Pricing Bundles of Products and Services in the High-Tech Industry
by
Juan-Carlos O. Ferrer
Civil Industrial Engineer, P. Catholic University of Chile (1995)
Master of Science in Engineering, P. Catholic University of Chile (1995)
Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Management
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2002
© Juan-Carlos O. Ferrer, MMII. All rights reserved.
The author hereby grants to Massachusetts Institute of Technology
permission to reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author .......................................................... Sloan School of Management
May 2002
Certified by ........................................................................ Gabriel R. Bitran
Deputy Dean; Nippon Telephone and Telegraph Professor of Management Science
Certified by ........................................................................ Erik Brynjolfsson
Professor of Management
Certified by ........................................................................ Yashan Wang
Robert N. Noyce Career Development Assistant Professor
Accepted by .......................................................................... Birger Wernerfelt
Professor of Management Science; Chair of the Doctoral Program
Pricing Bundles of Products and Services in the High-Tech Industry

by

Juan-Carlos O. Ferrer

Submitted to the Sloan School of Management on May 2002, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Management

Abstract

The High-Tech industry faces tremendous complexity in product design because of the large number of different products that can be offered and the mix of products and services that exists. Information Technology (IT) products and services available in the market are increasing exponentially. Bundling appears in this industry as a natural mechanism to reduce complexity for sellers and buyers and to reduce variability on the customer’s valuation of individual products. The first chapter of this dissertation discusses these issues.

Chapter Two addresses the real-world problem of pricing bundles of IT services and products contracts when there is a high setup cost. Customers pay a fixed monthly fee. The company finances the hardware (HW) and software (SW) while the services and support are paid in a monthly basis out of the fee. The solution approach computes the monthly fee to be charged for every offered bundle, taking into account that customers may defect before the end of the contract. The dynamics of the system account for defection of current customers and arrival of new ones, at each period. Optimal pricing policies and equilibrium points of the system are characterized.

Chapter Three addresses the determination of the optimal bundle’s composition and price while maximizing total expected profits. The setting is a high-tech company in a highly competitive environment that must build a bundle and put it out in the market. Bundles are built from a set of components that meet technical constraints. The customers’ choice among competitors’ bundles (not under the company’s control) and the company’s bundle (under its control) is modeled in a random utility framework. A nonlinear mixed integer programming formulation of the company’s decision problem is presented and solved.

Chapter Four analyzes telecommunications networks for broadband services. The network consists of several regions whose connecting links have a given bandwidth capacity at each point in time. Bandwidth is an intangible perishable commodity that has to be sold in advance. The network owner faces the problem of selling the available bandwidth capacity of the delivery period from the present to the last selling period. In the analysis, the network owner faces an external demand for each market or product (capacity between two cities) that can be fulfilled by assigning capacity of the corresponding arc (direct
link) or by assigning capacity of an alternative path (indirect link), incurring a higher cost. Customers do not distinguish between routings as long as the requested connection between the pair of cities fulfills the specified Quality of Service. The goal is to maximize the network owner’s revenues while deciding on each period whether or not to bundle links.

Research Head: Gabriel R. Bitran
Title: Deputy Dean; Nippon Telephone and Telegraph Professor of Management Science

Thesis Supervisor: Erik Brynjolfsson
Title: Professor of Management

Thesis Supervisor: Yashan Wang
Title: Robert N. Noyce Career Development Assistant Professor
Contents

1 Introduction to Bundle Pricing .................................................. 12
   1.1 Real world setting .......................................................... 13
   1.2 Background and literature review ...................................... 17
   1.3 Bundle pricing problems ..................................................... 20

2 Pricing Service-Bundle Contracts with High Setup Cost ................. 26
   2.1 Introduction ....................................................................... 26
   2.2 Bundle monthly fee ......................................................... 31
      2.2.1 Example: Extended Office .......................................... 31
      2.2.2 Financing costs ......................................................... 32
      2.2.3 Customer’s lifetime distribution ................................... 32
      2.2.4 The monthly fee ........................................................ 34
   2.3 Expected profit per customer .............................................. 34
      2.3.1 Single-period profit model .......................................... 35
      2.3.2 Multi-period profit model .......................................... 38
      2.3.3 Alternative way to model the per-customer expected profit . 39
      2.3.4 Geometric customer behavior ...................................... 41
   2.4 A one-step choice model approach ...................................... 45
      2.4.1 Consumer choice model ............................................. 46
      2.4.2 The model ................................................................ 51
      2.4.3 Analysis of the optimal expected profit ......................... 52
   2.5 A two-step choice model approach ...................................... 54
      2.5.1 Assimilation-contrast theory ...................................... 55
      2.5.2 The model ................................................................ 55
      2.5.3 Analysis of the optimal expected profit ......................... 57
   2.6 A model for a dynamic system ............................................ 62
      2.6.1 Single bundle model formulation .................................. 63
      2.6.2 Optimal pricing for the last period ............................... 65
      2.6.3 Bellman equation ......................................................... 66
      2.6.4 Certainty equivalent controller (CEC) ............................ 67
      2.6.5 CEC transient phase .................................................... 72
A Bundle Pricing Overview 140
A.1 Motivation .................................................. 141
A.2 Notion of Price Bundling ........................................ 142
A.3 Bundling Strategies ............................................. 143
A.4 Why Bundling? .................................................. 145
A.5 Economic Research .............................................. 146
A.6 Consumer Behavior Research ................................. 147
A.7 Decision Models Research ................................. 149

B Mathematical Appendix 154
B.1 Gumbel Distribution Properties ............................... 154
B.2 Lambert W Function ............................................ 155
   B.2.1 Properties and approximations ...................... 155
   B.2.2 C code for Lambert W function, principal branch .... 157

C Computational Experiments Code 160
C.1 Bundle bounds source code .................................. 160
C.2 Pricing policy source code .................................. 162
# List of Figures

1-1 Customized service offering ........................................... 13
1-2 Bundling opportunity research ........................................ 21
1-3 Single customer segment .............................................. 22
1-4 Multiple customer segment .......................................... 24

2-1 Two customer behavior models ....................................... 27
2-2 Choice probabilities in the one-step choice model (case $B = 3$) .... 28
2-3 Choice probabilities in the two-step model (case $B = 3$) ............ 29
2-4 Monthly payment based on contract length .......................... 33
2-5 Customers’ lifetime distribution ...................................... 33
2-6 Shift on customer lifetime pdf when changing $p$ .................... 37
2-7 Customer behavior in the one-step choice model .................... 46
2-8 Optimal prices determined by the first order conditions ............. 54
2-9 Path-following procedure in a 2-bundle case ........................ 54
2-10 Assimilation-Contrast Theory ....................................... 55
2-11 Choice probability vs. observed price .............................. 56
2-12 Customer behavior in a 2-step decision model ...................... 57
2-13 Example of individual profit functions for 3 products ............... 59
2-14 Total profit upper bound for the 3-product example ................. 60
2-15 Optimal prices vs. Lower Bound prices ............................ 62
2-16 Changes in the discount factor $\rho$ ................................ 64
2-17 Graphical solution for price $p^*_T(N_T)$ in equation (2.45) .......... 65
2-18 System convergence to steady state ................................ 67
2-19 $H(p)$ changes with the setup cost $K$ ............................. 71
2-20 Quasi-concavity of functions $g(N, p, w)$ ........................... 73
2-21 Example of constraint (2.52) for $T = 2$. .......................... 73
2-22 Optimal prices when the system reaches steady-state in $T$ periods. 74
2-23 Interest rate for which a fixed-price policy is optimal ............... 75
2-24 Optimal pricing policy and Value function .......................... 76
2-25 Simulations using optimal pricing policy when $r = 0.1$ .............. 77
2-26 Optimal pricing policy around $\hat{\rho} = 1/(1 + \hat{r})$ ............... 77
2-27 GAP between optimal pricing policy and fixed-price policy .......... 78
<table>
<thead>
<tr>
<th>Page</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-28</td>
<td>M/M/∞ queue with state-dependent rates</td>
</tr>
<tr>
<td>2-29</td>
<td>An example of an optimal pricing policy</td>
</tr>
<tr>
<td>2-30</td>
<td>Example of a two-price optimal policy</td>
</tr>
<tr>
<td>2-31</td>
<td>Customers in the system</td>
</tr>
<tr>
<td>2-32</td>
<td>Potential customers and marketing expenditure in segment i</td>
</tr>
<tr>
<td>2-33</td>
<td>Dynamics of the system</td>
</tr>
<tr>
<td>3-1</td>
<td>Composition of a feasible bundle</td>
</tr>
<tr>
<td>3-2</td>
<td>An example of reservation price pdf for all bundles in Ω</td>
</tr>
<tr>
<td>3-3</td>
<td>Reservation price pdf for offered bundles in the different price-ranges</td>
</tr>
<tr>
<td>3-4</td>
<td>Expected profit from bundle j in price range k</td>
</tr>
<tr>
<td>3-5</td>
<td>Controllable and non-controllable choices</td>
</tr>
<tr>
<td>3-6</td>
<td>Expected profit function shape</td>
</tr>
<tr>
<td>3-7</td>
<td>Increasing differences in the profit function</td>
</tr>
<tr>
<td>3-8</td>
<td>Optimal price vs. cost</td>
</tr>
<tr>
<td>3-9</td>
<td>Dominant (Pareto) Bundles</td>
</tr>
<tr>
<td>3-10</td>
<td>Example of Network behind CSP_v</td>
</tr>
<tr>
<td>3-11</td>
<td>Pre-optimized objective function</td>
</tr>
<tr>
<td>3-12</td>
<td>Optimal bundle composition</td>
</tr>
<tr>
<td>3-13</td>
<td>Profit maximization with rebate</td>
</tr>
<tr>
<td>3-14</td>
<td>Arrival process - Bundle vs. Bundle with accessory</td>
</tr>
<tr>
<td>4-1</td>
<td>Inline Network of n-1 links</td>
</tr>
<tr>
<td>4-2</td>
<td>One-loop Network of n links</td>
</tr>
<tr>
<td>4-3</td>
<td>Star Network of (n − 1) links</td>
</tr>
<tr>
<td>4-4</td>
<td>Network topologies for 3 nodes</td>
</tr>
<tr>
<td>4-5</td>
<td>Two views of the function $J_{ij}(P_{ij}, x_{ij})$</td>
</tr>
<tr>
<td>4-6</td>
<td>Stochastic formulation taxonomy</td>
</tr>
<tr>
<td>4-7</td>
<td>Network topologies for 4 nodes</td>
</tr>
<tr>
<td>A-1</td>
<td>Service Platform</td>
</tr>
<tr>
<td>A-2</td>
<td>Fixed and variable fees for the two-part tariff</td>
</tr>
<tr>
<td>B-1</td>
<td>Gumbel distribution pdf and cpf</td>
</tr>
<tr>
<td>B-2</td>
<td>Lambert W function</td>
</tr>
<tr>
<td>B-3</td>
<td>Lambert approximation error, where $\tilde{W}(z)$ denotes the approximation in (B.3)</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Customers’ reservation prices .................................................. 18
4.1 Function $J(\cdot)$ under different PDF for the reservation prices ........ 128
A.1 Customers’ reservation prices .................................................. 141
Acknowledgements

First of all, I want to give thanks to God for having given me the opportunity of pursuing my doctoral studies in such a wonderful place with such pleasant friends and colleagues.

I want to thank the members of my thesis committee, Gabriel Bitran, Erik Brynjolfsson, and Yashan Wang, for their very valuable contribution to my research and for every minute they spent with me.

Gabriel has gone far beyond his responsibility as Ph.D. advisor. He has been an outstanding friend supporting me during the different scenarios that I faced in my years as a Ph.D. student. He has been an inspiration not only around the actual research, but also around integrity and respect. I will never forget that every time I entered Gabriel’s office he would ask me about my family and myself before we started working. I will be always indebted to him.

The support from the Engineering School of the Catholic University of Chile, from the Chilean government, and from Fulbright IIE was instrumental in making my studies possible. I will do my best in the future to give back what was given to me.

The continuous prayers, contact, and care from my entire family in Chile have been an important source of support during all of these years. I want to thank all the unconditional help that I received from my in-laws, Patricio and Loreto. Last but not least, special thanks to my parents, Juan Carlos and Angélica, to whom I owe very much. They have shown me during my life that with effort, perseverance, and trust in God, we can always achieve our goals.

I want to thank all the friends I have made at M.I.T. for their constant support and care along the different stages of the Ph.D. My friends Rene Caldentey, Ariel Schilkrut, Paulo Oliveira, and Paulo Gonçalves always gave me outstanding feedback on my work, and more importantly, we became very close friends. I appreciate the valuable contribution and friendship of my friends Opher Baron, Hasan Arslan, Felipe Caro, Fernando Ordoñez, Alejandro Neut, Nico Stier, Rafa Correa, Christian Werner, Kevin Cowan, Carlos Osorio, Juan Pablo Montero, Claudio Raddatz, Vicente Smith, and Rodrigo Huidobro. I want to thank Dirk Beyer and Alex Zhang from HP Labs for their valuable feedback and insights about this research. I would like to highlight my appreciation to Anna Piccolo, Marguerite Baty, and Soulaymane Kachani, who helped me a lot during the last, intensive stage of this dissertation.
I want to give special thanks to the Fink family, who are very close friends of my wife, my children, and myself. Yoel, Zohar, Amit, and Noam have been some of the strongest pillars of support during my experience as a student. I have no doubt that they will be in our thoughts forever.

I would also like to thank my nuclear family. There is no question that my children Berni, Mañi, and Juan Carlitos have been my source of happiness along these years. I have been fortunate to have the opportunity of sharing this experience with them. Lastly, I want to thank to the most special person I have met, Sole, my wife, who has given all of herself to enrich this family experience and make it so smooth and exciting, who has cared about my children and me in an outstanding way, who left her career to support mine, and who has unconditionally loved me through the years. Sole, “mi osita,” I thank you.
Chapter 1

Introduction to Bundle Pricing

High-tech companies strive to attract and retain corporate customers by offering them customized web-based services. More precisely, companies sell bundles of services in order to differentiate their offers from those of their competitors. The present dissertation analyzes the role of the external environment, the customer interface, and the internal environment in the delivery process of a B2B service operation. The analysis places special emphasis on the design of service platforms for customized services. Specifically, it presents an overview of current bundle pricing policies to be considered in order to determine the profit-maximizing prices and configurations of service bundles.

The design of customized services is a dynamic process. The service interface should be a personalized interactive channel with the dual purpose of communicating and selling. Through this channel, information is gathered and services (or products) are delivered in the same package. During each interaction, the customer provides information (e.g., by answering a few questions) that is used to design and deliver dynamically the most suitable service offering. One of the objectives of delivering personalized services is to induce customer loyalty. Customers receiving high quality customized services face barriers when they consider switching to competitors. The size of these barriers depends on the amount of available information, the quality of the information, and the way in which that information is used.

A service delivery system must have two elements in place in order to provide effective highly customized services. First, there must be a service platform that provides the basic offerings common to most small and medium businesses. Second, this platform must be complemented by specific offerings for a given business (customer) that will shape the final service that is going to be delivered. The actual specifications of these offerings vary depending upon business characteristics such as industry, management style, and customer needs (see Figure 1-1).

