherein customers can choose a product from a set of products, enriches the revenue management studies. These studies started in the 1970s, and later in 1985, Ben-Akiva and Lerman published the "*Discrete choice analysis : theory and application to travel demand*" [5]. Afterwards, *Anderson et. al.* fused the idea of customer choice behavior and product differentiation [1]. Following this developement, researchers used this model in the area of revenue management to model the customers' behavior in a more realistic context [29][32] .

2.5 Contributions

To the best of our knowledge, this research is the first to study the concept of externalities and congestion from a revenue management perspective. We offer managerial guidance in the use of differentiation as a tool to alleviate the negative effects of conformity. If we look back at the literature, we will notice that the research studies about customer externalities are mostly looking at this issue from the economical point of view, and they do not offer any solution in the instance of monopoly competition. In the area of revenue management, the main focus is designated to the pricing strategies and the concept of strategic customers. In this study, we tried to merge these notions together and look at them from a game theoretic perspective to offer a new revenue management approach with a special focus on the entertainment services.

Chapter 3

A 2-Server Model

3.1 General Settings

In this chapter, we initially introduce the basic concepts and definitions, then we sketch a basic model of problem, and finally, we analyze this model to investigate: a. how conformity affects a service provider, and b. whether service differentiation counteracts these effects. This model includes a parsimonious monopoly service provider with a fixed demand, who tries to maximize its revenue. We divide the general settings into two sections: The Service Provider's Setting and the Customers' Setting.

3.1.1 Service Provider's Setting

The service provider can offer different services on its available servers. We differ services based on a measurable continuous attribute, e.g., the complexity of a game, position of an information kiosk in a mall, or even the spicy level of a meal which is offered in a restaurant. To establish a generic model, we map the quantity of this attribute to a 0-1 interval for all possible outcomes. With this assumption, we can measure the difference of two services in a normalized format. To illustrate this concept, we define a *specification line.*

Definition 1. Specification line is a limited line where each of its points represents a particular service by addressing its specification or attribute.

This limited attribute is the only characteristic that the service provider can control to offer different services. Therefore, we exploit this attribute as a decision variable of the service provider.

We define a set of servers $\mathcal{S} = \{s_1, s_2, ...s_m\}$. Size of this set shows the number of services which the service provider can simultaneously offer. Given x_i is the position of server i on the specification line, the decision of a manager for offering the different services will be shown by a decision set $X_s = \{x_1, x_2, ..., x_m\}$. Aside the position of each server on the specification line which indicates the service it offers, each server s_i has a net income coefficient p_i. This variable shows the marginal revenue of server i when there is no load on it. The value of this variable can be a function of either the service or the server itself.

The next important concept is the *cost of congestion.* We assume that the marginal revenue of a server decreases as the number of customers increases. We model this inverse relationship with a linear function. The slope of this linear function is determined by α, the congestion penalty coefficient.

Instead of offering a cost function and a revenue function, we directly address a net income function. This decision is made based on two assumptions: first, we assume that offering different services has a constant cost independent from the demand size, and second, changing the service attributes has zero or negligible cost.

If we look at this problem from a game theoretic prospective, the service provider is the first player in the game; it also has perfect information about the number of customers and their utility functions. Based on this data, it wants to assign a set of services to its available servers to not only maximize the revenue but also create a stable solution which is not fragile in the facing of small arbitrary changes.

3.1.2 Customers' Setting

In this game, there is another set of players; these rational players are the customers. As with the servers, each subset of customers is defined by an attribute which is derived from their ideal service they have described. This attribute is analogous to the service specification. For instance, if we determine each service by its complexity, we should determine each subset of customers by the ideal complexity they want. For later computations, we also normalize this attribute and map it to the specification line. We use y_j to represent the position of a subset of customers c_j with an identical attribute on the specification line.

By using specification line properties, we can simply calculate the *intrinsic utility* of each customer from choosing a particular service, by measuring the distance between them on the specification line.

Definition 2. The intrinsic utility is a utility which customers gain solely from receiving a service or consuming a product [6].

We define the intrinsic utility by function $\Gamma(y_j, x_i)$. This function represents the utility of customers in subset c_j from receiving a service offered by server i. Traditionally, using an inverse function of distance for stating the utility is common in the literature [17], and for this chapter, we will employ a similar model; however, we try to avoid spatial function labeling for our utility function. This labeling unnecessarily constraints our utility function to a specific set of functions.

It should be considered that the total utility of customer is formed from two distinct parts: the intrinsic utility, which we discussed, and *conformity utility*.

Definition 3. The conformity utility is the utility of customers from jointly receiving a service.

The conformity utility has a positive relationship with the number of customers that are using a server simultaneously. We simply assume a linear relationship between the conformity utility and demand of server. The previous studies offered different ways to formulate this utility, for instance, for the network effect there is a strong trend for using exponential functions [27] but it is important to distinguish between the conformity externality and network effects. However, as far as we keep a positive continuous relationship between these two variables the analogy of results will not dramatically change.

To model this linear function, we use the coefficient β to capture the level of conformity. A bigger β means that people are more interested in joining a group with more customers. This incident may especially happen when customers have a previous tie to each other, e.g., they are friends. The battle of sexes[1] is an example of this situation, in which the utility of being together is big enough to eclipse the idiosyncratic preferences of players [22]. In contrast, when β is zero, each customer simply chooses the closest open server.

As mentioned, the demand is known and fixed. We capture the demand by a continuous function over the specification line.

3.2 A 2-Server Model with Uniform Distribution of Customers

We begin by considering a parsimonious model where the service provider can open at most two heterogeneous servers with net income coefficients of p_1 and p_2. We assume that the difference between these two values originated from the servers, not the services.

3.2.1 Revenue Function

The revenue of each particular server depends on its net income coefficients, its demand, and the congestion penalty coefficient. To calculate the revenue for the server i, we employ the formulation below:

$$\pi_i = n_i(p_i - \alpha n_i) \tag{3.1}$$

[1]Battle of sexes is a well-known coordination game with two players. In this game each player can choose a place for a date night. Despite each of them has a different preference, they gain more utility when they share a strategy [11].

In this formulation, n_i and π_i are respectively the demand and the revenue of server i. We interpret $(p_i - 2\alpha n_i)$ as the marginal revenue of server, which has a negative linear relation with the server's demand. Based on the previous assumption, α and p_i are fixed and defined; hence, the revenue will be a quadratic function of the server's demand size.

The service provider is a monopoly that can satisfy all its demand, so the aggregation of all its servers' demand should be equal to the total demand. In the model with two servers where the demand size is N, we can write: $n_1 + n_2 = N$. In other words, opening two servers will partition our set of customers into two subsets and generates a *partition set*.

Definition 4. A partition set, $\mathscr{P} = \{c_1, c_2, ..., c_m\}$, is a set that divides customers into non-overlapping subsets; each corresponds to a server.

In this definition c_i is subset of customers and n_i is the size of this subset. The size of each of these subsets specifies the demand of each server; therefore, the revenue of all servers can be calculated when the partition set is given. In this chapter the partition set has only two members: $\mathscr{P} = \{c_1, c_2\}$. Since the calculation of revenue for each server is independent, the total revenue of a service provider is calculated by a summation over all servers' revenue:

$$\Pi = \pi_1 + \pi_2 \tag{3.2}$$

$$\Pi = n_1(p_1 - \alpha n_1) + n_2(p_2 - \alpha n_2) \tag{3.3}$$

Owing to the perfect partition set, we can write $n_1 + n_2 = N$, then we can replace n_2 with $N - n_1$, and calculate the maximum of Π by applying the first order condition. The results are:

$$n_1 = \frac{p_1 - p_2 + 2\alpha N}{4\alpha} \tag{3.4a}$$

$$n_2 = \frac{p_2 - p_1 + 2\alpha N}{4\alpha} \tag{3.4b}$$

The second order condition also assures that the converged supremum point is a maximum in the continuous revenue function.

Lemma 1. *If the difference between two servers' net income coefficient becomes greater than $2\alpha N$, then the service provider only needs to open a server with higher net income coefficient.*

Proof. To have a feasible solution, it is necessary to restrain n_i between 0 and N. If we solve this restriction for p_1 and p_2, we will obtain an inequality described by equation 3.5; this restriction secures us a non-negative partition set.

$$|p_1 - p_2| < 2\alpha N \tag{3.5}$$

□

Equation 3.5 states that when the difference of the net income coefficients is greater than $2\alpha N$, the congestion penalty cannot cover this income gap. Thus, even with the full congestion, the marginal income of one server remains higher than the other yet; in such cases, the trivial answer is opening one server with the higher value of p. We assume this constraint is always satisfied.