Customers will purchase the bundle of e-services and products that best meets their needs and gives them the most value. Therefore, policies that determine the pricing and composition of the bundles offered have a profound impact on profitability. Companies
want to maximize their revenue and help their customers choose bundles that best fulfill their needs. In order to attain these goals, companies must understand the notion of bundling, devise bundling strategies, and develop decision models for bundling.

![Diagram of Customized service offering](image)

Figure 1-1: Customized service offering

The remainder of this chapter is divided in three parts. §1.1 examines the current practices of bundling policies currently in use. §1.2 reviews the academic literature and identifies areas of research that can serve as building blocks in the development of bundling strategies. Finally, §1.3 provides examples of how mathematical decision models can aid companies as they strive to increase profitability through optimal bundling and pricing.

1.1 Real world setting

Bundling issues play a central role in today’s computer industry. A computer can be defined as a bundle consisting of several complementary components such as motherboard, RAM, hard disk, and processor among others. Computer companies have always faced the problem of deciding which configurations they will make available to their customers. Therefore, companies in this industry have been developing and refining their bundling strategies for a long time. This section examines the current state-of-the-art practice in the computer industry in order to gain insights into which issues are already adequately addressed by current practices and which deserve more careful analysis.1

Paying a “price” is the last act the consumer performs before taking possession of any product or service. Hence, it is the last message he receives from the firm that is providing the goods, and it is the message that stays with him the longest, from the beginning of his search-journey for a good until he buys it. Price is the last link in the supply chain (just before the delivery and/or the post-sale support) between the supplier and the customer.

---

1 This subsection was developed in a joint effort with two master students (Marco Barbier and Juan Fonseca) from the Sloan Fellows executive MBA program at the Sloan School of Management at MIT. For more details, refer to their master’s thesis [BF02], which was under the supervision of Professor Gabriel Bitran and myself.
In most cases, it is also the first and the strongest message received by the consumer when he searches for a product, much before he blends it with the other attributes of the product to determine the product’s “value.” Price is a fact; value is a judgment. Also, when a customer recalls a purchased good, price is the first characteristic of the good that is remembered.

Consumers have unlimited wants and limited means, so everyday they have to make choices to maximize their happiness within their budget. “Price” is the market’s means to allocate these choices efficiently. The purchases made by the consumer have a blend of both quantitative-rational components and emotional components. The consumer’s perception of a product is the result of a thinking process rich in rational and emotional considerations; “value” is the result and conclusion of all these considerations. Each consumer has a unique conclusion, and hence he attributes to each product a unique value that differs from that of each and every other consumer. We can then infer that the “reservation price” of each consumer is different. What he is willing to pay for a product is based on his perceived value of that specific product, on his needs, and on his particular budget and these characteristics are different for each and every person. That is the reason why we use a probability distribution of the reservation prices in this dissertation.

The ideal world for a supplier would be one in which he held a market monopoly and could price-discriminate every single consumer. In such a scenario, he would sell to each consumer at his reservation price up to the point where the price charged for the last product sold was greater than the marginal cost of producing it (first degree price discrimination). Bundling is one of the tools available to suppliers to price discriminate consumers. But the real world is not as simple as this; the market has competition, the individual reservation prices are not known, and there is a possibility of arbitrage among different consumers.

The Internet allows for an intimate relationship, a one-to-one interactive experience, between the supplier and the customer, at a very low cost. The Internet makes possible every marketing manager’s dream of one-to-one marketing and every sales manager’s dream of mass-customization. Also, thanks to the Internet, suppliers can perform controlled marketing experiments on their potential and present customers and can collect results in real time. For instance, by changing the price of the offering, either randomly or selectively, the supplier can construct a real-time demand curve, test for price sensitivity (elasticity of demand), or implement specific promotions to specific target audiences. Unfortunately, benefits do not come without costs. The Internet also allows consumers to gather plenty of information about different products, their substitutes, and their competitors, and to express publicly their opinions about products-services and suppliers through chat rooms or consumer bulletin boards. The consumer can finally compare and ponder the opportunity cost of every purchase. A very well informed consumer is much more difficult to manipulate than an uninformed one, and hence the Internet makes the
marketing and the sales manager’s jobs much more challenging and difficult.

As a consequence of the commoditization of PCs and their high elasticity of demand, the PC manufacturers that will survive will be those with clear, consistent, and powerful strategies. Barbier and Fonseca [BF02] analyzed four PC manufacturers (Sony, Dell, HP, and Compaq)\(^2\) and studied their corporate, business, marketing, and pricing strategies. They concluded that only Dell and Sony have coherent strategies in these areas. Dell is more focused than the other companies because “computers” are their only business and on-line is its only channel. Sony’s business is blurrier than Dell’s because of Sony’s product breadth and its presence in multiple sales’ channels. HP and Compaq, in their opinion, have serious issues with strategy-alignment and message-communication cohesiveness; in many cases, both firms say X and do Y. Examples of the previous assertion were seen in their corporate, business, and marketing strategies.

Their research also indicated that none of the brand PC manufacturers use a scientific bundling theory. They have noticed that even though basic PC configurations, strictly speaking, are bundles, no PC manufacturer considers them as such. They also indicate that for online shopping the pre-configured machines have a price equal to the price that would result from custom configuring the machine from separate components. One exception to the previous statement occurs when HP bundles its PCs with an HP printer and sells them as a bundled option; here HP uses discount bundling to push the sales of printers. In the retail channel, we have noticed discount bundling when consumers receive rebates from brand PC manufacturers if they add components of the same brand together with the basic PC configuration purchased. For example, if a consumer buys the PC “box” and the monitor from HP, he receives an “$X” rebate, and if he also purchases an HP printer he receives an additional “$Y” rebate.

Brand PC manufacturers consider the sum of the PC components as a mechanism to configure different machines to service different market segments. Firms first configure and price a few basic PC machine types (usually 4 or 5) to serve specific market segments such as those of technology buffs, simple-basic users, etc. Second, firms fine-tune consumer segmentation by pricing different versions (usually 3) of those basic configurations. At the end there are 12 (4×3) to 15 (5×3) different options that cover the complete demand curve. As an example, in Dell’s web site the visitor first chooses the market segment he belongs to, for instance “Home & Home Office”. With his second click, he chooses desktop PCs or laptop PCs. With his third click, he chooses one of these four basic machines: cutting edge technology, superior performance, affordable desktop solution, or simple system-value price. If at this point he chooses superior performance, he is given two additional choices to either customize or to take preconfigured systems whose prices are $669, $1,139, and $1539.

It is clear from the previous example, that Dell and other PC manufacturers use

---

\(^2\)There is no proprietary information in this section, so their names can be disclosed.
mostly second-degree price discrimination (versioning) to extract consumer surplus from a diverse consumer market. These configurations could be viewed as neutral bundling, creating new products to cater to different market segments. Because PC manufacturers do not specifically see PC configurations as bundling, bundling is not consciously used.

For sales through the retail channel, PC manufacturers use a “compact” version of the same strategy that is used for on-line sales. They offer PC bundles called configurations at different prices again to cover the ample consumer spectrum. In the retail channel, the previous technique does not work well because the “versioning” is drastically reduced by the difficulty of customization. In the on-line channel PC firms can have many price versions that almost achieve perfect price discrimination. In retail channels because of supply-chain issues and shelf-space issues, achieving the same degree of price-discrimination is impossible. In spite of the previous reality, there are things that the consumer clearly values and that have low marginal costs such as pre-sale and post-sale service, warranty, etc. These could be bundled with the PC to reduce the consumers’ dispersion of willingness to pay and to entice him to purchase off-line instead of on-line. To achieve a better off-line customer attraction, Barbier and Fonseca [BF02] suggest that PC manufacturers should emphasize the following: PC availability, pre-sale and post-sale assistance and service, warranty, and customer training. They also suggest that PC manufacturers should configure bundles with peripheral equipment. PC firms should directly orchestrate bundling policies at the retail level, not leaving the decision to the retail chain; these two actor’s objectives might not always be aligned. Peripheral equipment is present anyway at most stores and does not represent an added burden to the PC firm.

Only HP offers clear discount bundles with their monitors and printers both on-line and off-line. This pushes printer sales since “printing and imaging” is their most profitable division. HP sells the printer with the expectation of selling ink cartridges in the future (this particular type of bundling is called tie-in sales bundling).

Barbier and Fonseca also suggest that there is room in some PC configurations to take advantage of premium bundling, which is not presently done by any US PC manufacturer. Vobis, the German PC retail chain gains on average a 7-10% premium through “premium bundling,” as Fuerderer et al. [FHW99] point out. Bundling could be used to create new consumer needs and hence to expand the current market for PCs. Today, with the exception of Sony, bundling is only used to create new products, to achieve price discrimination, or to push sales of some specific accessories such as printers.

The consumer in general likes bundling because it feels like a “deal” (discount) and because, in most cases, the bundle does the job for him. This assertion was corroborated by interviews with retail stores, by a consumer survey, and by the fact that 50% of on-line purchases by consumers at Dell are of preconfigured machines. This is a great result since it tells PC manufacturers that there is plenty of room and receptiveness from consumers to do much in the currently underdeveloped field of bundling.

Barbier and Fonseca’s research also claims that PC manufacturers have two choices:
play Dell’s price-commodity game or play a differentiation game where they strive for a tight bonding with their customers and can charge a premium for their products (Sony, Apple). Additionally, it is difficult to believe that a surviving firm in the PC industry could prosper without scientific marketing tools that would allow it to connect all the segments of the “demand supply chain” in real time, and would allow it to manage those components dynamically, consistently showing the consumer the optimal price for its products.

1.2 Background and literature review

Bundling strategy is an important tool for companies that serve customers with heterogeneous preferences. When price discrimination is not possible, companies must rely to a great extent on pricing and design policies in order to maximize their profits. In its simplest form, bundling consists of collecting goods or services in a package and selling them at a usually discounted package price [GP88]. In order to understand the effectiveness of bundles, note that a customer’s reservation price (the maximum amount of money that customers are willing to pay) for a bundle is the sum of the reservation prices of the bundle’s components. Since the reservation prices for individual components varies from customer to customer, bundles allow companies to capture more consumer surplus from the buyers because excess consumer surplus is transferred from one component of the bundle to another. In this way, companies implicitly price discriminate. When implemented effectively, bundling strategies can accomplish a number of objectives such as extension of monopoly power, price discrimination, reduction in complexity costs, reduction in transaction costs, creation of barriers to entry and economies of scope and scale.

According to Simon and Wuebker [FHW99, part 1], bundling plays an increasingly important role in many industries. Some companies even base their business strategies on their bundling strategies. A renowned example is Microsoft. By smartly combining its application software into the “Office” bundle, Microsoft increased the market share of Access and PowerPoint. These two less attractive components were bundled with the more attractive components Word and Excel. Simon and Wuebker [FHW99, part 1] also mention that bundling is particularly popular in the service sector. Some examples are vacation packages (airline ticket, hotel accommodation plus rent-a-car), insurance packages, restaurant menus (appetizer, entrée, dessert), and telecommunication packages (local, long-distance, internet access, cellular phone). Altinkemer [Alt01] points out that since the 1996 Telecommunication Deregulation Act, telecommunication companies such as local and long-distance phone companies, Internet service providers (ISPs), cable companies, and entertainment companies have been battling to be the one-stop-shop for infotainment (information entertainment). Therefore, many companies are joining forces with others to round out their offerings, as evidenced by mergers between MCI WorldCom and Sprint,
SBC and Ameritech, AT&T and TCI, and Time Warner. [Duk00] and [Swa00] suggest that business customers like to bundle more than residential customers because business customers required more services and want only one bill. In telecommunications, bundling increases customers’ loyalty because of the high switching costs involved. Consequently, companies fiercely compete to become their customers’ one-stop communication partner.

How can a company decide whether to bundle its products and determine the profit-maximizing prices? Most companies do not know their customers’ reservation prices. However, they may be able to estimate the distribution of reservation prices by conducting market surveys. The results of these surveys can then be used to design a pricing strategy. A simple example can illustrate this concept and the mechanism behind it. Consider a company that needs to decide whether or not to bundle two services. Table 1.1 shows the reservation prices of each customer for each service.

<table>
<thead>
<tr>
<th>Customer 1</th>
<th>Service A</th>
<th>Service B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer 1</td>
<td>$12,000</td>
<td>$3,000</td>
</tr>
<tr>
<td>Customer 2</td>
<td>$10,000</td>
<td>$4,000</td>
</tr>
</tbody>
</table>

Table 1.1: Customers’ reservation prices

If the company does not bundle the services, the maximum revenue it can earn is $26,000. The optimal prices that maximize the revenues are \( P_A = \$10,000 \) and \( P_B = \$3,000 \), and in this case both customers will purchase both services. If the company has the option of offering both services as a bundle, they can earn $28,000, since \( P_{AB} = \$14,000 \) makes the bundle attractive for both customers. In economic terms, the consumer surplus from the highly valued service \( A \) is transferred to the less valued service \( B \) ($1,000 are transferred from service \( A \) to service \( B \) in customer 1). Since customers are heterogeneous, the company is better off bundling.

According to the literature, the seller faces three main alternative strategies to offer its services: (a) no bundling, where the seller prices and offers the component services as separate items, not as bundles; (b) pure bundling, where the seller prices and offers the component services only as a bundle and not as individual items; and (c) mixed bundling, where the bundle as well as the individual component services are priced and offered separately. No one strategy may always be the best. The optimal strategy and the bundle itself strongly depend on the distribution of the customers’ reservation prices.

Bundling has been broadly studied from consumer behavior, economic and marketing perspectives. In the area of consumer behavior, studies such as those by Yadav ([YM93] and [Yad94]) have primarily focused on the process by which consumers evaluate product bundles. Suri and Monroe [FHW99, part 3] looked into the effects of contextual factors, such as consumers’ prior purchase intentions on their evaluation of product bundles. Others have examined the extent to which bundling can stimulate demand under specific promotion conditions. Their results have shown that increased promotion activity on the
individual items can significantly lower buyers’ evaluations of bundle offers featuring the individual items. A recent study by Herrmann, Huber, and Coulter [FHW99, part 3] found that price discounts and complementarity of bundle components appear to be key drivers of purchase intention. Many conclusions from these studies are clearly compatible with Prospect Theory, (Kahneman and Tversky [KT79]) which postulates that the impact of perceived losses is greater than the impact of perceived gains of the same magnitude.

Adams and Yellen [AY76] provided a building block which proved to be fundamental to the understanding of bundling from the perspective of economics and marketing. In their paper, they introduced a two-dimensional graphical framework to analyze the effect of bundling as a price discrimination tool. More recent studies ([Sch82], [DC84], and [Sch84]) have presented numerical criteria to determine which bundling strategy is more profitable. This line of research is based on the examination of bundles either containing an unwanted component (value reducing) or providing additional value beyond the aggregated values of the individual items (value enhancing). Salinger [Sal95] extends the result of [AY76] and [Sch84] using a graphical analysis. He suggests that in the case of independent linear demand functions and relatively high costs, positively correlated reservation prices may increase the incentive to bundle if economies of scope prevail. Cready [Cre91] derived profitability conditions for premium bundling (bundles are sold at a premium price), which are possible only when individual products alone offer little benefit. Recently, Bakos and Brynjolfsson [BB99] used the Law of Large Numbers to obtain the intriguing result that bundling a large amount of unrelated and nearly costless information goods can be profitable. Customers have varying valuations for individual goods, but tend to average them when considering the bundle. Thus, there are more consumers with moderate valuations for the bundle than for individual components (see also [BB97] and [BB00]). Finally, [GS01] and [SG01] show that bundling services can hurt consumption. The sunk cost effect induces product consumption because people are more aware of costs. On the other hand, bundling offsets the sunk cost effect since it hides costs in a way, and thus, consumption decreases.

Literature on bundling in economics and marketing does not focus on suggesting specific approaches to meet the needs of a high-tech service provider in an integrative way. Most articles published on bundling primarily provide theoretical rationales for the bundling concept and project several contexts that are aptly suitable for bundling. However, there are a few studies that provide practitioners with optimization approaches for designing and pricing bundles, which we use as a motivation base to build our models. Hanson and Martin [HM90] present a mathematical programming formulation to determine the profit maximizing bundle configurations and prices, without explicitly considering the full range of feasible problem solutions. Venkatesh and Mahajan [VM93] consider two dimensions in the consumer decision-making process (time and money) and find the optimal price for a given bundle under different bundling strategies. Finally, Fuerderer et al. [FHW99, part 2 page 61] demonstrate how a single firm can optimize the
design and the price of its product line when reservation prices and consumer choice behavior are uncertain. In this study, they use an algorithm to solve a mix-integer stochastic programming model to determine bundle prices and configurations.

1.3 Bundle pricing problems

In this section, our objective is to clarify and establish the challenges that high-tech companies are facing in terms of bundling their services and products offered through the Internet. We present a list of questions and assumptions that provide enough insights to illustrate two simple models, and that will help managers make decisions about bundling and pricing their services.

Our objective is to propose a probabilistic approach that enables high-tech companies to determine optimal bundle prices and configurations under different bundling strategies. In order to model bundling problems there are several questions that have to be clarified. Some of them are as follows:

- Are there any constraints when making bundles? (e.g., every bundle must include Internet access; there are certain technologically incompatible services)
- Is there a correlation among services? (e.g., if one service is purchased, another service has to be included in the bundle)
- In case there is demand correlation between two bundles, does this correlation depend on the decision variables? (e.g., prices and selection)
- What factors drive the consumer decision-making process? How do customers decide whether to buy a bundle or not? (e.g., price, time, know-how)
- Do we want to consider different segments among customers? How many segments? How can a high-tech company determine to which segment a particular customer belongs?
- What are the costs involved? Are there any segment specific costs (e.g., shipping, support)? Are there any common across-segment costs for a particular bundle, or any setup costs?
- What specificities of an Internet-services setting should be taken into account in the model?

Although this dissertation does not limit its scope only to the physical or Internet channel, it is worth mentioning that Figure 1-2 just depicts an unexplored niche (bundling services over the Internet) which could be addressed in the future. Before we present the basic modeling tools used in the literature, we introduce some notations.
Figure 1-2: Bundling opportunity research

\[ B : \] number of available bundles.
\[ S : \] number of customer segments.
\[ \Pi : \] total profits.
\[ p_{ij} : \] price of bundle \( j \).
\[ \theta_{ij} : \] equal to 1 when customer segment \( i \) selects bundle \( j \), 0 when not.
\[ \psi_{ij}(\cdot) : \] pdf of customer segment \( i \)'s reservation prices for bundle \( j \).
\[ s_{ij} : \] expected consumer surplus when customer segment \( i \) choose bundle \( j \).
\[ \hat{s}_i : \] highest expected consumer surplus for customer segment \( i \), i.e., equal to \( \max_{1 \leq j \leq B} s_{ij} \)
\[ c_{ij} : \] cost of bundle \( j \) for customer segment \( i \).

Note: When dealing with just a single-segment the index \( i \) is dropped.

Because of real life simplifications, we make several assumptions based on the literature that are crucial for the development of the model. These are the following:

1. Prices for every bundle are determined by only one profit-maximizing firm.
2. All customer segments face the same price for bundles.
3. Every customer maximizes his consumer surplus (reservation price minus the product price), given the bundle prices.
4. The probability density functions (pdf) of reservation prices of all customer segments are known for all bundles.
5. Customer segment \( i \)'s utility for bundle \( j \) only depends on the realized bundle price \( p_{ij} \).
6. The benefit of a duplicate component is zero, and component resale is not possible.
A customer never purchases when consumer surplus is negative.

In what follows we explore formulations presented in the literature for both single and multiple segment cases. In the first case only one customer segment is considered, i.e., the whole customer population has the same reservation price distribution and is offered the same bundles (see Figure 1-3). As in Hanson and Martin [HM90], the objective is to find the price and the bundle that maximize a firm’s profits while at the same time maximizing consumer surplus for that given price. The trade-off faced here is price vs. consumer surplus. The higher the price, the higher the profits will be, but the lower the consumer surplus will be. It may cause another bundle to take the highest surplus position, yielding to very different profits.

![Figure 1-3: Single customer segment](image)

As shown in (1.1), profits are calculated as the selected bundle’s profit (price minus cost) times the proportion of the potential customers that will be willing to pay the price $p_j$. Thus, profits are

$$\Pi = \sum_{j=1}^{B} \theta_j (1 - \Psi_j(p_j))(p_j - c_j) \quad (1.1)$$

where $\Psi_j(\cdot)$ is the cumulative distribution function (cmf) of the reservation prices for bundle $j$, defined as

$$(1 - \Psi_j(p_j)) = \int_{p_j}^{\infty} \psi_j(r)dr. \quad (1.2)$$

Now from the consumers’ point of view we have to maximize their surplus. Given a price $p_j$ for bundle $j$, the expected consumer surplus of a customer who chooses bundle $j$ is given by

$$s_j = \int_{0}^{\infty} \max\{r - p_j; 0\} \psi_j(r)dr = \int_{0}^{p_j} \psi_j(r)dr + \int_{p_j}^{\infty} (r - p_j) \psi_j(r)dr,$$
which reduces to
\[ s_j = \int_{p_j}^{\infty} (r - p_j) \psi_j(r) dr. \] (1.3)

Using (1.3) we can establish the following conditions (1.4a-1.4c) that will ensure (given the prices) that only one bundle is selected and that this bundle is the one that leaves the highest consumer surplus.

\[ \hat{s} = \sum_{j=1}^{B} s_j \theta_j \] (1.4a)

\[ \hat{s} > s_j, \quad \forall j = 1, \ldots, B \] (1.4b)

\[ \sum_{j=1}^{B} \theta_j = 1 \] (1.4c)

Prices need to be nonnegative and the bundle selection indicators are binary variables. Let us impose some conditions to fulfill these requirements.

\[ p_j > 0, \quad \forall j = 1, \ldots, B \] (1.5a)

\[ \theta_j \in \{0, 1\}, \quad \forall j = 1, \ldots, B \] (1.5b)

Without considering the maximization of consumer surplus, the decision of which bundle to offer is given by the maximization of (1.1) subject to (1.4c), (1.5a) and (1.5b). It reduces to determining which bundle contributes with highest profit. That is, we can analyze each component of (1.1) separately since there is no relation among them, other than the constraint that just one bundle is chosen. The first-order conditions are given by

\[ \frac{\partial}{\partial p_j} [(1 - \Psi_j(p_j))(p_j - c_j)] = 0, \quad \forall j = 1, \ldots, B \]

which imply that the bundles’ optimal prices are the fixed points of the following conditions:

\[ p_j^* - c_j = \frac{1 - \Psi_j(p_j^*)}{\psi_j(p_j^*)}, \quad \forall j = 1, \ldots, B \] (1.6)

For most probability density functions, conditions (1.6) will not yield closed form solutions for the optimal prices \( p_j^* \). In order to find the optimal prices, we can use a graphical approach by simultaneously plotting the cmf \( \Psi_j(p_j) \) and the function \( Y(p_j) = 1 - \psi_j(p_j)(p_j - c_j) \) for every bundle. The intersection of these two functions will determine the optimal prices \( p_j^* \).

Now, let us consider multiple customer segments in order to take into account customers’ heterogeneity (see Figure 1-4). This problem is much more complex than the
previous one. Here, we still want to determine optimal prices for each bundle, but there is an implicit constraint in that all segments face the same prices.

The mathematical model formulation is basically the same as in the previous section, but now variables and parameters use an index to indicate the segments, except the price $p_j$ which is the same across all segments (implicit constraint).

Maximize $\Pi = \sum_{i=1}^{S} \sum_{j=1}^{B} \theta_{ij} (1 - \Psi_{ij}(p_j))(p_j - c_{ij})$

subject to

\begin{align*}
  s_{ij} &= \int_{p_j}^{\infty} (r - p_j)\psi_{ij}(r)dr, \quad \forall i = 1, ..., S \quad \forall j = 1, ..., B \\
  \hat{s}_i &= \sum_{j=1}^{B} s_{ij}\theta_{ij}, \quad \forall i = 1, ..., S \\
  \hat{s}_i &> s_{ij}, \quad \forall i = 1, ..., S \quad \forall j = 1, ..., B \\
  \sum_{j=1}^{B} \theta_{ij} &= 1, \quad \forall i = 1, ..., S \\
  p_{ij} &> 0, \quad \forall i = 1, ..., S \quad \forall j = 1, ..., B \\
  \theta_{ij} &\in \{0, 1\}, \quad \forall i = 1, ..., S \quad \forall j = 1, ..., B.
\end{align*}
Chapter 2

Pricing Service-Bundle Contracts with High Setup Cost

2.1 Introduction

This chapter addresses the problem of determining bundle pricing policies of IT services contracts when there is a high setup cost. The objective is to determine, for each bundle, the price that maximizes total profits. Total profits are the sum of individual profits over all segments, customers, and bundles. The first task in developing pricing policies is to understand consumer behavior. Specifically, we must define the manner in which consumers react to a set containing different bundles. Customers will undergo a judgement task where they determine which of the bundles best suits their preferences and a decision task where they decide whether or not they will make a purchase. These two tasks will be referred to as the “choice” decision and the “buy” decision hereafter. A mathematical formalization of the behavioral model must make assumptions concerning the sequence in which the consumer makes the “choice” and “buy” decisions. Since this sequence will be different depending on contextual variables such as industry and the type of product, it is important to understand the effect that the selected model will have on the final policies. To this effect, we model customer behavior in two different ways, which are described below.

The objective of the choice model that we use is to map different types of consumers and bundles into probabilities of purchase. One way to devise such a model is to assume that customers determine their preferred bundle (“choice”) from a given set and decide whether or not they will make a purchase (“buy”) simultaneously. This is denoted the “one-step model.” This model (depicted in Figure 2-1a) applies to the situation where an arriving customer faces the following choices: buy bundle 1, buy bundle 2, ..., buy bundle n, or do not buy any bundle. A second approach, called the “two-step model” assumes that the “choice” and “buy” decisions are made sequentially. As depicted in Figure 2-1b, the two-step model applies to situations when customers first choose among the bundles
and then, given that decision, decide whether to buy it or not. We now turn to a more
detailed analysis of how each decision process can be formalized mathematically. Our
objective is to derive choice probabilities and profit functions for each of the two decision
processes and understand how they relate to each other.

![Decision Process Diagram]

(a) One-step choice

(b) Two-step choice

Figure 2-1: Two customer behavior models

The one-step decision process can be formalized through a multinomial logit model
discussed in section 2.4.1). According to this model, customers choose the bundle that
maximizes their expected utility. Let \( \alpha_j \) be the probability of choosing bundle \( j \), and let
\( \alpha_0 \) denote the probability that the customer does not buy any of the available bundles.
It immediately follows that \( \sum_{j=0}^{B} \alpha_j = 1 \), where \( B \) is the number of available bundles.
Furthermore, let \( \pi_j \) denote the probability of choosing bundle \( j \) when the consumer is
confined to make a purchase (i.e., the option of not making a purchase is unavailable).
It follows that \( \sum_{j=1}^{B} \pi_j = 1 \). Figure 2-2 depicts the above probabilities for the case when
\( B = 3 \). Note that the shaded areas (probability of buying something) in (a) and (b) are
the same, i.e.,

\[
\alpha_1 + \alpha_2 + \alpha_3 = \pi_1(1 - \alpha_0) + \pi_2(1 - \alpha_0) + \pi_3(1 - \alpha_0).
\] (2.1)

By using the multinomial logit probabilities

\[
\alpha_j = \frac{e^{V_j}}{\sum_{k=0}^{B} e^{V_k}} \quad \forall j = 0, \ldots, B
\] (2.2)

and

\[
\pi_j = \frac{e^{V_j}}{\sum_{k=1}^{B} e^{V_k}} \quad \forall j = 1, \ldots, B
\] (2.3)

where \( V_j \) is the customer’s strict utility of choosing bundle \( j \). This formulation allows
us to make a stronger statement than (2.1), which is also important in understanding
the relationship between the one-step and the two-step models. This statement will be
presented later in Proposition 1.
The choice probabilities derived in the previous paragraph enable us to define the profit function. Let $\tilde{U}_j(p_j)$ be the i.i.d. random variables representing the utility functions for purchasing each bundle, and let $\tilde{U}_0$ denote the utility of not purchasing any bundle. Furthermore, let $\tilde{d}$ be a random variable which indicates the index of the bundle that gives the highest utility to the customer, i.e.,

$$\tilde{d} = \arg \max \{\tilde{U}_1(p_1), \tilde{U}_2(p_2), \ldots, \tilde{U}_B(p_B), \tilde{U}_0\}.$$ 

The profit received from a particular customer from segment $i$ is denoted by $\tilde{V}_{id}$. The expected value of $\tilde{V}_{id}$ is given by

$$E[\tilde{V}_{id}] = \sum_{j=1}^{B} V_{ij}(p_j) \Pr(\tilde{d} = j).$$

Summing the expected profit over all customer segments yields the total expected profits:

$$\Pi(p) = \sum_{i=1}^{s} \sum_{j=1}^{B} V_{ij}(p_j) \Pr(\text{choose and buy } j). \quad (2.4)$$

The profit function is constructed using customer choice among the offered bundles as in Hanson and Martin [HM96] (Multinomial logit profit function). This concludes the formulation of the one-step model.

In the two-step model, the probability that a given bundle will be chosen and bought
can be derived using conditional probabilities:

\[
\Pr(\text{choose and buy } j) = \Pr(\text{choose } j) \Pr(\text{buy } j|\text{chose } j).
\]

Let us first define some notations that will be essential in making the distinction between the “choice” process and the “buy” process:

\[
V_{ij}(p_j) = \text{profit generated when a customer from segment } i \text{ buys bundle } j \text{ at price } p_j
\]

\[
X_{ij}^k = \begin{cases} 
1 & \text{if customer } k \text{ from segment } i \text{ chooses bundle } j \\
0 & \text{otherwise}
\end{cases}
\]

\[
Y_{ij}^k = \begin{cases} 
1 & \text{if customer } k \text{ from segment } i \text{ buys bundle } j \text{ given he chose } j \\
0 & \text{otherwise}
\end{cases}
\]

We now turn to a key assumption in the development of the choice model:

**Assumption 1** A customer will buy a bundle only if the following two conditions hold:
1. The bundle maximizes his/her utility given all the available bundles, and (2) the customer is willing to pay the bundle’s price.