3.2.2 Customers' decision model

The total utility of customers will be the aggregation of the intrinsic utility and the conformity utility. For defining the intrinsic utility function, we limit the specification line between zero and one. Then, we use a quadratic model to define the intrinsic utility function based on the distance of customer's ideal service from an offered service:

$$\Gamma(c_j, s_i) = 1 - (y_j - x_i)^2 \tag{3.6}$$

where x_i and y_j are respectively the position of server s_i and subset c_j of customers. This definition ensures the highest utility for a customer when the attribute of a server perfectly matches with the customer's expectation or simply when $y_c = x_i$ on the specification line. The utility starts decreasing proportionally as the gap between the customer and the server increases. This function secures a non-negative utility for the customers with the maximum utility of one. We can alter this model by changing the possible interval of specification line and the maximum value of utility.

The function of conformity utility is determined by a conformity coefficient. This coefficient demonstrates the level of customers dependency on the other customers' decision when they are choosing between the two servers. Based on this coefficient, we build the conformity utility, Λ, as a function of server's demand:

$$\Lambda(s_i) = \beta n_i \tag{3.7}$$

Finally, the total utility of a customer or a subset of customers from a server is defined as:

$$U(c_j, s_i) = 1 - (y_j - x_i)^2 + \beta n_i \tag{3.8}$$

Definition 5. Indifference point (y_{ind}) is a point on the specification line where the utility of two servers are equal for the possible customers in that position.

The indifference point is driven by calculating the crossover of utility curves for both servers:

$$1 - (y_{ind} - x_1)^2 + \beta n_1 = 1 - (y_{ind} - x_2)^2 + \beta n_2 \tag{3.9}$$

Given c_{ind} is an imaginary subset of customers at the indifference point; therefore, the above formulation can be stated as:

$$U(c_{ind}, s_1) = U(c_{ind}, s_2). \tag{3.10}$$

By the current definition of utility function, the indifference point is a unique point on the specification line for a known distribution of customers, position of servers, and a set of net income coefficients. For all known servers, this statement is always valid except when both servers offer the exact same service, which in that case there will be infinite indifference point on the line. For now we only focus on the case where $x_1 \neq x_2$.

After solving equation 3.9 for y_{ind}, there is one specific answer for the indifference point:

$$y_{ind} = \frac{[(x_1)^2 - (x_2)^2] - \beta(n_1 - n_2)}{2(x_1 - x_2)} \tag{3.11}$$

Lemma 2. *In any arbitrary interval on the specification line, all customers have same server preference if there is no indifference point in that interval.*

Proof. Given two points, y_a and y_b on the specification line, customers at y_a, and y_b respectively prefer server A, and B. Due to continuity of utility functions, we apply intermediate value theorem to prove the existence of a point like y_c between y_a and $y_b$which the possible customers at that point receive equal utility from both servers. □

Without loss of generality, assume $x_1 < x_2$. By using the Lemma 1, when y_{ind} is the indifference point, we can derive the partition values by using Equation 3.12:

$$n_1 = \int_0^{y_{ind}} f(y), \; n_2 = \int_{y_{ind}}^1 f(y) \tag{3.12}$$

Where $f(y)$ is the distribution of customers over the specification line. When the distribution is uniform, the above equation can be simplified to:

$$n_1 = N \times y_{ind} \; and \; n_2 = N \times (1 - y_{ind}) \tag{3.13}$$

There are two possible positioning arrangements for a given indifference point: it may rest either between the two servers, or outside of them. In the both situations and under the previous assumption, y_{ind} will be exactly $\frac{n_1}{N}$.

After plugging Equation 3.13 into Equation 3.11 and solve it for x_1 and x_2, we can find the position of servers with respect to the position of the indifference point. The result will be a hyperbolic curve defined by Equation 3.14:

$$[(x_1)^2 - (x_2)^2] - 2y_{ind}(x_1 - x_2) - \beta N(2y_{ind} - 1) = 0 \tag{3.14}$$

If we solve equation 3.14 for y_{ind}, we will find the position of indifference point only by knowing the position of servers. The equation below formulate this finding, consider that by adding absolute condition to this formulation, there is no need to keep the $x_1 < x_2$ constraint.

$$y_{ind} = \frac{|x_1^2 - x_2^2| + \beta}{2(|x_1 - x_2| + \beta)} \tag{3.15}$$

Equation 3.15 is always exclusively true for all value of (x_1, x_2); However, when $x_1 = x_2$ it is not the only answer. When we place two servers in a same position as far as half of the customers choose server 1 and other half choose server 2, the answer is feasible and there is no obligation to specify a particular indifference point on the specification line. Indeed, if we look back to Equation 3.11, we can argue that all the points in the utility line are the indifference point when the service provider offers identical services.

In figure 3.1, each axis represents the position of one of the servers on the specification

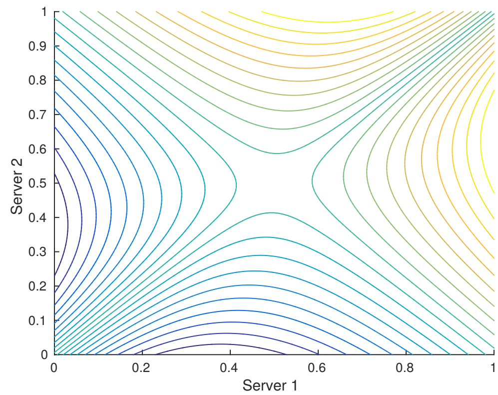

Figure 3.1: Position of indifference points with regard to the servers' position in utility line

line, and the contours are representing the position of indifference points based on the quadratic utility function and a fixed value for β and N. Contours are symmetrical with respect to the line $x_1 = x_2$, which means replacing the services of two server will not change the position of the indifference point. Also, as both servers share a homogeneous conformity coefficient, the isoprofit contours keep an analogous structure as figure 3.1; however, the symmetry reference changes from the line $x_1 = x_2$ to the center point of the utility space. This means by swapping the position of servers, the indifference point remains the same but the revenue will change; even though, mirroring the position of servers based on center of specification line will not change the revenue.

When both servers are in the exact same position –servers are offering homogeneous services– two possible outcome can happen: either all the customers choose one of the servers or each exact half choose one. In the later case, there is no specific indifference point and the decision of clients is not a function of their position in the utility line.

3.2.3 Partition Set Feasibility

We analyzed the relation of server's position on the revenue of service provider, and how our solutions produce the indifference points and consequently the partition sets. In this section, we discuss how the positioning set generates the partition set, and what are the limitations.

In Equation 3.15, as x_1 and x_2 are limited between 0 and 1, some limitations take place for the possible position of the indifference point. These limitations can be translated to restrictions on the segmentation ratio of partition set's subsets size to the demand size, which from now we just refer to it as the s*egmentation ratio*. For instance, when an indifference point is limited between 0.2 and 0.8, it is not possible to form a subset of customers with the size of less than 20 percent of whole demand size.

To find out how these limitations work, we calculate the maximum value of indifference point by solving the following non-linear set of equations:

$$\max \ y_{ind} = \frac{|x_1^2 - x_2^2| + \beta}{2(|x_1 - x_2| + \beta)} \tag{3.16a}$$

$$X_s \in [0, 1] \tag{3.16b}$$

As With, we can find the minimum value, and repeat it for various amount of β to investigate the adjustments of solution space with regard to the changes of the conformity level.

Figure 3.2 illustrates the maximum and minimum of possible segmentation ratios with respect to the value of β. The lower line shows the minimum position of the indifference

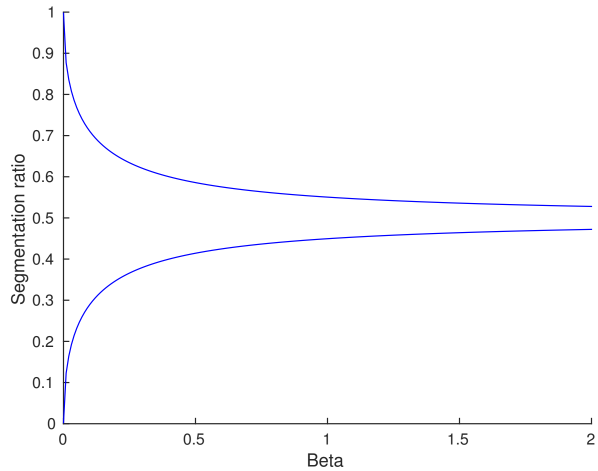

Figure 3.2: Max and Min value of indifference point based on β ($N = 5, \alpha = 0.14$)

point and the upper line likewise demonstrates the maximum position. As is shown in the figure when the conformity level is zero, there is no segmentation limit for customers but by intensifying of the conformity level the solution space becomes more restricted. The growth of demand has the same effect on the feasible boundaries.

This constraint can make our optimal answer, which is derived from Equation 3.4a and 3.4b, infeasible, and generates a new optimal answer on the constraint lines. When the optimal answer is an extreme point, any changes on conformity level affect the optimal value.

To calculate the revenue, we need to solve a highly nonlinear optimization model. In this chapter, owing to the small size of the problem, we use a computational model to find the optimal solutions. We will scrutinize this issue in the next chapter; nevertheless, we use the output of this methodology to analyze the behavior of revenue in the several cases.

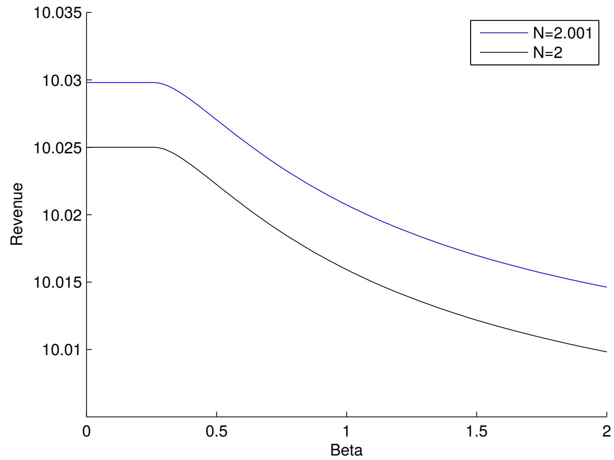

Figure 3.3: Maximum revenue based on β ($\alpha = 0.14,\ P_1 = 2.25,\ P_2 = 2.1$)

Figure 3.3 demonstrates the revenue of a service provider regards to the value of conformity level. The axes x and y respectively represent the value of β and the maximum feasible revenue of our sample system. As it has been shown in the figure, in this case, the growth of β does not affect the best answer at first, where β is less than 0.4, after that there is a continuous decrease in the maximum revenue, which shows the optimal answer is in the feasible edge region. Note that the breakdown point depends on various parameters and in some cases it can be at the very starting point. This figure also illustrates the changes in revenue based on demand compared to the changes based on conformity level. As it is seen the small changes in the demand can significantly affect the demand compare to the conformity level.

To study the changes of revenue with regard to demand, we solve the maximum revenue problem for different sizes of the demand. As demand increases, we pass through three phases. First, the constraint 3.5 is not active yet; therefore, the optimal answer is satisfied

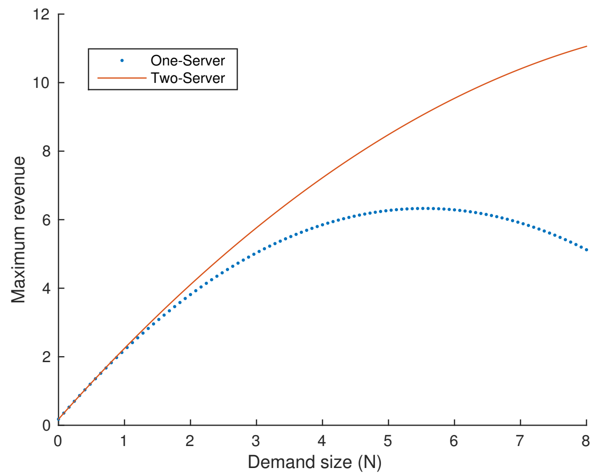

Figure 3.4: Maximum revenue of the service provider with regard to the size of demand $(\beta, \alpha = 0.2,\ P_1 = 2.25,\ P_2 = 2.1)$

by only one server but after a threshold the two server answer will be more beneficial, and finally due to quadratic characteristic of objective function the revenue starts decreasing.

Figure 3.4 illustrates the effect of demand size on the optimal revenue of a service provider. The solid line shows the optimal revenue of a service provider with 2 available servers ($p_1 = 2.25$ and $p_2 = 2.1$) while the dotted line depicts the revenue of another service provider with one server ($p_1 = 2.25$) for comparison. As it is depicted in the figure, the optimal answer passes 3 phases. At first, N is too small and the constraint 3.5 is active; therefore, the service provider opens only one server. When N is big enough to pass this constraint, the service provider opens both servers to keep the growth speed and counteract the negative effect of conformity. The final stage happens when the revenue reaches its highest and congestion become so intensive that the revenue starts decreasing.

Note that we never see a breakpoint such what we had in figure 3.3 while the demand grows. When conformity level grows, the feasible space starts to shrink, but the objective function does not change. However, when demand grows the optimal solution moves faster than the feasible space boundaries.

In this chapter, we elaborately describe the outline of the model, then we used an example to illustrate the restrictions and to investigate the effect of conformity level on the system. We found out the differentiation can be helpful when the level of conformity passes a certain threshold. In next chapter, we will talk about more general problems beside the stability of an answer and the degree of differentiation.

Chapter 4

The Generic Model

In the previous chapter, we introduced and analyzed a model with two servers in which the utility line is limited between 0 and 1; in addition, customers were distributed uniformly on the utility line. To generalize the model, we start by changing these assumptions one by one in each section. In the first section, we change the initial interval of the specification line and we try to study the cases when the customers are not distributed uniformly over the specification line.

In the next section, we offer a generalized model with a finite number of available servers. We reconsider some definitions based on this new setting and rewrite the mathematical model for this problem. Subsequently, we investigate the complexity of the problem and possible optimization methods to solve it. Finally, we apply this model in an empirical study in the area of online gaming.

4.1 Analysis of a Model With an Arbitrary Continuous Distribution

The main extension of the model we are interested in is related to the distribution of customers. In this section, we will assume $f(y)$ – as the customers' distribution function over the specification line– has two properties: first, if we take the integral of $f(y)$ over the specification interval limits, the result will be equal to demand size, and second, $f(y)$ is a continuous function.

We also extend the specification line from the [0,1] to the [0,M] interval. To carry through with this change, we need to adjust the intrinsic utility function to $\Gamma(y_j, x_i) = M^2 - (y_j - x_i)^2$; this change ensures a non-negative utility for all customers on the specification line. Although this function behaves very similarly to the previous one, it helps us adjust M whenever needed.

4.1.1 The indifference point and partition set relation

The function $f(y)$ is the distribution of customers over the specification line, and $F(y)$ is its cumulative function. For any continuous distribution, and when $x_1 \leq x_2$, we can rewrite Equation 3.9 as follows:

$$\Gamma(x_1, y_{ind}) + \beta F(y_{ind}) = \Gamma(x_2, y_{ind}) + \beta(N - F(y_{ind})) \tag{4.1}$$

If we solve this equation for y_{ind}, we can find the position of the indifference point based on the position of both servers.

We should be cautious about two main issues: first, losing the symmetry of solutions when we apply a general distribution, and second, losing the unity of the indifference point. In the previous cases, the mirror solutions of a position set based on the middle of the specification line would result in the same output, but now we lose this attribute as long as $f(y)$ is not symmetrical. Moreover, we will lose the uniqueness of the indifference point. Therefore, based on a fixed position set of the servers, we may have zero, one, two or more indifference points, which lead us to different outputs for one solution.

4.1.2 Stability of a partition set with 2 servers

In the previous chapter, we found that when customers are distributed uniformly on the specification line, there is only one indifference point while we offer two heterogeneous services. However, this does not mean that our answer is the only possible solution. When conformity is intense enough – either demand, conformity level, or both are high– all customers cascade toward one of the servers with a small arbitrary entropy. Therefore, in addition to the feasibility of solutions, we should also consider the stability of each solution.

Definition 6. A partition set is *stable* when all of the customers at each of its subsets exclusively prefer a unique server over the other options, and sustain this preference even if an arbitrary subset of customers change their preference.

Before we investigate the effect of stability on our solution, we need to find when the solution is stable. To this end, we first define a special case of stability which we call ϵ-stability, in which the solution is stable only when the changes are arbitrarily small and adjacent to the indifference point. Next, we expand this notion to the limited changes in the neighborhood of the indifference point. Finally, we expand these limits to secure a perfect stable solution.

Lemma 3. *For any indifference point y_{ind}, its corresponding partition set is ϵ-stable with regards to arbitrarily small changes in the neighborhood of y_{ind}, if and only if $f(y_{ind}) < \frac{2(|x_1 - x_2|)}{\beta}$, where $f(y)$ is the continuous distribution function of customers.*

Proof. In this problem we have a finite distribution of customers, which generates a finite and uncountable set of players. These players can choose their individual strategy from a set of finite and countable strategies, i.e., for a two-server problem they can choose either server 1 or server 2.

Given 2 subsets of customers: $C_\delta^+ = \{c_j | c_j \in \mathscr{C}, y_j \in [y_{ind}, y_{ind} + \Delta y]\}$ and $C_\delta^- = \{c_j | c_j \in \mathscr{C}, y_j \in [y_{ind}, y_{ind} - \Delta y]\}$, where $\mathscr{C}$ is the set of all customers, and c_j is a subset of customers who are at y_j, also the Δy is an arbitrarily small distance on the specification line.