Figure 2-3 depicts the choice probabilities derived from the two-step model. The probabilities \( \pi_j \) correspond to the first part of the assumption. Part (2) of the assumption is accounted for by establishing that consumers make the “buy” decision based on the reservation price. The shaded area (probability of buying) in this approach is driven by the reservation price distribution functions \( \Psi_j(\cdot) \).

![Figure 2-3: Choice probabilities in the two-step model (case \( B = 3 \)](image-url)
Total profits are given by

\[ \Pi(p) = \sum_{i=1}^{S} \sum_{j=1}^{B} \sum_{k=1}^{N_i} X_{ij} Y_{ij} V_{ij}(p_j) \]

\[ = \sum_{i=1}^{S} \sum_{j=1}^{B} V_{ij}(p_j) N_i \sum_{k=1}^{N_i} X_{ij} Y_{ij} / N_i \]

The expression, \( \sum_{k=1}^{N_i} X_{ij} Y_{ij} / N_i \), can be interpreted as the proportion of customers from segment \( i \) that will choose and buy bundle \( j \). Equivalently, it is equal to the probability that a customer from segment \( i \) will choose and buy bundle \( j \).

Figure 2-3 is not the same as Figure 2-2. In particular, \( \alpha_0 \) depends on the vector of prices \( p \), while each reservation price distribution function \( \Psi_j \) only depends upon its corresponding price \( p_j \). This difference between Figures 2-2 and 2-3 makes clear the similarities and differences of the two decision models. The proposition below helps formalize this idea.

**Proposition 1** Given the strict utility functions \( \hat{V}_j(p_j) \) and the choice probabilities \( \alpha_j \) and \( \pi_j \) in (2.2) and (2.3), we claim that \( \alpha_j = \pi_j(1 - \alpha_0) \) for all \( j = 1, \ldots, B \), i.e., the probabilities \( \pi_j \) are a normalization of the \( \alpha_j \).

**Proof.** Using (2.2) and (2.3) this proof is straightforward.

\[ \alpha_j = \frac{e^{\hat{V}_j}}{e^{\hat{V}_0} + \sum_{k=1}^{B} e^{\hat{V}_k}} \]

\[ = \frac{(e^{\hat{V}_j} / \sum_{k=1}^{B} e^{\hat{V}_k})}{(e^{\hat{V}_0} / \sum_{k=1}^{B} e^{\hat{V}_k}) + 1} \]

\[ = \pi_j \frac{\sum_{k=1}^{B} e^{\hat{V}_k}}{e^{\hat{V}_0} + \sum_{k=1}^{B} e^{\hat{V}_k}} \]

\[ = \pi_j(1 - \alpha_0) \]

closing this proof. ■

The remainder of this chapter is organized as follows. In §2.2 we introduce the basic concepts that will be used in our models such as financing costs, customer’s lifetime distribution, and computation of the monthly fee. In §2.3 the per-customer expected profit is modeled and computed under some assumptions. The one-step decision process is presented and analyzed in §2.4, alongside some numerical results. In §2.5 we develop the two-step model, presenting an approximation that allows a deep analysis and numerical experiments of the model. In §2.6 we explore the evolution of the customer base through the time, defining pricing policies for each possible state of the system. At each period
current customers can defect with probability $1 - q$ (function of the price) and the arrival rate of new customers also depends upon the price. Finally, in §2.7, we provide some extensions and directions for further research.

### 2.2 Bundle monthly fee

Recall that this chapter addresses the problem of pricing bundles of IT services contracts when there is a high setup cost. The problem is motivated by a real world situation with the following characteristics where the bundle has four fully integrated elements: hardware (HW), software (SW), services, and support. Customers must sign a contract agreeing to pay a fixed monthly fee. The company must finance the HW and SW while the services and support are paid by a monthly fee. This section addresses the problem of computing the monthly fee to be charged for every offered bundle. This computation takes into account the fact that customers can defect before the end of the contract by considering a distribution of the customers’ lifetime.

#### 2.2.1 Example: Extended Office

The quotations below\(^1\) illustrate the Extended Office bundle, which has hardware (HW), software (SW), services, and support components.

“Make your life simpler. With the Extended Office solution you get computer hardware, software, services, and support fully integrated and ready to go. This is a perfect solution for business owners who do not want to invest in technology infrastructure but simply subscribe to one. Let us take care of the computer headaches while you focus on your core business.”

“Extended Office combines Internet-delivered services with state-of-the-art hardware, popular software and our outstanding support to give you a comprehensive computing solution. Now even the smallest companies can have the level of expertise and support found only in large corporations with full-time information technology departments.”

“The service includes Microsoft Office 2000 Professional software, guaranteed 90-second response at our 24/7 hardware and software support center, virus protection, email, unlimited nationwide Internet access, secure online data backup, instant bug fixes, and automatic software updates - all on powerful desktops and/or laptops backed by our comprehensive service guarantee.”

“Our company offers subscription plans for desktop and notebook users at affordable and predictable monthly rates. And of course, you can tailor each solution with optional features and services.”

\(^1\)Extracted from sponsor company’s documentation.
“The Extended Office service takes your business to the next level by extending your computing power and providing a worry-free solution.”

2.2.2 Financing costs

Customers who buy the “Extended Office” bundle receive the HW and SW components at once. The payment, on the other hand, is made in monthly payments. Therefore, the company has to finance HW and SW for each customer. Services and support do not need to be financed because these components are delivered and paid for on a monthly basis. In this section we compute the monthly payment \( x_n \) for HW and SW given the total cost \( K \) and the interest rate \( r \). We consider the case where the company is considering contracts of 36 months. The monthly payment will be a function of \( n \), the number of periods or payments.

As in classic finance theory, the monthly payment is computed as a difference of two perpetuities: one starting in period 0 and the other starting in period \( n \). The difference (in today’s money) has to be equal to the total HW and SW costs, as in (2.5).

\[
K = \frac{x_n}{r} - \frac{x_n}{r} \frac{1}{(1 + r)^n} \tag{2.5}
\]

Solving for \( x_n \) yields

\[
x_n = K r \frac{(1 + r)^n}{(1 + r)^n - 1}, \quad \forall n \geq 1. \tag{2.6}
\]

The exponential nature of (2.6) is most vividly captured through a graph. Figure 2-4 illustrates the monthly payment fee for different contract lengths. It assumes that total initial costs for HW and SW is $100 \( (K) \), the annual interest rate is 8%, (i.e., a monthly interest rate \( (r) \) of approximately 0.64%).

In this example, a customer who signs a contract for 36 months would have to pay $3.12 per month. If this customer stops paying before the end of the contract (e.g., he goes out of business), the service provider will lose money. Therefore, it is important to consider the expected customer’s lifetime when deciding what the monthly payment fee should be.

2.2.3 Customer’s lifetime distribution

The portfolio of customers exhibits significant internal differences in the customers’ lifetimes. Customers leave the service without warning. The company finds out that a valuable customer left the service after a few months of not receiving payments or after a notification that the customer went out of business. If the monthly fee is \( x_{36} \) and the customer has not finished the 36-month contract, the company will lose money. If there
Figure 2-4: Monthly payment based on contract length

is a critical mass of customers that leaves the service prior the contract expiration, this event must be considered in the monthly fee.

Figure 2-5 depicts the probability density function (pdf) of the customers’ lifetimes. A high proportion of customers will stay longer than the contract length (36 months), but also a non negligible proportion of customers does not finish the contract.

Figure 2-5: Customers’ lifetime distribution

The dark-shaded area in Figure 2-5 represents the probability that a customer will be in the service more than \( i \) months but less than \( i + 1 \) months. It can be represented by

\[
P(i \leq t < i + 1) = \int_{i}^{i+1} f(t) dt.
\]
A weighted average of monthly payments for different contract lengths can be computed using this distribution.

2.2.4 The monthly fee

This section shows how to determine the optimal monthly fee. This computation is done so as to minimize the costs incurred when customers leave the service before the end of their contracts. The analysis builds on the previous sections by taking into account the customers’ lifetimes through a pdf such as the one depicted in Figure 2-5.

Let $x_1$ denote the monthly fee to be charged to customers who stay more than a month and less than two months, $x_2$ denote the fee for customers who stay more than two months and less than three months, and so on. If the contract duration is 36 months, then let $x_{36}$ denote the first 36 months. This fee should decrease substantially beginning in the 37th month because the HW and SW costs have already been covered. There is no way to charge the HW and SW costs if the customers stay for less than one month. For that reason we assume that $x_0 = K$, so the company will charge everybody a little bit more to cover the costs incurred by these delinquent customers.

The monthly fee $\bar{x}$ can be written in the form

$$\bar{x} = x_0 P(t < 1) + \sum_{i=1}^{35} x_i P(i \leq t < i+1) + x_{36} P(t \geq 36)$$

which is equivalent to

$$\bar{x} = \left[ K \int_0^1 f(t) dt \right] + \left[ \sum_{i=1}^{35} x_i \int_i^{i+1} f(t) dt \right] + \left[ x_{36} \int_{36}^\infty f(t) dt \right] . \quad (2.7)$$

The company should charge its customers a monthly fee composed of the sum of the expression in (2.7), plus the services and support fee $S$, plus a margin $\pi$. Therefore, $\bar{x} + S + \pi$ is the monthly fee that takes into account the fact that some customers will defect before the expiration of the service contract. This fee can be rewritten as $\bar{x} + S + \pi = c + \pi$, which makes explicit the fact that $(\bar{x} + S)$ is a cost-related component and $\pi$ is a profit-related component.

2.3 Expected profit per customer

In Section 2.2 we presented a methodology to compute the monthly payment for a bundle composed of services and products under a contract with a fixed duration. This monthly payment takes into account the costs incurred when customers leave the service before the end of the contract.
In this section we include the previous idea in a price bundling model for eServices. We assume that the bundle is composed exclusively of services and that there is a high setup cost in order to implement the service delivery system for each customer. This setup cost (e.g., the cost of a PC) is equivalent to the hardware and software cost presented in Section 2.2. The idea now is to spread the setup cost over time in order to avoid having the customer to pay a high initial cost. New customers sign contracts for an agreed period of time. These customers will pay a monthly fee that includes the service cost and part of the total setup cost.

The cost-related component of the monthly fee depends on the prices of the bundles. This is because an increase in prices will result in shorter customer lifetimes. As depicted in Figure 2-6, an increase in price shifts the customers’ lifetime distribution curve to the left. A shorter customer lifetime means that more customers will turn over before the contract’s expiration date. Consequently, if companies want to recover the initial investment they made through the setup costs, they will have to increase the monthly fees.

2.3.1 Single-period profit model

Let us recall some definitions from previous sections in order to build our eServices price bundling model.

The following definition establishes what the monthly payment corresponding to setup cost should be, taking into account customers’ lifetime.

**Definition 2**  Given total costs $K$, contract length $T$, customers’ lifetime distribution $f(t)$, and price $p$, the monthly payment per customer, assuming that some customers will turn over before the end of the contract, is given by

$$
\bar{x}(K, T, f(t|p)) = \left[ K \int_0^1 f(t)dt \right] + \left[ \sum_{k=1}^{T-1} x_k \int_k^{k+1} f(t)dt \right] + \left[ x_T \int_T^{\infty} f(t)dt \right]
$$

where

$$
x_n = Kr \frac{(1+r)^n}{((1+r)^n - 1)}, \quad \forall n \geq 1.
$$

**Note:** See Section 2.2 for a complete explanation of Definition 2.

We now formulate for the first time our stochastic bundling problem as follows:

$$
P \) \quad \text{max} \quad \sum_{i=1}^{S} \sum_{j=1}^{B} (p_j - c_{ij}(p_j))N_i \text{Pr}(\text{buy } j)
$$

subject to

$$
p_j \geq 0, \quad \forall j
$$
where \( S \) is the number of customer segments and \( B \) is the cardinality of the set of available bundles.

In (2.8) total profits across all customer segments are maximized. In each segment \( i \), for each bundle \( j \), profits are computed as the product of the unit margin \((p_j - c_j(p_j))\), the segment size \( N_i \), and the proportion of customers \( \Pr(\text{buy } j) \) willing to choose and buy bundle \( j \).

There are no fixed costs associated with offering a bundle because we are dealing with an Internet setting. Offering any bundle built out of the available eServices adds no cost to the company.

There are no constraints on the set of bundles offered, which makes the problem less difficult. In order to compare different bundling strategies (i.e., separate pricing, pure bundling, and mixed bundling), the model should be run with different bundle sets according to each strategy.

What is new in our formulation is that costs are a function of price. Costs are composed of both services cost and setup cost. Let \( S_j \) be the total monthly cost of all services in bundle \( j \), and let \( \bar{x}_{ij}(K, T, f(t|p_j)) \) be the monthly setup cost (Definition 2). As we can see, the latter cost component is a function of price. Therefore, total monthly costs \( c_{ij}(p_j) \) are also a function of price, as shown in the following definition.

**Definition 3** Total monthly cost of bundle \( j \) for customer segment \( i \) is given by

\[
c_{ij}(p_j) = S_j + \bar{x}_{ij}(K, T, f(t|p_j)).
\]

**Assumption 2** When the price increases, the probability that a customer will be in the system at least \( n \) periods decreases, i.e., if \( p'' < p' \) so it makes \( F(t|p''') < F(t|p') \) for all \( t \), where \( F(\cdot) \) is the cumulative probability function of the reservation prices. Figure 2-6 shows the shift of the customer lifetime pdf \( f(t|p_j) \) when increasing the price.

**Proposition 4** The higher the prices \( p_j \), the higher the costs \( c_j(p_j) \).

**Proof.** Since \( S_j \) does not depend on the price, it is sufficient to show that the monthly payment per customer increases with \( p_j \). For simplicity, we call this payment \( \bar{x}(p) \). We start by rewriting it.

\[
\bar{x}(p) = \left[ K \int_0^1 f(t) dt \right] + \left[ \sum_{k=1}^{T-1} x_k \int_k^{k+1} f(t) dt \right] + \left[ x_T \int_T^\infty f(t) dt \right]
\]

\[
= \sum_{k=0}^{T-1} x_k (F(k+1|p) - F(k|p)) + x_T (1 - F(T|p)) \quad \text{where } x_0 = K. \tag{2.9}
\]
Figure 2-6: Shift on customer lifetime pdf when changing $p$.

Let $\bar{p}$ be a price greater than $p$, so by Assumption 2 we can write $F(t|\bar{p})$ as $F(t|p) + \Delta_t$ where $\Delta_t > 0$ for all $t \geq 1$ is the positive increase in probability. Note that $\Delta_0 = 0$. Now, the monthly payment for this higher price is given by

$$
\bar{x}(\bar{p}) = \sum_{k=0}^{T-1} x_k (F(k+1|\bar{p}) - F(k|\bar{p})) + x_T (1 - F(T|\bar{p}))
$$

$$
= \sum_{k=0}^{T-1} x_k (F(k+1|p) + \Delta_{k+1} - F(k|p) - \Delta_k) + x_T (1 - F(T|p) - \Delta_T). \quad (2.10)
$$

Subtracting (2.9) from (2.10) we get

$$
\bar{x}(\bar{p}) - \bar{x}(p) = \sum_{k=0}^{T-1} x_k (\Delta_{k+1} - \Delta_k) - x_T \Delta_T
$$

$$
= \sum_{k=1}^{T} (x_{k-1} - x_k) \Delta_k
$$

From (2.6) and Figure 2-4 we know that $x_k$ is strictly decreasing on $k$, and by definition $\Delta_k > 0$. Therefore, $\bar{x}(\bar{p}) - \bar{x}(p) > 0$ for all $\bar{p} > p$, concluding the proof of this proposition.