Without loss of generality, assume that $x_1 \leq x_2$, and customers at C_δ^- and C_δ^+ respectively prefer server 1 and server 2.

By using Definition 6, a partition set is stable if a slight haphazard irregularity does not affect all customers' preferences. For example, if we force a small number of customers who initially prefer server 1 to join another server and then allow them to choose again, they should return to server 1 if the solution is stable.

Given $U(s_i, c_j)$ is the utility of clients at y_j for choosing s_i when the partition set is balanced [1], and $U'(s_i, c_j)$ is the utility function when customers at C_δ^+ choose the server that contradicts their preference. Based on these definitions, this balanced partition set is ϵ-stable if and only if both Equations 4.2a and 4.2b are satisfied.

$$U'(C_\delta^+, s_1) < U(C_\delta^+, s_2) \tag{4.2a}$$

$$U'(C_\delta^+, s_2) < U(C_\delta^+, s_1) \tag{4.2b}$$

If either one or both of these strict inequality conditions relax to non-strict inequalities, the associated strong Nash Equilibrium will change to a weak Nash Equilibrium.

Given $F(y)$ is the cumulative function of the customers' distribution, we define $\Delta F(y) = F(y + \Delta y) - F(y)$ and $\Delta F(y)^- = F(y + \Delta y) - F(y - \Delta y)$. Based on this definition we can calculate the size of any set of customers:

$$||C_\delta^+|| = \Delta F(y_{ind})^+ \tag{4.3a}$$

$$||C_\delta^-|| = \Delta F(y_{ind})^- \tag{4.3b}$$

Now, we apply these definitions in Equations 4.3a and 4.3b to solve the inequalities 4.2a and 4.2b. By expanding Equation 4.2a and adding the equality of balance point to both sides, we will reach the following inequalities:

$$\beta(\Delta F^+(y_{ind})) < 2\Delta y(x_2 - x_1) \tag{4.4a}$$

$$\beta(\Delta F^-(y_{ind})) < 2\Delta y(x_2 - x_1) \tag{4.4b}$$

[1] We refer to the balanced partition as a partition set that generated by an indifference point.

and then,

$$\frac{\Delta F^{+}(y_{ind})}{\Delta y} < \frac{2(x_2 - x_1)}{\beta} \tag{4.5a}$$

$$\frac{\Delta F^{-}(y_{ind})}{\Delta y} < \frac{2(x_2 - x_1)}{\beta} \tag{4.5b}$$

Owing to the distribution continuity assumption, we can take a limit from the LHS of both equations where $\Delta y \to 0$, and the results will be the following equations:

$$\frac{d_{+}F(y_{ind})}{dy} < \frac{2(x_2 - x_1)}{\beta} \tag{4.6a}$$

$$\frac{d_{-}F(y_{ind})}{dy} < \frac{2(x_2 - x_1)}{\beta} \tag{4.6b}$$

Because the distribution function is continuous, we can also conclude that the cumulative function is differentiable; therefore, the left-derivative and right-derivative will be equal. Moreover, we can relax the $x_1 < x_2$ constraint by using the absolute function, and then we can conclude that:

$$f(y_{ind}) = \frac{dF(y_{ind})}{dy} < \frac{2(|x_2 - x_1|)}{\beta} \tag{4.7}$$

□

This lemma simply captures the stability of a partition set with regard to changes in the neighboring indifference point, and we extend this lemma by using the Δ-stability notion.

Definition 7. A partition set is Δ-*stable if by changing the destination server of all customers between the indifference point* (y_{ind}) *and* $y_{ind} + \Delta$ *or* $y_{ind} - \Delta$*, the partition set* *remains stable or in other words:*

$$\forall \delta \in [-\Delta, +\Delta],\ |F(y_{ind} + \delta) - F(y_{ind})| < \frac{2\delta}{\beta}(|x_2 - x_1|) \tag{4.8}$$

Lemma 4. *When a partition set is Δ-stable, the stability of this partition set will not break due to any arbitrary changes in the customers' preference in the $y_{ind} \pm \Delta$interval.*

Proof. Given $\mathcal{C}$ is an arbitrary subset of C_Δ^+, where C_Δ^+ is the set of all customers between the indifference point (y_{ind}) and $y_{ind} + \Delta$. If this subset changes its server preference, it can make a cascade and reinforce the conformity utility of the other server such that all customers move to that server. However, the partition set is Δ-stable, and deviating C_Δ^+ will not change the partition set. We assume y_j is the most remote point of set $\mathcal{C}$ from the indifference point. If function $U(c_j, s_i)$ shows the utility function when we have the partition set based on the indifference point, and $U_C(c_j, s_i)$ shows the utility when set C is deviating from the expected choice, we can write:

$$U(c_j, s_1) \quad < \quad U_{\mathcal{C}}(y_j, s_2) \tag{4.9}$$

$$U_{\mathcal{C}}(c_j, s_2) \quad \leq \quad U_{C_j^+}(y_j, s_2) \tag{4.10}$$

So it can be concluded that:

$$U(c_j, s_1) < U_{C_j^+}(y_j, s_2) \tag{4.11}$$

which contradicts the definition of Δ-stability. □

Proposition 1. *A partition set is perfectly stable and the system has only a Nash equilibrium when partition set is* Δ*-stable, where* $\Delta = \max\{y_{ind}, M^2 - y_{ind}\}$.

This proposition suggests that if we move all customers to one server, they will still prefer to return to their original server and build the balanced partition set. In other words, a partition is perfectly stable when the system has only one Nash equilibrium.

4.2 Stability of a 2-server Model with Uniform distribution of customers

To create a better understanding of the stability concept, we investigate the effect of stability on the problem we introduced earlier in the previous chapter. For this purpose, we simplify the stability condition for the case in which customers are uniformly distributed over a specification line which is limited in a 0-1 interval. Based on these assumptions the distribution function of customers will be:

$$f(y) = N \quad \forall x \in [0, 1]$$

Therefore, we can rewrite Equation 4.7 for an uniform distribution between 0 and 1 as follows:

$$2(|x_1 - x_2|) > \beta N \tag{4.12}$$

The formulation above gives us a better picture of stability. The left-hand side of

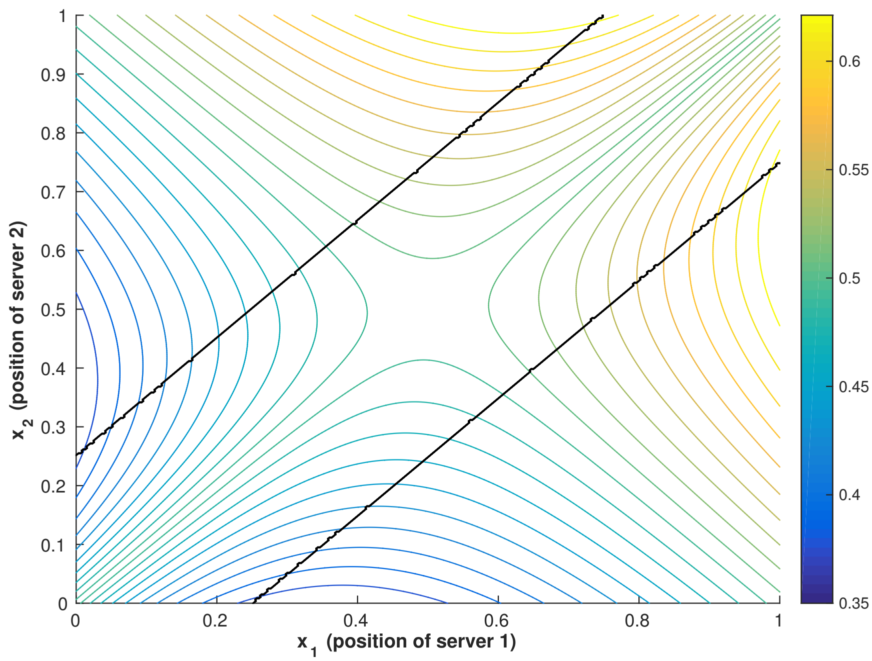

Figure 4.1: Revenue contour and stability constraint

equation demonstrates the differentiation needed to secure a stable solution, and right-hand side demonstrates the intensity of conformity.

Proposition 2. *When distribution of customers is uniform, ϵ-stability of a partition set results in the perfect stability.*

By using this proposition and Equation 4.12, we can illustrate the stability constraint by separating the unstable and stable area on the solution space. Figure 4.1 includes a set of contours which depict the isoprofit curves, and two black lines which capture the stability constraints.