Formulation (2.8) is a non-linear optimization problem which can be extremely difficult to solve, since the objective function in (2.8) may be non-concave. An example could be shown here to demonstrate the absence of concavity. Conventional solution methods may get stuck in poor local maxima. In order to avoid this pitfall, we could employ a path-following technique such as the one presented in Hanson and Martin [HM96]. The
key idea is to define a smooth transition between a concave approximation function and the original non-concave objective function. Then, starting with the well-known global maximum of the approximation, we draw a path along the maxima of the intermediate functions with a small constant step size.

An assumption in this formulation is how we are going to model the costs $c_{ij}(p_j)$, i.e., how the customer lifetime pdf $f(t|p_j)$ will change while the prices change.

In the following subsection we formulate the problem considering the whole contract period. Since we are considering that there is a discount factor for cash flows along the contract horizon and that there are customers who turn over before the end of the contract, it is more realistic to maximize profits considering all periods at the same time.

### 2.3.2 Multi-period profit model

The following formulation maximizes the profits obtained from the contracts signed at a specific period of time, i.e., all contracts which start operating at time $t = 0$ and expire at time $t = T$. We do not consider previous contracts nor future ones that will be signed during the next $T$ periods.

Let $N_i$ be the number of potential contracts to be signed by customers from segment $i$ to start at $t = 0$. Then, considering customers’ reservation price distribution, their choices among all different bundles, and a vector of prices, the expected number of contracts to be signed by customer segment $i$ for bundle $j$ is given by

$$\hat{N}_{ij0} = N_i \Pr(\text{buy } j). \quad (2.11)$$

We know that some customers will turn over in each period. Let $\hat{N}_{ijt}$ be the number of remaining customers from segment $i$ in period $t$ who signed a contract for bundle $j$. Using the customer lifetime cumulative distribution $F_i(t)$ and the initial number of signed contracts in (2.11), we can easily get

$$\hat{N}_{ijt} = \hat{N}_{ij0}(1 - F_i(t)). \quad (2.12)$$

Note that $\hat{N}_{ijt}$ is nonincreasing in $t$.

The total revenue to be collected during the contract time from customers from segment $i$ who buy bundle $j$ in today’s money is given by

$$TR_{ij} = \sum_{t=1}^{T} p_j \hat{N}_{ijt}(1 + r)^{-t} \quad (2.13)$$
and the corresponding service \((SC_{ij})\) and setup \((FC_{ij})\) costs are given by

\[
SC_{ij} = \sum_{t=1}^{T} S_j \hat{N}_{ijt}(1 + r)^{-t} \tag{2.14}
\]
\[
FC_{ij} = K_j \hat{N}_{ij0}. \tag{2.15}
\]

Therefore, the profit from customer segment \(i\) for bundle \(j\) is \(TR_{ij} - SC_{ij} - FC_{ij}\). Summing over all bundles and over all segments in order to come up with the overall total profits \(\Pi\), we get

\[
\Pi = \sum_{i=1}^{S} \sum_{j=1}^{B} (TR_{ij} - SC_{ij} - FC_{ij}).
\]

After replacing (2.12), (2.13), (2.14) and (2.15), and rearranging some terms, we get

\[
\Pi = \sum_{i=1}^{S} \sum_{j=1}^{B} \left[ \sum_{t=1}^{T} (p_j - S_j) \hat{N}_{ij0} (1 - F_i(t))(1 + r)^{-t} - K_j \hat{N}_{ij0} \right].
\]

Finally, replacing (2.11) into (2.16) we can formulate the following optimization problem:

\[
\text{P) } \max_{p \geq 0} \sum_{i=1}^{S} \sum_{j=1}^{B} N_i \Pr(\text{buy } j) \left[ \sum_{t=1}^{T} \frac{(p_j - S_j)(1 - F_i(t))}{(1 + r)^t} - K_j \right]. \tag{2.17}
\]

The term between brackets in (2.17) is the customer’s expected profit from segment \(i\) buying a bundle \(j\) which we denote by

\[
V_{ij}(p_j) = \sum_{t=1}^{T} \frac{(p_j - S_j)(1 - F_i(t))}{(1 + r)^t} - K_j. \tag{2.18}
\]

### 2.3.3 Alternative way to model the per-customer expected profit

In order to simplify the formulation in (2.17) or perhaps to come up with a more mathematically tractable expression, we propose an alternative way of modeling the per-customer expected profit \(V_{ij}\).
Let \( \tau_i \) be a random variable representing the customer segment \( i \)'s lifetime, which has a probability distribution \( f(t) \) as mentioned before. The profit obtained from any particular customer from segment \( i \) who chooses bundle \( j \) which is also a random variable would be given by

\[
\text{Profit}_{ij}(p_j) = \sum_{t=1}^{\tau_i} (p_j - S_j) \rho^t - K_j
\]

\[
= (p_j - S_j) \rho \frac{1 - \rho^\tau_i}{1 - \rho} - K_j
\]

where \( \rho = (1 + r)^{-1} \) is the discount factor. Taking expectation to this expression with respect to \( \tau_i \), we get

\[
V_{ij}(p_j) = E[\text{Profit}_{ij}(p_j)]
\]

\[
= E \left[ (p_j - S_j) \rho \frac{1 - \rho^\tau_i}{1 - \rho} - K_j \right]
\]

\[
= (p_j - S_j) \rho \frac{1 - E[\rho^\tau_i]}{1 - \rho} - K_j. \tag{2.19}
\]

As we can notice \( E[\rho^\tau_i] \) is the moment-generating function of the customer lifetime pdf, i.e., the s-transform\(^2\) or z-transform depending upon how we model the lifetime (as a continuous or discrete random variable). In this case we model it as a discrete random variable which can take values from 0 to \( T \) with probabilities \( p_{\tau_i}(\cdot) \) where \( p_{\tau_i}(\tilde{\tau}_i) \) is the probability that a customer stays in the system an amount of time in between \( \tilde{\tau}_i \) and \( \tilde{\tau}_i + 1 \).

**Proposition 5** Both approaches to compute the per-customer expected profit in (2.18) and (2.19) are equivalent, i.e.,

\[
\sum_{t=1}^{T} (p_j - S_j) \rho^t (1 - F_i(t)) - K_j = (p_j - S_j) \rho \frac{1 - E[\rho^\tau_i]}{1 - \rho} - K_j. \tag{2.20}
\]

**Proof.** For the sake of simplicity we do not use subindices in this proof. First, we add \( K \) and divide by \( (p - S) \) on both sides of (2.20). We need to prove that \( V_1 = V_2 \), where \( V_1 = \sum_{t=1}^{T} \rho^t (1 - F_i(t)) \) and \( V_2 = \frac{\rho}{1 - \rho} (1 - E[\rho^\tau_i]). \) By working first with \( V_1 \) and

\(^2\)In order to model the lifetime as a continuous random variable we just need to change \( \rho \) with \( e^{-s} \) to come up with the s-transform of \( f(t) \).
expanding it, we obtain

\[ V_1 = \sum_{t=1}^{T} \rho^t - \sum_{t=1}^{T} \rho^t \left( \sum_{t=0}^{t-1} p_r(t_0) \right) \]

\[ = \frac{\rho - \rho^{T+1}}{1 - \rho} - \left[ \rho p_r(0) + \rho^2 (p_r(0) + p_r(1)) + \ldots + \rho^T (p_r(0) + \ldots + p_r(T-1)) \right] \]

\[ = \frac{\rho - \rho^{T+1}}{1 - \rho} - \left[ (\rho + \rho^2 + \ldots + \rho^T) p_r(0) + (\rho^2 + \ldots + \rho^T) p_r(1) + \ldots + \rho^T p_r(T-1) \right] \]

\[ = \frac{\rho - \rho^{T+1}}{1 - \rho} - \left[ p_r(0) \frac{\rho - \rho^{T+1}}{1 - \rho} + p_r(1) \frac{\rho^2 - \rho^{T+1}}{1 - \rho} + \ldots + p_r(T-1) \frac{\rho^T - \rho^{T+1}}{1 - \rho} \right] \]

\[ = \frac{\rho - \rho^{T+1}}{1 - \rho} - \frac{\rho}{1 - \rho} \left( \sum_{t=0}^{T-1} \rho^t p_r(t) - \rho^T \sum_{t=0}^{T-1} p_r(t) \right). \tag{2.21} \]

On the other hand, by expanding \( V_2 \), we get

\[ V_2 = \frac{\rho}{1 - \rho} - \frac{\rho}{1 - \rho} \left( p_r(0) + \rho p_r(1) + \rho^2 p_r(2) + \ldots + \rho^T p_r(T) \right) \]

\[ = \frac{\rho}{1 - \rho} - \frac{\rho}{1 - \rho} \left( \sum_{t=0}^{T-1} \rho^t p_r(t) + \rho^T \left( 1 - \sum_{t=0}^{T-1} p_r(t) \right) \right) \]

\[ = \frac{\rho - \rho^{T+1}}{1 - \rho} - \frac{\rho}{1 - \rho} \left( \sum_{t=0}^{T-1} \rho^t p_r(t) - \rho^T \sum_{t=0}^{T-1} p_r(t) \right), \]

which is identical to the expression for \( V_1 \) shown in (2.21).

\[ \square \]

2.3.4 Geometric customer behavior

Let us assume that in every period there is a fixed probability \( q_i \) (which strictly increases with \( p_j \)) that a customer from segment \( i \) will leave the system (a Bernoulli process). We only care about customers leaving the system prior to \( T \), the end of the contract, so we concentrate in \( p_{r_i}(T) \) the probability of leaving after \( T \), where \( \tau_i \) is the random variable representing customers’ lifetime.

**Assumption 3** Customer’s lifetime has a truncated geometric probability distribution with parameter \( q_i \), given by

\[ p_{r_i}(\tau) = \begin{cases} 
q_i(1 - q_i)^{\tau-1} & \tau = 1, 2, 3, \ldots, T - 1 \\
(1 - q_i)^{\tau-1} & \tau = T 
0 & \text{otherwise}.
\end{cases} \]

**Remark 6** If \( T \rightarrow \infty \) the above pmf turns into a geometric pmf.
In order to compute the per customer expected profit we need to derive the moment-generating function of \( \tau_i \).

\[
E[\rho^{\tau_i}] = \sum_{k=1}^{T} \rho^k p_{\tau_i}(k)
\]

\[
= \rho q_i + \rho^T (1 - q_i)^T (1 - \rho)
\]

Thus, plugging it into (2.19) we get the following closed form expression for the per customer expected profit to be used in our optimization formulation in (2.17):

\[
V_{ij}(p_j) = \frac{(p_j - S_j)\rho}{1 - \rho} \left( 1 - \frac{\rho q_i + \rho^T (1 - q_i)^T (1 - \rho)}{1 - \rho(1 - q_i)} \right) - K_j
\]

and after replacing \( \rho \) in it we get

\[
V_{ij}(p_j) = \frac{(p_j - S_j)}{r} \left( 1 - \frac{q_i + (1 + r)^{-T} (1 - q_i)^T r}{q_i + r} \right) - K_j
\]

\[
= \frac{p_j - S_j}{q_i + r} \left( 1 - \left( \frac{1 - q_i}{1 + r} \right)^T \right) - K_j.
\] (2.22)

Its corresponding derivative with respect to \( p_j \), in order to reconstruct the first-order conditions developed in the previous subsection, is

\[
V_{ij}' = \frac{\partial V_{ij}}{\partial p_j}
\]

\[
= \frac{q_i + r - (p_j - S_j)q_i'}{(q_i + r)^2} (1 - D) + \frac{(p_j - S_j)Tq_i'}{(q_i + r)(1 - q_i)} D,
\] (2.23)

where \( D = \left( \frac{1 - q_i}{1 + r} \right)^T < 1 \) for all \( T \).

**Assumption 4** Prices can be anything above the service costs, i.e., \( p_j \in [S_j, \infty) \).

**Assumption 5** The parameter \( q_i \in [q_i, \bar{q}_i] \) (probability of leaving the system at any period) has a maximum value \( \bar{q}_i \) which cannot exceed 1. In addition, \( q_i \) is a nondecreasing and concave function of \( p_j \), i.e., \( \frac{\partial q_i}{\partial p_j} \geq 0 \) and \( \frac{\partial^2 q_i}{\partial p_j^2} \leq 0 \) \( \forall p_j \geq S_j \).

The following two lemmas will be helpful in proving that the per customer expected profit is non-decreasing in price.

**Lemma 7** The expression \( M = \frac{(q_i + r)TD}{(1 - q_i)(1 - D)} \) is less than or equal to 1 \( \forall T \geq 1 \) and \( \forall p_j \geq S_j \).
Proof. After replacing $D$ into it and rearranging some terms, we get

$$M = \frac{T(q_i + r)(1 - q_i)^{T-1}}{(1 + r)^T - (1 - q_i)^T}. $$

Now, we proceed by induction. For $T = 1$, we easily prove that

$$M_1 = \frac{q_i + r}{(1 + r) - (1 - q_i)} = 1 \leq 1.$$

Next, we assume (inductive hypothesis) that $M_T \leq 1$ is true, that is,

$$T(q_i + r)(1 - q_i)^{T-1} \leq (1 + r)^T - (1 - q_i)^T, \quad (2.24)$$

thus all we have to do is to prove that $M_{T+1} \leq 1$, that is,

$$(T + 1)(q_i + r)(1 - q_i)^T \leq (1 + r)^{T+1} - (1 - q_i)^{T+1}$$

$$T(q_i + r)(1 - q_i)^{T-1}(1 - q_i) + (q_i + r)(1 - q_i)^T \leq (1 + r)^{T+1} - (1 - q_i)^{T+1}. \quad (2.25)$$

Subtracting (2.24) from (2.25) and rearranging terms, we get

$$(q_i + r)(1 - q_i)^T - T(q_i + r)(1 - q_i)^{T-1}q_i \leq r(1 + r)^T + q_i(1 - q_i)^T$$

$$(q_i + r) - \frac{T(q_i + r)q_i}{(1 - q_i)} \leq r \left( \frac{1 + r}{1 - q_i} \right)^T + q_i$$

$$r \left( 1 - \left( \frac{1 + r}{1 - q_i} \right)^T \right) \leq \frac{T(q_i + r)q_i}{(1 - q_i)} \quad (2.26)$$

which is true because the left hand side of (2.26) is less than zero while its right hand side is positive for all $p_j \geq S_j$. Therefore $0 \leq M \leq 1$. $\blacksquare$

Lemma 8 The derivative of $M$ is non-positive for all $T$ and for all $p_j \geq S_j$.