The upper-left and bottom–right triangles respectively capture the stable areas where

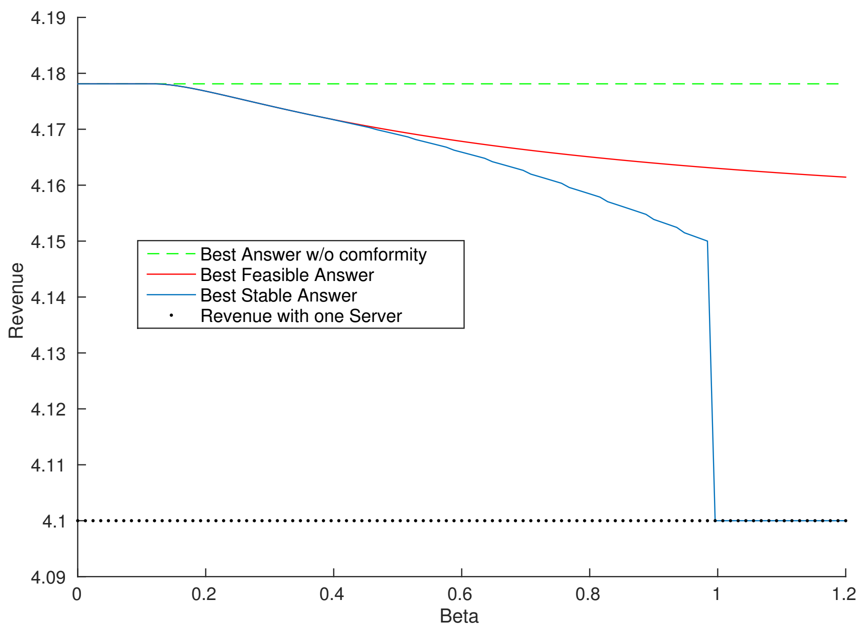

Figure 4.2: Revenue of service provider subjected to different constraints

$x_1 \geq x_2$ and $x_2 \leq x_1$. Any solution between these two triangles is unstable, which means that all the answers in this area have a tendency to break down. In that area, the intrinsic utility gap between two servers is small. As a result, conformity can overcome the idiosyncratic preferences of customers and create a snowball effect towards one of the servers to seek a stable answer. Clearly, when $\beta \times N$ is greater than 2, the service provider cannot arrange any stable 2-server partition set; however, when βN is small enough (e.g. in Figure 4.1), it does not invalidate any feasible solutions. Note that as β grows the stability boundary lines move toward the corners and the stable region starts shrinking.

Figure 4.2 compares the optimal revenue of the service provider in four different situations. Starting from the top, the first line shows the optimal solution derived without considering feasibility; the second line represents the optimal solutions that consider feasibility but not stability constraint; the third line captures the best stable solution; and finally, the last line represents the optimal solution with only one server. The first break point (close to 0.2) occurs when the best answer drops out of feasible region when the feasible region shrinks by an increase of β. The second break point (close to 0.4) happens when the best feasible answer falls into the unstable region described by Equation 4.12. The last break point (at 1.0) occurs when the feasible stable solution space is empty. If we set different values for these parameters we might not see these 4 phases clearly, but the general trend will remain the same.

4.2.1 Mathematical Model

In the previous chapter, we mentioned even for a 2-server service provider, the mathematical optimization model of the problem is nonlinear. In this section, we formulate the generic version of the optimization model for a 2-server model with uniformly distributed set of customers. We start with modelling our original problem:

$$\max_{X_s} \Pi(X_s) \quad s.t. \tag{4.13a}$$

$$0 \leq X_s \leq 1 \tag{4.13b}$$

$$2(|x_2 - x_1|) \geq \beta N \tag{4.13c}$$

The first equation indicates our objective function, the second one takes care of differentiation constraints, and the last one secures a stable solution. This model may not be sufficiently self-explanatory. Hence, we break it down into more details, while limiting the feasible space to half. For this purpose, we again assume that $x_1 \leq x_2$, and y_{ind} is the indifference point driven from Equation 3.15, or its simplified version based on our assumption: $y_{ind} = \frac{x_2^2 - x_1^2 + \beta}{2(x_2 - x_1 + \beta)}$. Based on these assumptions we rewrite the model as:

$$\max_{X_s} \Pi(X_s) = N y_{ind}(p_1 - \alpha N y_{ind}) + N(1 - y_{ind})(p_2 - \alpha N(1 - y_{ind})) \qquad s.t. \tag{4.14a}$$

$$0 \leq x_s \leq 1 \qquad \forall s \in \mathcal{S} \tag{4.14b}$$

$$x_1 \leq x_2 \tag{4.14c}$$

$$2(x_2 - x_1) \geq \beta N \tag{4.14d}$$

To solve this model, we can either apply a computational model to find an optimal solution or apply the Karush–Kuhn–Tucker conditions to build a set of equations and solve them by a commercial solver. First, we analyze the computational model.

To obtain the optimal solution without considering the constraints, we need to produce the optimal partitioning which is driven from Equations 3.4a, and 3.4b, then find the corresponding hyperbola of server's position based on this partition set by using Equation 3.9. After passing these steps, we reach the following hyperbola:

$$x_1^2 - x_2^2 + (\frac{4\alpha + p_2 - p_1}{4\alpha})(x_2 - x_1 - \beta) = 0 \tag{4.15}$$

If we find a set of services, $X = \{x_1, x_2\}$, which satisfies the aforementioned hyperbola constraint, as well as the feasibility and stability constraints, that set will be the optimal solution of the model. If there were no set of services between 0 and 1 which satisfies

Algorithm 4.1 Computational model of 2-server optimization

Form the equation 4.15
Form the optimal solution set $\mathcal{O}$
Form the Feasible Region $\mathcal{F}$
$\mathcal{O} = \mathcal{O} \cap \mathcal{F}$
If $\mathcal{O} = \emptyset${
set $x_1 = 0$, find the optimal solution set $\mathcal{O}'$
set $x_2 = 0$, find the optimal solution set $\mathcal{O}$"
$\mathcal{O} = \arg\max(\Pi(\mathcal{O}')\&\Pi(\mathcal{O}"))$}
Form the Stable Region $\mathcal{S}$
$\mathcal{O} = \mathcal{O} \cap \mathcal{S}$
If $\mathcal{O} = \emptyset${
Set $|x_1 - x_2| = \beta N/2$, find the optimal solution set $\mathcal{O}$}
Return $\mathcal{O}$

these conditions, we narrow down our search space to the feasibility boundaries where either $x_1 = 0$ or $x_2 = 1$. To this end, first, we assume $x_1 = 0$, with this assumption the initial constraint 4.14c will be redundant. However, two other constraints remain valid. Inequalities $x_2 \leq 1$ and $2x_2 \geq \beta N$ are the transformed version of those constraints. If the solution violates the constraint 4.14d, we repeat the same procedure, which leads us to another search space on the line where $2(x_2 - x_1) = \beta N$. With this procedure, we can find the optimal position set. Consider that, when either of our constraints is activated, our optimal solution changes from a hyperbola to a single node where the optimal hyperbola and one of the constraints are tangent.

On the other hand, by using the KKT conditions, we can offer a more generalized procedure to solve this problem. To this end, we first need to write the Lagrangian form the model:

$$\mathcal{L}(X) = \Pi(X) + \mu^T g(X) \tag{4.16}$$

Where $\mathcal{L}$ is the Lagrangian function of our optimization model, μ is the set of inequality Lagrange multipliers, and $g(X)$ is the set of inequality constraints, i.e., inequalities 4.14b, 4.14c, and 4.14d. For this model, we will have 6 inequality constraints 4 deriven from Equation 4.14b, 2 from the others, and all of them must be reformed to $g(X) \leq 0$. With these assumptions, the KKT condition of the model will be:

$$\nabla_X \Pi(X_s) + \mu^T \nabla_X g(X) = 0 \tag{4.17a}$$

$$\mu^T g(X) = 0 \tag{4.17b}$$

$$\mu \geq 0 \tag{4.17c}$$

$$g(X) \leq 0 \tag{4.17d}$$

The KKT produces 8 equality equations, 8 inequality equations, and 8 variables. This methodology is more general compared to our previous procedure but solving a nonlinear set of equations is not always easy; nonetheless, there are many numerical methods to help.

4.3 Analysis of a Multi-server Model

All the discussed cases before assumed that the service provider can open at most 2 servers. In this section, we relax this assumption and let the service provider open more than 2 servers simultaneously. However, we keep the uniform distribution assumption to keep the problem away from the unnecessary complexity of other distributions. We also borrow several of 2-server model concepts to build the current one. The following equations illustrate the general format of the optimization model:

$$\max_{X_s} \Pi(X_s) \qquad s.t. \tag{4.18a}$$

$$0 \leq X_s \leq 1 \tag{4.18b}$$

$$2(|x_i - x_j|) \geq \beta N \qquad \forall s_i, s_j \in \mathcal{S} \tag{4.18c}$$

Note that by the growth of available servers, the solving time of this model with previous procedure exponentially increases. Moreover, defining the objective function is not as easy as the 2-server model. To solve this problem, we first need to elaborate it's optimization model.

We can formulate the mathematical optimization model of multi-server problem with two different approaches. First, we can extend the 2-server model which in the set of servers' position, X, is the only decision variable, or we can add the partition size as an auxiliary decision variable to make the problem more clear. The former case generates a very complex and highly non-linear objective function, and the latter method may include more decision variables, but it generates a more affable optimization model with a quadratic objective function. In this chapter we only focus on the second methodology.

Proposition 3. *There will be no stable solution for a service provider with m active server if $m \geq M\frac{2}{\beta N} + 1$.*

Proof. If the number of active servers is greater than $\frac{2M}{\beta N} + 1$, there will be two adjunct server with distance of less than $\beta n/2$, and based on Equation 4.12 these servers cannot create a stable solution. $\square$

From now on, we assume that $m \leq \frac{2M}{\beta N} + 1$ is always valid. If a service provider has more possible servers, we only keep the top m with the highest net income coefficient. This assumption keeps the complexity of our problem limited.

We base our model on the 2-server optimization model where the revenue was function of servers' demand. To this end, we define the revenue of service provider equal to aggregation of all servers' revenue. If we calculate the revenue based on demand sizes, and without considering the feasibility of differentiation and stability of solution, the initial model will be:

$$\max_{\mathcal{N}} \Pi = \sum_{i \in \mathcal{S}} \pi_i(n_i) \tag{4.19}$$

$$\sum_{i \in \mathcal{S}} n_i = N \tag{4.20}$$

$$n_i \geq 0 \quad \forall i \in \mathcal{S} \tag{4.21}$$

Where $\mathcal{N}$ is the set of demands for all servers, and $\mathcal{S}$ is the set of all servers. Solving this model gives us an optimal solution, but we cannot be sure if this solution can be implemented regards to our differentiation constraints, but we can be sure that the output of this model is an upper-bound of our problem. Also, by solving this model we are able to omit the servers with zero optimal demand, and reduce the complexity of further calculations by reducing the number servers.

Solving this model generates a solution in which only the demands of all servers are specified, but to check the feasibility and the stability of solution, all possible positions of servers that result in this demand partition should be calculated.

To avoid repeating the equations, we do not rewrite the aforementioned objective function and constraints, and in each step we add the necessary constraints to complete the model. In the first step, we add the feasibility constraint related to the position of servers –i.e. Equation 4.22, and calculate all possible indifference points in Equation 4.23, and at the end we add the stability constraint –i.e. Equation 4.24.

$$X \in [0, M] \tag{4.22}$$

$$U(y_{i,j}, x_i) = U(y_{i,j}, x_j) \quad \forall i, j \in S - \{i, j | i = j\} \tag{4.23}$$

$$2(|x_i - x_j|) \geq \beta N \quad \forall i, j \in \mathcal{S} \tag{4.24}$$

In the next step, we need to relate the optimal demand sizes to the *active indifference points* in which the active indifference points are defined as the indifference points which in the dominant preference of customers is different for each side of the indifference point. For every pair of services, we will have an indifference point where only $m - 1$ of them are active and shaping the partition set. To find out the active indifference points, we use a binary variable $r_{i,j}$ in the following constraint which is an optimization problem:

$$\max \quad \sum_{i,j \in S} r_{i,j} \quad s.t. \tag{4.25}$$

$$U(x_i, y_{i,j}) \geq U(x_k, y_{i,j}) \times r_{i,j} \quad \forall i, j, k \in S \tag{4.26}$$

$$\frac{\partial U(x_i, y_{i,j})}{\partial y_{i,j}} r_{i,j} \leq \frac{\partial U(x_j, y_{j,i})}{\partial y_{j,i}} r_{j,i} \quad \forall i, j \in S \tag{4.27}$$

Based on this model, we can find the indifference points. Constraint 4.26 specifies the intersections that the preference of customers change, if $y_{i,j}$ is active both $r_{i,j}$ and $r_{j,i}$ can be one. The next constraint shows the direction of change, i.e., when the preference changes from server i to server j, this constraint keeps $r_{i,j}$ as 1 and forces $r_{j,i}$ to zero. Now that we found our active indifference point, we need to map them to the demand.

$$n_i = (\sum_{k \in S} r_{k,i} y_{k,i} - \sum_{k \in S} r_{i,k} y_{i,k}) \times N/M \tag{4.28}$$

Figure 4.3 shows us an example of this map, but we need to check all possibilities of servers arrangement.

Up to now, we built an optimization problem constrained by the other optimization problem. By solving this nonlinear model we can find the optimal solution; however, owing to the nonlinearity structure of problem, we could not offer an approach to find the exact solution of this problem in the polynomial time.

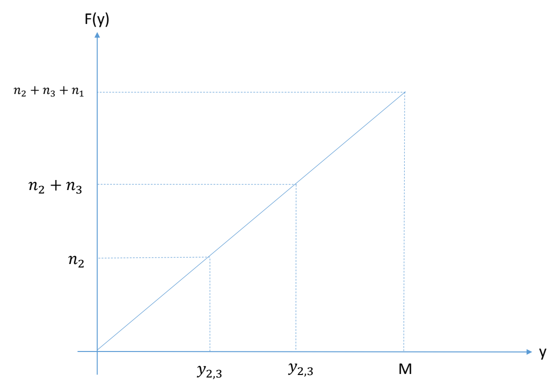

Figure 4.3: Relationship between indifference points and partition set with 3 server

4.4 Empirical Study

We investigate an online game service who can offer different range of games on the limited number of servers. We can differentiate the games based on the complexity of them; however, adding an extra service with a different setting causes a surplus cost to the service provider. This surplus cost is fixed and determined. The firm offers more than one game type with different setting. We assume that for each game type, demands, revenues and costs are independent; thus, we can independently optimize them.

Each game on a server has a specific net income in a predefined time window, which is usually two hours. This net income depends on the server, the number of people playing them and the congestion penalty of the game. The player is charged for each round of the game, and the length of each round has a positive relationship with the number of people playing the game. Therefore, as the number of people grows the number of rounds per time-unite decreases, as with the net revenue of server per player. On the other hand, the players find a game more "fun" if there are more crowd of players in the game.

Because of the gap between the optimal number of players and capacity of servers, the capacity strategy does not help the system optimally. Also, changing the capacity cannot improve the revenue dramatically; and moreover, altering the capacity frequently is not an option for the firm. Here, we offer the differentiation solution to counteract this problem and maximize their revenue. By using a time series analysis, the firm forecasts the demand for each game based on data of the previous weeks and data of a same period of a year ago. In this game, players team together to play against an AI (Artificial Intelligence), and the firm offers 3 level of AI's difficulty for each game type. With some approximation, we assume that the demand for these levels is uniformly distributed, and we also extend the 3 levels of difficulty to a continuous spectrum, which maps the limited difficulty of the level 1 to 0 and level 3 to 1. For each game, they open with up to 3 different setting, and by opening a new server with the new server the initial income of new server is 15 percent less than the previous one. The only missing elements of the model that needs more investigation to find is the conformity level and the differentiation size, by assuming a fixed value for the differentiation size, we only need to find the conformity level. Fortunately, on the peak hour as all servers serve almost the same amount of player, we can profile the preference of the customer, and later on investigate the effect of conformity on their decision. Based on this data we can approximate the value of beta for customers, this value varies from customer to customer sometimes, but we use the median of it. However, we calculated all parameters for all game types individually. Therefore, we can see the effect of different loads on the final result.

The current policy of the firm is based on offering 3 services, at 3 stages of difficulty 24/7, and if demand of a game exceeds a certain threshold they start opening a new server with the same service specification. We compared the result of their strategy, with the optimal solution we calculate based on the differentiation model. The detailed result of these comparisons will be discussed in the next chapter.

Chapter 5

Results

This chapter is twofold, the first part contains the result of a two-server model which is implemented with the sample data. The second part is the result that is generated from an online game service provider data, to show the applicability of model for a more complex problem. For the 2-server model, we tried to make examples as general as possible to illustrate the different phases of problem, and carry out the concept clearly.

5.1 Differentiation Strategies for 2-server model

We tie the optimality of service differentiation to the three important model attributes: the conformity incentive, the total demand, and the cost of congestion. As a benchmark, we compare the results with the revenue resulting from placing undifferentiated servers. Finally, we also assess the relationship between these elements and the optimal degree of differentiation which may not be implementable because of the conformity incentive.