Proof. We want to prove that

$$\frac{\partial}{\partial p_j} \left( \frac{T(q_i + r)(1 - q_i)^{T-1}}{(1 + r)^T - (1 - q_i)^T} \right) \leq 0$$

43
which after some algebra reduces to
\[
\left( q'_i - \frac{(T-1)(q_i + r)q'_i}{1 - q_i} \right) \left( (1+r)^T - (1-q_i)^T \right) - T(q_i + r)(1-q_i)^{T-1}q'_i \leq 0
\]
\[
((1+r)^T - (1-q_i)^T) - (q_i + r) \left( T(1-q_i)^{T-1} + \frac{T-1}{1-q_i}((1+r)^T - (1-q_i)^T) \right) \leq 0
\]
\[
(1-D) - (q_i + r) \left( \frac{TD}{1-q_i} + \frac{T-1}{1-q_i}(1-D) \right) \leq 0
\]
where \( D = \left( \frac{1-q_i}{1+r} \right)^T \). Dividing by \( (1-D) \) we end up with the following condition:
\[
\left( \frac{T}{1-D} - 1 \right) \geq \frac{1-q_i}{q_i + r}
\]
\[
\frac{T}{1-D} \geq \frac{1+r}{q_i + r}
\]
\[
T \left( \frac{q_i + r}{1+r} \right) + \left( \frac{1-q_i}{1+r} \right)^T \geq 1. \tag{2.27}
\]

We prove by induction that condition (2.27) holds. For \( T = 1 \), it is obvious that (2.27) is true. We assume (inductive hypothesis) that (2.27) is true for \( T \), and we prove it for \( T + 1 \), that is,
\[
(T + 1) \left( \frac{q_i + r}{1+r} \right) + \left( \frac{1-q_i}{1+r} \right)^{T+1} \geq 1.
\]
Subtracting (2.27) from this condition, we get
\[
\frac{q_i + r}{1+r} + \left( \frac{1-q_i}{1+r} \right)^T \left( \frac{1-q_i}{1+r} - 1 \right) \geq 0
\]
\[
\frac{q_i + r}{1+r} - \left( \frac{1-q_i}{1+r} \right)^T \frac{q_i + r}{1+r} \geq 0
\]
\[
\left( \frac{1+r}{1-q_i} \right)^T \geq 1.
\]
which concludes this proof. ■

**Proposition 9** The per customer expected profit in (2.22) is non-decreasing in price, that is,
\[
V'_{ij} \geq 0
\]
Proof. Using (2.23) we establish the above condition as follows:

\[
\frac{q_i + r - (p_j - S_j)q_i'}{(q_i + r)^2} (1 - D) + \frac{(p_j - S_j)Tq_i'}{(q_i + r)(1 - q_i)} D \geq 0
\]

\[
(q_i + r) - (p_j - S_j)q_i' + \frac{(q_i + r)(p_j - S_j)Tq_i'}{(1 - q_i)} \left( \frac{D}{1 - D} \right) \geq 0
\]

\[
(q_i + r) + (p_j - S_j)q_i'(M - 1) \geq 0.
\]

By Lemma 7 and Lemma 8 and assumptions (4) and (5), we can see that the second component of the preceding condition is negative. Because of \((p_j - S_j)q_i'M \geq 0\), it would be enough to show that \((q_i + r) - (p_j - S_j)q_i' \geq 0\), thus avoiding dealing with \(M\) which makes the math more complicated. Let \(\zeta_1(p_j) = q_i + r\) and \(\zeta_2(p_j) = (p_j - S_j)q_i'\), so the first derivative (regarding to \(p_j\)) of these two functions would be \(\zeta_1'(p_j) = q_i'\) and \(\zeta_2'(p_j) = q_i' + (p_j - S_j)q_i''\) respectively. By the fundamental theorem of integral calculus we know that \(\zeta_k(x) = \zeta_k(S_j) + \int_{S_j}^{x} \zeta_k'(p_j) dp_j\) for \(k = 1, 2\), so we can rewrite our condition as follows:

\[
\zeta_1(x) - \zeta_2(x) \geq 0
\]

\[
\zeta_1(S_j) + \int_{S_j}^{x} \zeta_1'(p_j) dp_j - \zeta_2(S_j) - \int_{S_j}^{x} \zeta_2'(p_j) dp_j \geq 0
\]

\[
(q_i + r) + \int_{S_j}^{x} q_i' dp_j - 0 - \int_{S_j}^{x} (q_i' + (p_j - S_j)q_i'') dp_j \geq 0
\]

\[
(q_i + r) \geq \int_{S_j}^{x} (p_j - S_j)q_i''' dp_j
\]

where \(q_i = q_i(S_j)\) and we can easily see that the left hand side is strictly positive while the right hand side is nonpositive (by assumption 5). Therefore \(V_{ij}' \geq 0\). ■

2.4 A one-step choice model approach

In this section we formulate the first bundle pricing optimization model. This formulation is very similar to the one in section 2.5. They only differ in how customers’ purchase behavior is modeled. We first analyze the case when consumers make a decision in a single step, i.e., they decide whether or not to buy and which one to buy (see Figure 2-7) in one step. The underlying demand is captured by a consumer discrete choice model, which we present and discuss in detail in the following section.
2.4.1 Consumer choice model

Consumers are assumed to be random utility maximizers. This assumption, shared by many other models of consumer choice, implies that the preferred option is the one that gives the highest perceived utility. Among all theories about “consumer as a random utility maximizer” the most cited ones are those of Thurstone [Thu27] in psychology, and McFadden [McF74] in economics. They say that consumers see each choice as a bundle of attributes. Then, consumers make an overall evaluation of each choice based on these attributes, by using the so-called utility function. Finally, the choice with the highest overall evaluation is the one chosen by the consumer.

True utility is not fully observable because of consumers’ differences in taste and many other unmeasurable external factors. For this reason the utility is modeled as the sum of a deterministic utility and a stochastic utility$^3$ (these utilities are also called ‘strict utility’ and ‘random utility’ respectively, in McFadden [McF81]). The deterministic utility, denoted by $V(\cdot)$, is computed using all those measurable choice $j$’s attributes which we denote by the vector $X_j$. $V(X_j)$ can be any multivariate function (see [And81], [LM81], [Lyn85], [Lou88]), but it is commonly assumed to be a linear weighted average of all measurable attributes for analytical and interpretation purposes. The stochastic utility is an independent disturbance called $\varepsilon_j$, which reflects the observed tastes of the individual with respect to choice $j$. Thus, we can represent the consumer’s utility by the following random utility model:

$$U_j = V(X_j) + \varepsilon_j,$$

where the index $j$ represents one of the $B$ choices in the consideration set. Note that the consideration set (also called in the literature “evoked set” or “choice set,” see [RL91]) of our problem is an unordered set, i.e., the offered bundles have no specific order.

$^3$Among the underlying sources of the randomness of the utilities are: i) unobserved attributes, ii) unobserved taste variations, iii) measurement errors and imperfect information, and iv) instrumental (or proxy) variables.
As mentioned above, if a consumer chooses bundle \( j \), we assume that \( U_j \) is the maximum utility among the \( B \) utilities (consumers are random utility maximizers). Thus, the probability \( \pi_j(X) \) that bundle \( j \) is chosen (given the consideration set of \( B \) bundles) will be given by

\[
\pi_j(X) = \Pr(U_j > U_k) \quad \forall k \neq j, \quad k \leq B,
\]

which could be rewritten in terms of (2.28) as:

\[
\pi_j(X) = \Pr(\epsilon_k < V(X_j) - V(X_k) + \epsilon_j) \quad \forall k \neq j, \quad k \leq B,
\]

(2.29)

where \( X = [X_1, \ldots, X_B] \) is a matrix of attributes for the \( B \) alternatives in the consideration set.

As Meyer and Kahn [MK90] and Greene [Gre00] point out, this statistical model is made operational by a particular choice of probability distribution for the random disturbances \( \epsilon_j \) and \( \epsilon_k \). Then, by integrating (2.29) over a continuum of possible values of \( \epsilon_j \) we get

\[
\pi_j(X) = \int_{-\infty}^{\infty} \Pr(\epsilon_k < V(X_j) - V(X_k) + \epsilon)dF(\epsilon) \quad \forall k \neq j, \quad k \leq B,
\]

(2.30)

where \( \epsilon \) is a constant of integration and \( F(\epsilon) \) is the assumed cumulative probability distribution of the disturbances (or random utility).

Even though this distribution can take many forms, it is commonly assumed to be either Gaussian or Gumbel for reasons discussed below. If the utility disturbances have a Gaussian distribution (as in Thurstone’s early work on comparative judgement [Thu27]) we end up with the probit model, which is analyzed in Ben-Akiva and Lerman [BAL85]. Researchers based this assumption on the interpretation of the disturbances as the sum of many unobserved but independent components. By the central limit theorem, it can be said that the distribution of the disturbances tends to be Gaussian (or normal). The primary limitation of the normality assumption is that (2.30) cannot be developed in closed-form, and problems with more than two choices in the consideration set have proved to be computationally difficult. In fact, Daganzo [Dag79] provides approximate solutions for the parameters, but their accuracy has been subject of major controversy in the literature, particularly when the model is applied to consideration sets with more than three alternatives (see Horowitz [Hor80] and Wrigley [WLD84]).

If the disturbances \( \epsilon_j \) and \( \epsilon_k \) in (2.29) have independently and identically distributed Gumbel distributions (also called Type I Extreme Value distribution or double exponential distribution), the choice probability in (2.30) has a close-form solution. Under this assumption, we end up with the well-known multinomial (if more than two choices) logit model of choice. The following proposition summarizes this model first formulated by Luce [Luc59], then extended by Theil [The69], and finally formalized by McFadden
Proposition 10  Multinomial Logit Model (MNL). In a world of consumers that are random utility maximizers, where there is a consideration set of $B$ alternatives or choices, where consumers’ utilities for each alternative is given by (2.28), where the utility’s disturbances have independently and identically distributed Gumbel\(^4\) distributions, then, the choice probability for the alternative $j$ in the consideration set will be given by

$$\pi_j(X) = \frac{e^{V(X_j)}}{\sum_{k=1}^{B} e^{V(X_k)}}, \quad \forall j = 1, \ldots, B,$$

which leads to what is called the Multinomial Logit model.

Proof. Starting from (2.30) we can rewrite the probability of choosing bundle $j$ as

$$\pi_j(X) = \int_{-\infty}^{\infty} \left[ \Pr(\epsilon_k < V(X_j) - V(X_k) + \epsilon_j) \right] f(\epsilon_j) d\epsilon_j$$

$$= \int_{-\infty}^{\infty} f(\epsilon_j) \left[ \prod_{k \neq j} \frac{1}{\sqrt{2\pi} \sigma} \exp \left( -\frac{(V_j - V_k + \epsilon_j)^2}{2\sigma^2} \right) \right] d\epsilon_j,$$

where for notation simplicity we call $V_j$ to the deterministic utility $V(X_j)$. Now by recalling

\(^4\)If the random variable $x$ has the following pdf,

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right], \quad \forall -\infty < x < \infty$$

then it is said to have the Gumbel distribution with location parameter $\mu$ ($-\infty < \mu < \infty$) and scale parameter $\sigma$ ($\sigma > 0$). The case where $\mu = 0$ and $\sigma = 1$ is called the standard Gumbel distribution, and the pdf reduces to $f(x) = e^{-x}e^{-e^{-x}}$. For properties of this distribution refer to appendix B.1. Note that the Gumbel distribution is related to the Weibull distribution in a logarithmic way.
the standard Gumbel distribution we can write $\pi_j(X)$ as

$$
\pi_j(X) = \int_{-\infty}^{\infty} e^{-\epsilon_j} e^{-e^{-\epsilon_j}} \left[ \prod_{k \neq j} e^{-e^{-\epsilon_j - V_j - V_k + \epsilon_j}} \right] d\epsilon_j
$$

By performing a change of variables, where $\beta = (\sum_{k=1}^{B} e^{V_k}) / e^{V_j}$, and $-e^{-\epsilon_j} = -\frac{z_j}{V_j}$, we derive that $z_j = \epsilon_j - \ln \beta$ and get the following expression:

$$
\pi_j(X) = \int_{-\infty}^{\infty} e^{-z_j + \ln \beta} e^{-e^{-z_j}} dz_j
$$

$$
= e^{-\ln \beta} \int_{-\infty}^{\infty} e^{-z_j} e^{-e^{-z_j}} dz_j
$$

$$
= \frac{1}{e^{\ln \beta}} = \frac{1}{\beta}.
$$

Since the integral in the above expression is equal to 1 (integral of Gumbel pdf), the probability of choosing bundle $j$ is given by

$$
\pi_j(X) = \frac{e^{V_j}}{\sum_{k=1}^{B} e^{V_k}},
$$

which concludes the MNL model proof.

Although many other assumptions about the distribution of the disturbances have been made, Meyer and Kahn [MK90] point out that none of them yields probability expressions, which are as computationally tractable as the multinomial logit. They say that because of its simplicity, it has emerged as the most widely used form of individual
choice model\(^5\). The choice of the transportation system to get to work (drive, bus, subway, walk, bicycle), housing (buy house, buy condominium, rent), primary shopping location (downtown, mall A, mall B, other), brand of toothpaste, political party preference, choices of packaged goods, or occupation have all been described by using the Logit model. In general, the model provides a good account of the relationship between the attributes of sets of alternatives and the choices made from these sets.

An important aspect of the MNL model that has been extensively discussed in the literature is the “constant ratio rule” or the “independence of irrelevant alternatives” (IIA) property as called by Luce [Luc59]\(^6\). We can easily verify in the MNL model that the ratio of the choice probabilities of any two alternatives is not affected by the deterministic or strict utilities of any other alternatives, i.e.,

\[
\frac{\pi_j}{\pi_k} = \frac{e^{V_j} / \sum_{k=1}^{B} e^{V_k}}{e^{V_k} / \sum_{k=1}^{B} e^{V_k}} = \frac{e^{V_j}}{e^{V_k}}.
\]

The main implication is that if a new option is added to a choice set, the shares of existing options will always decrease in direct proportion to the size of their original shares [MK90]. If we add a new alternative essentially identical to the \(k^{th}\) alternative, the new alternative might reasonably be expected to split \(k^{th}\)’s probability and leave the others untouched. However, Proposition 10 asserts that the probabilities of all alternatives will be reduced. Counter-examples such as the red bus/blue bus paradox ([Deb60], [BAL85]) have demonstrated the limitations of the IIA assumption in the real world.

Effective logit models describe choice situations where the negative effects of the IIA property are minimized. In the bundling problem, it is important to ensure that the attributes that distinguish bundles from each other are relevant to the consumer’s evaluation process. As McFadden [McF74] puts it, “the application of the model should be limited to situations where the alternatives can plausibly be assumed to be distinct and weighted independently in the eyes of each decision-maker.” Many schemes that have been proposed for solving this potential difficulty (see McFadden [McF81]) involve arranging the alternatives into a hierarchy that groups similar alternatives, so in order to avoid these known pitfalls it is key to have them in mind when setting up a specific instance of the model.

The strict utility is usually modeled as a linear weighted average of the attributes, where the calibration of the weights is usually done by maximum likelihood (Ben-Akiva

\(^5\)The word model correctly implies a certain degree of approximation and imperfection. The strength of a good model is to abstract the important characteristics of the world that are relevant to the problem at hand, and, just as important, omit details of the world that are not particularly influential or relevant. Tests on the logit model have shown that it describes a surprisingly large number of purchase phenomena quite well.

\(^6\)Luce argues that the IIA property can be viewed as a probabilistic version of the concept of transitivity.
[BA73] provides a good algorithm for doing this) since this method provides parameter estimates which are consistent, asymptotically efficient and normally distributed under very general conditions. Empirical experience indicates that the estimate is quite good even in relatively small samples \((n \geq 100)\). Guadagni and Little [GL83] present a good list of statistics that can be used to measure the quality of fit.