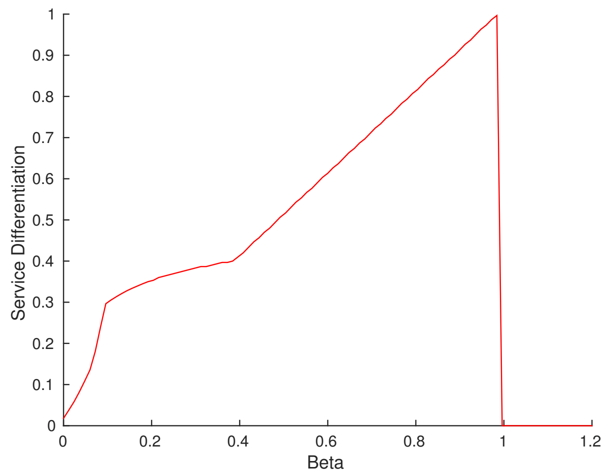

Figure 5.1: Optimal Differentiation with respect to the conformity level,β, ($\alpha = 0.12, N = 2$)

Based on the 2-server service provider with uniform distribution, we show how differentiation studies can counteract the possible congestion in the system caused by conformity externalities. If we get back to the Figure 4.2, we can see four different phases, where in each of them revenue exposes a different response to changes of conformity level. By looking at Figure 5.1, we can see the exact same phases in this graph. This graph demonstrates the optimal level of differentiation for a sample example with respect to the level of conformity. At the first phase, as the conformity grows we can mostly counteract the negative effects of congestion. In the next two phases we face the different constraints that reduce the possible differentiation strategies, and in the last phase we see the case that conformity level is so intense in which differentiation cannot help the system.

The other attribute that can increase the rate of conformity is the demand size. Figure 5.2 shows the level of differentiation for the optimal solution, this graph passes 5 different

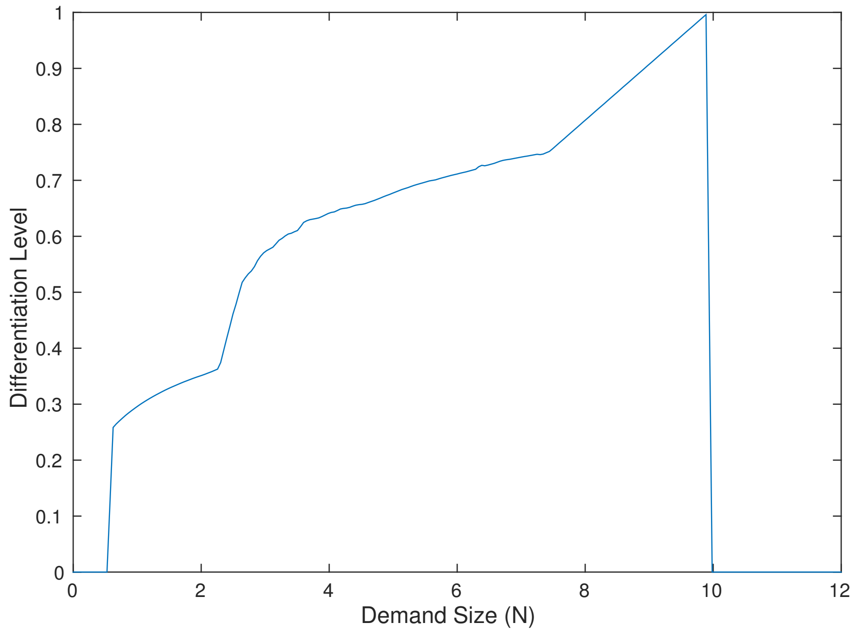

Figure 5.2: Optimal differentiation level with respect to the demand size $(N)(\alpha = 0.09, \beta = 0.2)$

stages. In the first stage, N is small enough that the optimal solution will be gained by only one server, after some threshold the optimal solution moves to one of the boundaries (either $x_1 = \{0, 1\}$ or $x_2 = \{0, 1\}$). In the third stage, where $N = [2, 8]$ here, the optimal solution is not different from the model without any constraint. As the solution can be a hyperbola, there is no specific solution; however, we kept the maximum level, the minimum level will be the mirror of curve respect to the line $x = y$. At the forth stage, the optimal solution will be tangent to the stability constraint, and at the end the conformity is so intense that all customers cascade to one server.

On the other hand, if we compare the optimal revenue of differentiated services with

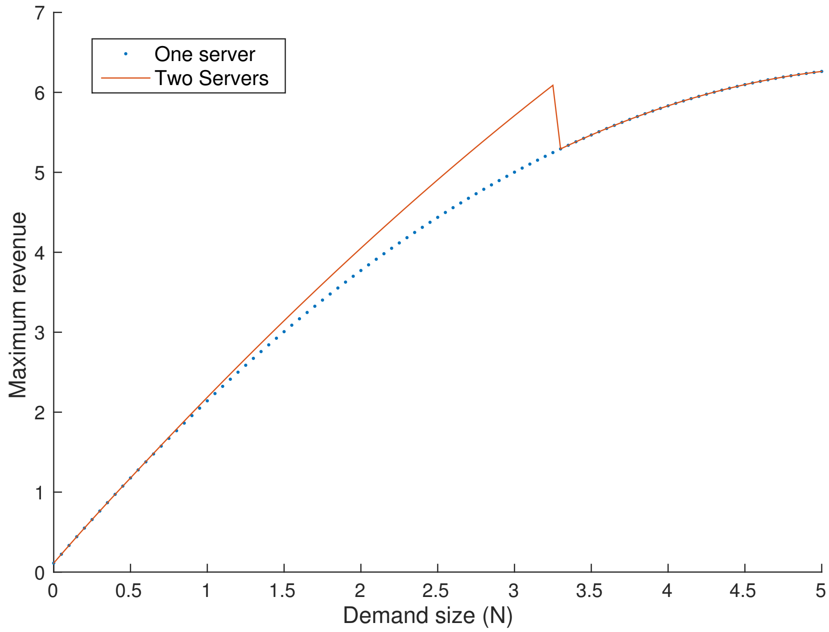

Figure 5.3: Maximum revenue of the service provider with regard to the size of demand ($\alpha = 0.08, \beta = 0.4$)

the homogeneous services, we observe that the differentiation can only help the service provider to enhance its revenue in a limited interval. If the demand is low, the congestion is low enough that one server can handle all the demand, and when demand is high enough, the limits on differentiation lead the system to choose one server again. In the example that Figure 5.3 illustrated the improvement is between 0-13% in total.

The last attribute of model we study is the congestion penalty coefficient. The figure 5.4 clearly shows three phases of differentiation regarding the value of α. Up to some threshold, there is no active differentiation policy. This situation happen when either α is less than $\frac{|p_1 - p_2|}{2N}$ or the 2-server optimal solution cannot be conducted in the feasible area.

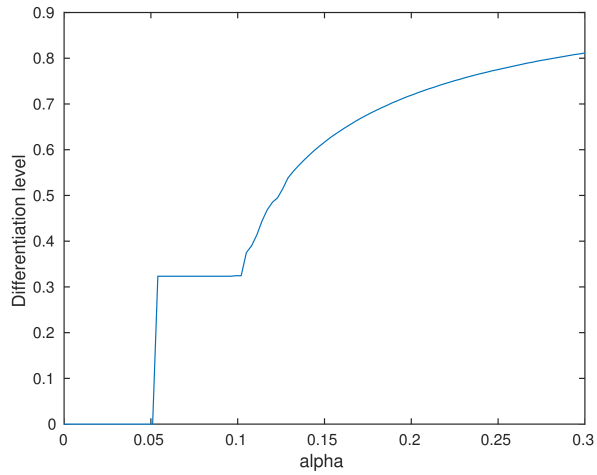

Figure 5.4: Optimal differentiation level with respect to the congestion penalty intensity $(\alpha)[N = 2, \beta = 0.2]$

5.2 Differentiation Strategies for multi-server model

In this section we compare the result of the firm's current policy, *virtual capacity* policy, with the result of our current model for 168 sample time units. This data is driven from a 2-week time period, and contains a full range of demands. We have the same amount of data and exactly the same procedure for all 5 game types. In this analysis, two patterns has been captured for all game types, which we call them *over-sized* and *under-sized* situation. In four games, we observed the *over-sized* situation. In these cases, firm stays too long to open new servers which cause the congestion significantly affects the revenue. This phenomena happens when the time of games highly depends on the number of player, which causes the higher congestion penalty.

In the figure 5.5, the continuous line shows the revenue of current policy for an over-sized situation. As it can be seen in the graph, there is a jump right after demand passes

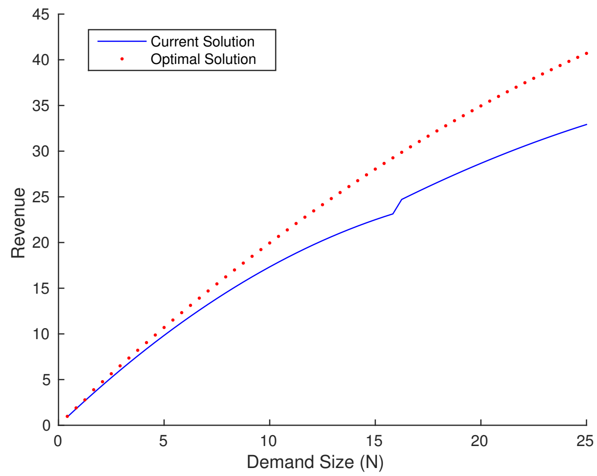

Figure 5.5: Difference of current policy with the optimal policy for the game type #1

15. This jump was caused by the new opening, and this alerts us we need to open the new server earlier.

On the other side, one game showed the opposite result; in the moment the new server is become available, there is a small plunge in the revenue. This behavior shows us that we are opening new servers sooner than we reach the full utilization capacity of current servers. Figure 5.6 demonstrates this plunge around the 45 units of demand. In this case, the congestion penalty is estimated as 0.015, while in the previous case the coefficient was 0.027. This coefficient solely depends on the structure of the game, and how congestion affects the average time of the game. The detailed result of all games is documented in the Appendix A.