### 2.4.2 The model

As mentioned before, profit maximization is the objective of this model. Recall that \(\alpha_j(p)\) denotes the probability that bundle \(j\) is chosen with \(\alpha_0\), the non-purchase probability, and \(V_j(p)\) denotes the per-customer expected profit when choosing bundle \(j\). Therefore, for this case, the total expected profit function to be maximized in \((2.4)\) is given by

\[
\Pi(p) = \sum_{j=1}^{B} \alpha_j(p)V_j(p),
\]

for the case of a single segment of customers.

The component \(\alpha_j(p)\) corresponds to the underlying demand function for bundle \(j\). As in Hanson and Martin [HM96] and Basuroy and Nguyen [BN98], we assume that the customers’ utility function for bundle \(j\) at price \(p_j\) with vector \(X_j\) of non-price attributes is defined by

\[
U(X_j,p_j) = g(X_j) + h(p_j) + \epsilon_j,
\]

where \(\epsilon_j\)'s are independent and identically distributed Gumbel random variables with variance \(\pi^2/6\mu^2\). Thus, as explained in the previous section, we are modeling the consumer choice behavior with a multinomial logit model (MNL), where the probability that a customer purchases product \(j\) given price vector \(p = (p_1, \ldots, p_B)\) and matrix of non-price attributes \(X = (X_1, \ldots, X_B)\) is given by

\[
\alpha_j(X, p) = \frac{\exp(g(X_j) + h(p_j))}{\sum_{k=0}^{B} \exp(g(X_k) + h(p_k))}.
\]

From Proposition 1, we know that \(\alpha_j = (1 - \alpha_0)\pi_j\), where \(\pi_j\) corresponds to the probability of choosing bundle \(j\) given that at least one product will be purchased. By letting \(R_j(p)\) be equal to \((1 - \alpha_0)V_j(p)\), we can rewrite our optimization model as follows:

\[
\max_{p \geq 0} \Pi(p) = \sum_{j=1}^{B} \pi_j(p)R_j(p).
\]
2.4.3 Analysis of the optimal expected profit

Considering the optimization model in (2.33) we proceed to derive the first-order conditions of the profit model in order to characterize the optimal price vector \( \mathbf{p}^* \). Before doing that, we first need to find the partial derivatives of each choice probability with respect to all prices.

\[
\frac{\partial \pi_j}{\partial p_j} = \frac{\partial}{\partial p_j} \left( e^{U_j} / \sum_{l=1}^{B} e^{U_l} \right) = \left[ U_j e^{U_j} / \sum_{l=1}^{B} e^{U_l} \right] \left[ 1 - e^{U_j} / \sum_{l=1}^{B} e^{U_l} \right] = h'_j \pi_j (1 - \pi_j)
\]

\[
\frac{\partial \pi_k}{\partial p_j} = \frac{\partial}{\partial p_j} \left( e^{U_k} / \sum_{l=1}^{B} e^{U_l} \right) = -U'_j e^{U_j} e^{U_k} \frac{1}{\left( \sum_{l=1}^{B} e^{U_l} \right)^2} = -h'_j \pi_j \pi_k \quad \forall k \neq j,
\]

where \( h_j(p_j) \) is the price response function of the customer’s utility function for product \( j \), that we will assume to be equal to \(-p_j\) for the sake of mathematical clarity. Thus, the derivative of \( h_j(p_j) \) with respect to \( p_j \) will be \( h'_j = -1 \). Now, deriving the total profit with respect to \( p_j \) we get:

\[
\frac{\partial \Pi(p)}{\partial p_j} = \frac{\partial \pi_j}{\partial p_j} R_j + \pi_j \frac{\partial R_j}{\partial p_j} + \sum_{k \neq j} \left( \frac{\partial \pi_k}{\partial p_j} R_k + \pi_k \frac{\partial R_k}{\partial p_j} \right)
\]

\[
= -\pi_j (1 - \pi_j) R_j + \pi_j \frac{\partial R_j}{\partial p_j} + \sum_{k \neq j} \left( \pi_j \pi_k R_k + \pi_k \frac{\partial R_k}{\partial p_j} \right)
\]

\[
= -\pi_j R_j + \pi_j \frac{\partial R_j}{\partial p_j} + \pi_j \sum_{k} \pi_k R_k + \sum_{k \neq j} \pi_k \frac{\partial R_k}{\partial p_j}
\]

\[
= \pi_j \left( \frac{\partial R_j}{\partial p_j} - R_j + \Pi(p) \right) + \sum_{k \neq j} \pi_k \frac{\partial R_k}{\partial p_j}. \quad (2.34)
\]

Before continuing developing these first-order conditions we need to compute \( \frac{\partial R_k}{\partial p_j} \) and \( \frac{\partial \alpha_0}{\partial p_j} \).

\[
\frac{\partial \alpha_0}{\partial p_j} = \frac{\partial}{\partial p_j} \left( \frac{1}{1 + \sum_{k=1}^{B} e^{U_k}} \right)
\]

\[
= \frac{-U'_j e^{U_j}}{\left( 1 + \sum_{k=1}^{B} e^{U_k} \right)^2}
\]

\[
= \alpha_0 \pi_j \quad (1 - \alpha_0)
\]

52
\[
\frac{\partial R_j}{\partial p_j} = \frac{\partial}{\partial p_j} ((1 - \alpha_0)V_j(p_j)) \\
= -\frac{\partial \alpha_0}{\partial p_j} V_j(p_j) + (1 - \alpha_0) \frac{\partial V_j(p_j)}{\partial p_j} \\
= -\alpha_0 \pi_j (1 - \alpha_0)V_j(p_j) + (1 - \alpha_0) \frac{\partial V_j(p_j)}{\partial p_j}.
\]

Now using these result in (2.34) we get
\[
\frac{\partial \Pi(p)}{\partial p_j} = \pi_j \left( \frac{\partial R_i}{\partial p_j} - R_j + \Pi(p) \right) - \frac{\alpha_0}{1 - \alpha_0} \left( \sum_k \pi_k (1 - \alpha_0)V_k - \pi_j (1 - \alpha_0)V_j \right) \\
= \pi_j \frac{\partial R_i}{\partial p_j} + \left( \frac{\alpha_0}{1 - \alpha_0} - 1 \right) \pi_j R_j + \left( \pi_j - \frac{\alpha_0}{1 - \alpha_0} \right) \Pi(p) \\
= \pi_j \frac{\partial R_i}{\partial p_j} + (\alpha_0 \pi_j - 1) \pi_j R_j + \pi_j (1 - \alpha_0) \Pi(p) \\
= \pi_j \left( \frac{\partial R_i}{\partial p_j} + (\alpha_0 \pi_j - 1) R_j + (1 - \alpha_0) \Pi(p) \right).
\]

Therefore, the first-order condition suggests that either \( \pi_j = 0 \) or the expression between the parenthesis is equal to zero. Thus, for all active bundles \( (\pi_j \neq 0) \) the following first-order conditions must hold:
\[
\frac{\partial R_j}{\partial p_j} + (\alpha_0 \pi_j - 1) R_j + (1 - \alpha_0) \Pi(p) = 0 \\
-\alpha_0 \pi_j (1 - \alpha_0)V_j + (1 - \alpha_0) \frac{\partial V_j}{\partial p_j} + (\alpha_0 \pi_j - 1) (1 - \alpha_0)V_j + (1 - \alpha_0) \Pi(p) = 0,
\]
which can be reduced to
\[
V_j - \frac{\partial V_j}{\partial p_j} = \Pi(p). \tag{2.35}
\]

This result is very surprising. It states that in the optimum all bundles will have the same value for their corresponding \( V_j - \frac{\partial V_j}{\partial p_j} \) function, and this value is equal to the optimal profit. That means that once one bundle’s price is determined, all the others are also determined. This problem can be seen as a unidimensional pricing problem (see Figure 2-8).

For completeness, it is worth mentioning a technique to solve non-linear functions with logit probabilities proposed by Hanson and Martin [HM96] based on the work of Garcia and Zangwill [GZ81]. Basically they characterize the optimization problem in terms of a parameter which, roughly speaking, corresponds to the amount of randomness in consumer choice. Since the original problem is hard to solve (not concave), it can be perturbed into an easy (concave) problem by artificially making the choices more random and less responsive to product attributes and prices. They develop a new solution technique for these problems based on a path-following procedure that does not require global concavity. It works by creating a path of solutions between an easy to solve model (concave due to a perturbation) and the difficult-to-solve true model (nonconcave). Figure 2-9 shows how our optimization problem in (2.33) is being perturbed while changing the parameter \( t \).
Figure 2-8: Optimal prices determined by the first order conditions

Figure 2-9: Path-following procedure in a 2-bundle case

2.5 A two-step choice model approach

The model we present in this section is basically the same structure as the previous one, but it differs in the way we capture the consumer behavior. Here consumers arrive to the store (or web site) with a predefined choice of which bundle they will buy. That choice will only change if the price they face at the arrival time is out of an acceptable range, i.e., small changes in prices do not affect customer choices. Then, customers decide whether or not to buy the bundle based on their reservation prices (maximum amount of money they are willing to pay).

Before we present this model, we introduce the Assimilation-Contrast Theory that supports this type of customer behavior.
2.5.1 Assimilation-contrast theory

This is a theory closely identified with Sherif and Hovland [SH61]. They hypothesize a price range internal to consumers called “latitude of acceptance” that is a range of acceptable prices. According to the theory, if a consumer sees a brand’s price that is within the latitude of acceptance, the price is assimilated into the range and becomes acceptable. It also becomes less noticeable relative to prices outside the range. A price that is outside the range is contrasted to the acceptable range and becomes noticeable (see Figure 2-10 adapted from [SH61]).

![Assimilation-Contrast Theory Diagram](image)

Figure 2-10: Assimilation-Contrast Theory

Within the latitude of acceptance, there is no price difference on the reference price (or what Sherif and Hovland [SH61] term “the median judgment of the stimulus”).

For prices outside the range, the reference price is affected in approximately a linear fashion. In other words, assimilation-contrast theory predicts roughly cubic relationships between observed and reference price with the flat spot occurring in the latitude of acceptance. Again, as with the other theories, assimilation-contrast theory specifies a relationship between the observed stimuli, market prices, and the internal reference prices.

2.5.2 The model

Applying the above theory to our case, the purchase probability will not be affected by changes in the latitude of acceptance of the bundle price (see Figure 2-11). These kind of demand functions are quite difficult to manipulate because they are not differentiable for every price. For that reason we look for an approximation of these choice probabilities.

As in the previous section, we use the multinomial logit model to model these choice probabilities for the bundles. This model has very nice properties, and it is very attractive
for its mathematical tractability. Strengths and drawbacks of the multinomial logit model in this bundling setting are widely discussed in section 2.4.1. Recall from the introduction of this chapter that the choice probabilities for this two-step model are given by

$$\pi_j(p) = \frac{e^{(g(X_j)+h(p_j))}}{\sum_{k=1}^{B} e^{(g(X_k)+h(p_k))}} \quad \forall j = 1, ..., B, \quad (2.36)$$

where $g(X_j)$ is a function of the non-price attributes of bundle $j$, and $h(p_j)$ is the price response function such as $\frac{\partial h(p_j)}{\partial p_j} < 0$.

On the other hand, within a customer segment, customers have different reservation prices for each of the offered bundles.

**Definition 11** $\Psi_j(\cdot)$ is the cumulative distribution function (cmf) of the reservation prices for bundle $j$, defined as $(1 - \Psi_j(p_j)) = \int_{p_j}^{\infty} \psi_j(r)dr$, where $\psi_j(\cdot)$ is the probability density function of the customers’ reservation prices for bundle $j$.

The expression in Definition 11 allows us to compute the proportion of customers willing to pay a given bundle price, which will be used in the profit function as the probability that a customer will purchase a given bundle. Figure 2-12 illustrates the decision process.

As we have seen, the objective function is composed of three major expressions. These are the proportion of customers willing to pay a certain price, the probability that such customers choose a certain product, and the per-customer expected profit. Hence, using (2.36) and Definition 11 we can rewrite our optimization model as follows:

$$\begin{align*}
\max_{p \geq 0} & \sum_{i=1}^{S} \sum_{j=1}^{B} \pi_{ij}(X_i, p) R_{ij}(p_j),
\end{align*}$$

\[ (2.37) \]
Figure 2-12: Customer behavior in a 2-step decision model

where \( R_{ij}(p_j) = N_i(1 - \Psi_{ij}(p_j))V_{ij}(p_j) \) corresponds to the total expected profit for product \( j \) in customer segment \( i \). Note that \( R_{ij}(p_j) \) is function of \( p_j \) only, exactly what makes this formulation different from the previous one (1-step model).

### 2.5.3 Analysis of the optimal expected profit

Let us start by characterizing the function \( R_{ij}(p_j) \) and looking at the implications it has in the convex combination functional form of (2.37).

**Definition 12** Function \( f(x) \in \mathbb{R} \) is called **quasi-concave** if for any two points \( x \neq y \) and for all \( 0 \leq \alpha \leq 1 \) the following condition holds:

\[
\min\{f(x), f(y)\} \leq f(\alpha x + (1 - \alpha)y).
\]

It is clear that \( R_{ij}(p_j) \) has a maximum value because \((1 - \Psi_{ij}(p_j))\) is a strictly decreasing function and \( V_{ij}(p_j) \) is a strictly increasing function (see Proposition 9) and both are always positive and reach zero at some point. There is a lot of empirical evidence that \( R_{ij}(p_j) \) turns out to be concave, but in order to guarantee a unique maximum we only need it to be quasi-concave, i.e., to have a positive slope for all prices less than the price that maximizes \( R_{ij}(p_j) \) and a negative slope otherwise.

**Assumption 6** Within the range of problems we face in reality, parameters are such that either \( R_{ij}(p_j) \) is quasi-concave, i.e., \((1 - \Psi(p_j))V'(p_j) - \varphi(p_j)V(p_j) \) is always positive until it reaches zero and then it is always negative, or \( R_{ij}(p_j) \) is concave, i.e., \((1 - \Psi(p_j))V''(p_j) - \varphi'(p_j)V(p_j) \leq 2\varphi(p_j)V'(p_j) \). The concavity of \( R_{ij}(p_j) \) strongly depends upon the reservation price distribution.