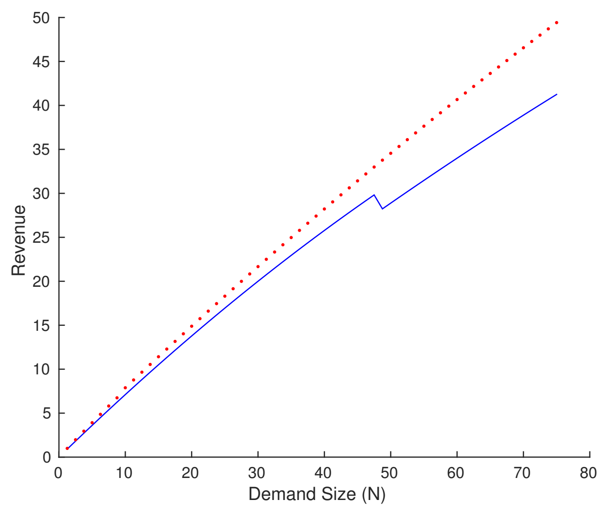

Figure 5.6: Difference of current policy with the optimal policy for the game type #4

Chapter 6

Conclusions

6.1 Conclusions and Summary

This study offered a model in which the service provider struggles with an unwanted congestion caused by externalities among the customer, and offered the differentiation as a strategy to hedge this congestion in particular situations.

We found that the optimal server differentiation is driven by two effects:

1. Parameters that reinforce crowding require additional differentiation to optimally partition customers.

2. Parameters that reinforce crowding shrink the space of stable customer partitions requiring the provider to settle for *second best* solutions.

For instance, if the crowding parameter is increased, relative utility for the more crowded server will increase, leading some customers to switch servers. The service provider may be able to maintain the current partitioning by increasing differentiation but at some point, the level of differentiation will be maximized and this particular partitioning will no longer be stable for any pair of server types. When the crowding parameter is large enough a single server with maximum efficiency will provide higher stable revenue than any pair of servers. This results in the following transitions:

1. At low incentives to crowd, the first best partitioning of customers can be implemented with minimal differentiation.
2. At moderate incentives to crowd, only a second best solution can be implemented and requires high levels of differentiation.
3. At high incentives to crowd, it is most profitable to operate only the most efficient server.

In the muti-server model, we compared the solution of limited capacity strategy with optimal differentiation strategy. For the online game case study, we found that when level of conformity is not very high and differentiation can be done without any unwanted cascade, the differentiation can be very beneficial for the firm. Although, when the demand gets very high –or in the case that conformity level is high, for keeping the solution stable we need to decrease the number of servers, and in this case we may reach the point that the capacity can offer a better stable solution.

6.2 Future Research and Recommendations

We introduced a novel motive for a monopolist to differentiate services which is of interest to a variety of service industries where customers benefit from receiving the service with their cohort. We tried to take a comprehensive look at this problem, but as we moved further new questions arose. We believe the problem can be expanded in two different degrees.

First, we are interested to offer a more complex model with more case studies. For instance:

a. Investigating the behavior of model with different distribution of customers.

b. Investigating the case when the revenue of each server is not only affected by the server, but also by the service it offers.

Second, as either the number of customers increases or the distribution of customers becomes more complicated, we need better algorithms to find the optimal level of differentiation. In this thesis, we used different approaches to solve the problems in a reasonable time, but the existence of a more general algorithm can be beneficial for the future works.

APPENDIX

Appendix A

Empirical Study Parameters

Table A.1 shows the estimated attributes of all 5 games we investigated in the empirical study section. All games can be roughly divided into two sections. Games 1,2, and 3 have lower demand, lower capacity and higher initial marginal net income. In the text we referred to the games 1 and 4 as a representative of each group.

	Game #1	Game #2	Game #3	Game #4	Game #5
Maximum Demand	25	25	25	80	75
Congestion Penalty(α)	0.015	0.017	0.03	0.027	0.012
Conformity Level (β)	0.03	0.03	0.03	0.02	0.02
Capacity	5	5	5	15	15
Initial net Income Coefficient(p_1)	2.6	3.1	2.8	2.4	2.4

Table A.1: Game Settings for all Game Types

References

[1] Simon P Anderson, Andre De Palma, and Jacques François Thisse. *Discrete choice theory of product differentiation*. MIT press, 1992.

[2] Solomon E Asch. Effects of group pressure upon the modification and distortion of judgments. *Groups, leadership, and men. S*, pages 222–236, 1951.

[3] Michael Bacharach. *Economics and the Theory of Games*. Macmillan, 1976.

[4] Gary S Becker. A note on restaurant pricing and other examples of social influences on price. *Journal of Political Economy*, pages 1109–1116, 1991.

[5] Moshe E Ben-Akiva and Steven R Lerman. *Discrete choice analysis: theory and application to travel demand*, volume 9. MIT press, 1985.

[6] B Douglas Bernheim. A theory of conformity. *Journal of political Economy*, pages 841–877, 1994.

[7] Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

[8] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, pages 992–1026, 1992.

[9] John R Birge and Francois Louveaux. *Introduction to stochastic programming.* Springer, 2011.

[10] James M Buchanan and Wm Craig Stubblebine. Externality. *Economica*, pages 371–384, 1962.

[11] Blanche Capel. The battle of the sexes. *Mechanisms of development*, 92(1):89–103, 2000.

[12] Edward Hastings Chamberlin. *The theory of monopolistic competition: A re-orientation of the theory of value*, volume 38. Harvard University Press Cambridge, MA, 1962.

[13] Richard Cornes. *The theory of externalities, public goods, and club goods.* Cambridge University Press, 1996.

[14] Peter J Danaher. Customer heterogeneity in service management. *Journal of Service Research*, 1(2):129–139, 1998.

[15] Nicholas Economides. Network externalities, complementarities, and invitations to enter. *European Journal of Political Economy*, 12(2):211–233, 1996.

[16] Guillermo Gallego and Garrett Van Ryzin. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research*, 45(1):24–41, 1997.

[17] Isabel Grilo, Oz Shy, and Jacques-François Thisse. Price competition when consumer behavior is characterized by conformity or vanity. *Journal of Public Economics*, 80(3):385–408, 2001.

[18] Michael L Katz and Carl Shapiro. Network externalities, competition, and compatibility. *The American economic review*, pages 424–440, 1985.

[19] Harvey Leibenstein. Bandwagon, snob, and veblen effects in the theory of consumers' demand. *The Quarterly Journal of Economics*, pages 183–207, 1950.

[20] Jeffrey I McGill and Garrett J Van Ryzin. Revenue management: Research overview and prospects. *Transportation science*, 33(2):233–256, 1999.

[21] Georgia Perakis and Guillaume Roels. The price of anarchy in supply chains: Quantifying the efficiency of price-only contracts. *Management Science*, 53(8):1249–1268, 2007.

[22] Matthew Rabin. Incorporating fairness into game theory and economics. *The American economic review*, pages 1281–1302, 1993.

[23] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

[24] Hiroo Sasaki and Manabu Toda. Two-sided matching problems with externalities. *Journal of Economic Theory*, 70(1):93–108, 1996.

[25] William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk*. *The journal of finance*, 19(3):425–442, 1964.

[26] Arch Wilkinson Shaw. *Some problems in market distribution*. Harvard University Press, 1915.

[27] Oz Shy. *The economics of network industries*. Cambridge University Press, 2001.

[28] Wendell R Smith. Product differentiation and market segmentation as alternative marketing strategies. *The Journal of Marketing*, pages 3–8, 1956.

[29] Kalyan Talluri and Garrett Van Ryzin. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):15–33, 2004.

[30] Kalyan T Talluri and Garrett J Van Ryzin. *The theory and practice of revenue management*, volume 68. Springer Science & Business Media, 2006.

[31] Thorstein Veblen. *The theory of the leisure class*. Oxford University Press, 2007.

[32] Dan Zhang and William L Cooper. Revenue management for parallel flights with customer-choice behavior. *Operations Research*, 53(3):415–431, 2005.