For the sake of simplicity we first analyze the case of a single customer-segment. Our problem in (2.37) reduces to a single convex combination among the total expected profit
for each product, so for simplicity we write Π = Σ^B_j=1 π_j R_j. In order to explore the first-order conditions of the objective function, we find the partial derivatives of each choice probability with respect to all prices.

\[ \frac{\partial \pi_j}{\partial p_j} = \frac{\partial}{\partial p_j} \left( e^{U_j / \sum B_{l=1} e^{U_l}} \right) = \frac{\left[ U'_j e^{U_j} / \sum B_{l=1} e^{U_l} \right] \left[ 1 - e^{U_j} / \sum B_{l=1} e^{U_l} \right]}{\left( \sum B_{l=1} e^{U_l} \right)^2} = h'_j \pi_j (1 - \pi_j) \]

where \( h_j(p_j) \) is the price-dependent component of the customer’s utility function for product \( j \). Now we compute the derivatives of the total profit \( \Pi \) with respect to all prices, establishing the first-order conditions of the optimal solution to this problem.

\[ \frac{\partial \Pi(p)}{\partial p_j} = \pi_j \frac{\partial R_j}{\partial p_j} + \sum_{k=1}^B \frac{\partial \pi_k}{\partial p_j} R_k \]

\[ = \pi_j \frac{\partial R_j}{\partial p_j} + h'_j \pi_j (1 - \pi_j) R_j - \sum_{k \neq j} h'_j \pi_j \pi_k R_k \]

\[ = \pi_j \frac{\partial R_j}{\partial p_j} + h'_j \pi_j R_j - h'_j \pi_j \sum_{k=1}^B \pi_k R_k \]

\[ = \pi_j h'_j \left( \frac{1}{h'_j} \frac{\partial R_j}{\partial p_j} + R_j - \sum_{k=1}^B \pi_k R_k \right). \]

For the above condition to hold we need \( h'_j \neq 0 \) for all products. This is a very reasonable assumption since customers’ utility functions are most likely to be function of prices. Thus, the set of optimal prices \( p^* \) is the solution to the following system of equations:

\[ \pi_j \left( \frac{1}{h'_j} \frac{\partial R_j}{\partial p_j} + R_j - \Pi(p) \right) = 0 \quad \forall j \quad (2.38) \]

If we know the products with \( \pi_j = 0 \), we would just make \( p_j = 0 \) or just remove those variables from the system of equations in order to have the same number of unknowns and equations. Without lost of generality prices could be normalized such that utility functions depend on price in an inverse one-to-one way, i.e., \( h_j(p_j) = -p_j \), thus, conditions in (2.38) reduce to

\[ \frac{\partial R_j}{\partial p_j} = R_j - \Pi(p) \quad \forall j : \pi_j \neq 0. \quad (2.39) \]

From (2.39) we can clearly see that \( \frac{\partial R_j}{\partial p_j} \) will be greater than zero for some products and less than zero for others since \( \Pi(p) \) is a convex combination of the \( R_j \)s (see Figure 2-13).
Proposition 13  The objective function in (2.37) is not concave.

Proof. We start this proof by computing the second derivatives of the objective function in order to analyze the associated Hessian matrix.

\[
\frac{\partial^2 \Pi(p)}{\partial p_j^2} = \frac{\partial}{\partial p_j} \left( \pi_j R_j - \pi_j \frac{\partial R_j}{\partial p_j} - \pi_j \Pi(p) \right)
\]
\[
= \frac{\partial \pi_j}{\partial p_j} \left( R_j - \frac{\partial R_j}{\partial p_j} - \Pi(p) \right) + \pi_j \left( \frac{\partial R_j}{\partial p_j} - \frac{\partial^2 R_j}{\partial p_j^2} - \frac{\partial \Pi(p)}{\partial p_j} \right)
\]
\[
= \pi_j \left( \frac{\partial R_j}{\partial p_j} - \frac{\partial^2 R_j}{\partial p_j^2} \right).
\]

Then the cross derivatives for every \( k \neq j \) are given by:

\[
\frac{\partial^2 \Pi(p)}{\partial p_j p_k} = \frac{\partial}{\partial p_j} \left( \pi_k R_k - \pi_k \frac{\partial R_k}{\partial p_k} - \pi_k \Pi(p) \right)
\]
\[
= \frac{\partial \pi_k}{\partial p_j} \left( R_k - \frac{\partial R_k}{\partial p_k} - \Pi(p) \right) + \pi_k \frac{\partial \Pi(p)}{\partial p_j}
\]
\[
= 0.
\]

As we can see above, the Hessian matrix \( H \) is a diagonal matrix, so the determinant of it is the product of all its components in the diagonal.

\[
H = \begin{bmatrix}
\pi_1(R'_1 - R''_1) & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \pi_B(R'_B - R''_B)
\end{bmatrix}
\]

In order to have a negative semidefinite Hessian matrix we need negative odd-index subdeterminants and positive even-index subdeterminants, i.e., \( \Delta_1 = \pi_B(R'_B - R''_B) \leq 0 \),
\[ \Delta_2 = \pi_{B-1} (R'_{B-1} - R''_{B-1}) \Delta_1 \geq 0, \quad \Delta_3 = \pi_{B-2} (R'_{B-2} - R''_{B-2}) \Delta_2 \leq 0, \] and so on and so forth. For this condition to hold we would need \( \pi_j (R'_j - R''_j) \leq 0 \) for every \( j = 1, \ldots, B \), which is not true. Therefore, the objective function is not concave.

In what follows we present bounds for the total expected profit given by the first-order conditions in (2.39). As we see in Proposition 13, we cannot guarantee that a feasible set of prices that satisfies the first-order conditions will be the optimal solution, thus these bounds are extremely helpful in determining how good a local optimum can be.

**Proposition 14** The following expression is an upper bound of the total expected profit \( \Pi(p^*) \).

\[
UB_\Pi = \min_{j=1, \ldots, B} \{ \max R_j - \frac{\partial R_j}{\partial p_j} \}
\]

**Proof.** This proof is straightforward. In optimality, necessary conditions in (2.39) show that \( R_j - \frac{\partial R_j}{\partial p_j} \) are all equal to the optimal objective value \( \Pi(p^*) \). It means that \( \Pi(p^*) \) has to be less than or equal to the maximum value of all functions \( R_j - \frac{\partial R_j}{\partial p_j} \). Therefore the minimum of all these maximum values corresponds to an upper bound for the total optimal profit (see Figure 2-14 for a graphical interpretation of \( UB_\Pi \)).

![Figure 2-14: Total profit upper bound for the 3-product example](image)

Before continuing we define some extra notation. Let \( k \) be the product’s index that set the upper bound \( UB_\Pi \), i.e., \( k = \arg \min_{j=1, \ldots, B} \{ \max R_j - \frac{\partial R_j}{\partial p_j} \} \). Let also \( \hat{p}_k \) be the price that maximizes \( R_k - \frac{\partial R_k}{\partial p_k} \), thus \( UB_\Pi = R_k (\hat{p}_k) - \frac{\partial R_k(\hat{p}_k)}{\partial p_k} \). Then, for every \( j \neq k \) there will be two solutions to the equation \( R_j - \frac{\partial R_j}{\partial p_j} = UB_\Pi \), which we denote by \( \hat{p}_j^L \) and \( \hat{p}_j^R \), left and right respectively (see Figure 2-14).

As we can see, any price vector \( \hat{p} \) built out of any combination of \( \hat{p}_1^L \) or \( \hat{p}_1^R, \ldots, \hat{p}_k, \ldots, \hat{p}_B^L \) or \( \hat{p}_B^R \) makes all \( R_j - \frac{\partial R_j}{\partial p_j} \) to reach the value \( UB_\Pi \) but by any means it corresponds to the optimal solution, because it does not fulfill the first-order conditions in (2.39). Without loss of generality let us pick any of the previous \( \hat{p} \) for our next proposition.
Proposition 15 A lower bound for the total profits is given by

\[ LB_\Pi = \Pi(\hat{p}) = UB_\Pi + \sum_{j=1}^{B} \pi_j(\hat{p}) \frac{\partial R_j(\hat{p}_j)}{\partial p_j}, \]

where \( \hat{p} \) is a price vector that fulfills the conditions \( R_j - \frac{\partial R_j}{\partial p_j} = UB_\Pi \) \( \forall j. \)

Proof. This proof is also very straightforward. It is only about replacing into the profit function at price levels \( \hat{p} \) the respective conditions in (2.39).

\[ \Pi(\hat{p}) = \sum_{j=1}^{B} \pi_j(\hat{p}) R_j(\hat{p}_j) = \sum_{j=1}^{B} \pi_j(\hat{p}) \left( UB_\Pi + \frac{\partial R_j(\hat{p}_j)}{\partial p_j} \right) = UB_\Pi + \sum_{j=1}^{B} \pi_j(\hat{p}) \frac{\partial R_j(\hat{p}_j)}{\partial p_j}. \]

Thus, we conclude this proof.

From the previous proof we can see that the gap in between these two bounds \( (UB_\Pi - LB_\Pi) \) is given by

\[ gap = - \sum_{j=1}^{B} \pi_j(\hat{p}) \frac{\partial R_j(\hat{p}_j)}{\partial p_j}, \]

which has to be non-negative to avoid infeasibility. Since all probabilities \( \pi_j \) are always non-negative, then at least one \( \frac{\partial R_k}{\partial p_k} \) has to be negative in order to be able to have a positive gap. It turns out that \( \frac{\partial R_k}{\partial p_k} \) is always negative (\( k \) is the product’s index that set the upper bound) because the first-order condition for \( R_k - \frac{\partial R_k}{\partial p_k} = UB_\Pi \) is given by \( \frac{\partial R_k}{\partial p_k} - \frac{\partial^2 R_k}{\partial p_k^2} = 0 \), and by assumption 6, \( \frac{\partial^2 R_k}{\partial p_k^2} < 0 \) then \( \frac{\partial R_k}{\partial p_k} < 0 \). For instance, observe product 2 in Figure 2-14 for a graphical interpretation of the previous result. In particular, if \( gap = 0 \) then \( \hat{p} \) is the optimal set of prices.

An experiment with the above bounds shows their efficiency. The gap between these bounds was always below 0.1%, which is really good. Figure 2-15 depicts the vector of optimal prices (\( \text{Price}_i \)) and the vector of prices suggested by the lower bound (\( \hat{p} \)) for different contract lengths. It is clear that the longer the contract the more accurate the bound is in suggested near optimal prices.

Returning to the general case in (2.37) where multiple segments of customers are considered, the following set of first-order conditions can be specified, assuming homogeneous-
normalized product utility functions across segments, i.e., \( \pi_{1j} = \ldots = \pi_{Sj} \) for all products.

\[
\sum_{i=1}^{S} R_{ij} - \sum_{i=1}^{S} \frac{\partial R_{ij}}{\partial p_j} = \Pi(p) \quad \forall j = 1, \ldots, B.
\] (2.40)

Since summation of concave functions preserves this property, then, by using assumption 6, the left hand side of conditions in (2.40) are concave functions in \( p_j \) respectively. Therefore, the same procedure used for the single-segment case to compute the upper and lower bounds can be used in this multi-segment case.

If we consider heterogeneous customer’s utility functions across segments the first-order conditions turn out to be much more complicated and not very insightful, so other mechanisms should be explored to come up with good bounds.

### 2.6 A model for a dynamic system

In this section we address the issue of customer base evolution in a system with contracts like the ones considered in this chapter. In a nutshell, we control the customer base by setting prices at different levels, depending upon the seller’s objectives. We analyze the case where a single bundle is offered and prices can fluctuate up and down from period to period. We consider different approaches which provide us with the whole picture of the system behavior. Figure 2-31 in the extensions of this chapter illustrates the customer base throughout the time.
2.6.1 Single bundle model formulation

In the following stochastic and dynamic programming formulation we consider the stationary discrete-time dynamic system

$$N_{k+1} = l_k(p_k, N_k) + w_k(p_k)$$ \hspace{1cm} (2.41)

where for all period $k$, the state $N_k$ (number of old customers at the beginning of period $k$) is an element of a space $S$, the control $p_k$ (price at period $k$) is an element of a space $C$ and depends on the current state $N_k$, and the random disturbances $l_k$ (old customers who continue in the system during period $k$) and $w_k$ (customer demand in period $k$) are elements of space $D$ which we assume to be a countable set. The old and new customer demand is characterized by probabilities $\Pr(l_k, w_k | N_k, p_k)$ defined on $D$, which represent the probability of occurrence of $l_k$ and $w_k$ when the current state and control are $N_k$ and $p_k$ respectively. The function $q(p_k, N_k)$ represents the probability that any particular old customer will continue in the system for at least one more period when the price is $p_k$ and there are $N_k$ customers in the system at the beginning of period $k$.

In period $k$ there are $N_k$ customers who pay $p_k$ each one with probability $q(p_k, N_k)$, i.e., the random variable $l_k$ has a Binomial distribution with parameters $N_k$ and $q(p_k, N_k)$. There are $w_k$ new customers who cost to the company an initial setup cost of $K$ each one.

The random variable $w_k$ follows a Poisson distribution with rate $\lambda(p_k) = \lambda(1 - F(p_k))$, where $F(p_k)$ is the reservation price cumulative probability distribution, and $\lambda$ is the arrival rate of potential customers. Thus, the net profits of period $k$ is a function $g : S \times C \times D \rightarrow \mathbb{R}$ given by

$$g(N_k, p_k, l_k, w_k) = l_k p_k - w_k K.$$

Given an initial state $N_0$ our objective is to find a policy $\pi = \{\mu_0, \mu_1, \ldots\}$, where $\mu_k(N_k)$ is a function $\mu_k : S \rightarrow C$, such that maximizes the total expected profit function

$$J_\pi(N_0) = \lim_{T \rightarrow \infty} E\left(\sum_{k=0}^{T-1} \rho^k g(N_k, p_k, l_k, w_k)\right)$$

subject to the dynamics of the system given by (2.41). The parameter $\rho$ represents the discount factor.

Let $J_k(N_k)$ be the maximum discounted future expected profit if there are $N_k$ customers at the beginning of period $k$. Thus, the expected profits at period $k$ are given by the immediate profits defined by $g(N_k, p_k, l_k, w_k)$ plus the discounted profits from period $k + 1$ onwards, i.e.,

$$J_k(N_k) = \max_{p_k} E \left\{ g(N_k, p_k, l_k, w_k) + \rho J_{k+1}(N_{k+1}) \right\}$$ \hspace{1cm} (2.42)
which corresponds to the respective dynamic programming (DP) algorithm.

Taking expectation in (2.42) we have

\[
J_k(N_k) = \max_{p_k} \left[ \bar{l}_k p_k - \bar{w}_k K + \rho \sum_{j=0}^{N_k} \Pr(l_k + w_k = j) J_{k+1}(j) \right],
\]

(2.43)

where \(\bar{l}_k = q(p_k, N_k) N_k\), \(\bar{w}_k = \lambda(p_k)\), and \(N\) is a large number such that the number of customers that can be in the system at any particular period never will be larger\(^7\).

The optimal pricing policy to be derived from (2.43) will depend upon the discount factor level. If future profit does not matter at all, i.e., \(\rho = 0\), we face a “last period problem,” while if it fully matters, i.e., \(\rho = 1\), the value function will go to infinity at any price level. In this case, it will be more suitable to solve a problem which maximizes profit per unit of time as it is the case of an \(M/M/\infty\) queueing system (see Figure 2-16).

![Figure 2-16: Changes in the discount factor \(\rho\)](image)

**Remark 16** An interesting result we have observed in this formulation is that for a discount factor smaller than a certain threshold \(\hat{\rho}\), the optimal pricing policy is counter-intuitive, i.e., the more customers there are in the system the lower the optimal price. Conversely, for a discount factor greater than \(\hat{\rho}\), we have an intuitive optimal pricing policy.

In the remainder of this subsection examining the dynamic system, we explore first the last period problem. Then, a certainty equivalent formulation is presented and compared with the stochastic one. Computational experiments are performed in order to contrast the optimal pricing policy with a fixed price policy. Finally, the queueing model is presented along with the explanation of how it relates with our discrete-time model.

\(^7\)This is mild approximation, just made for computational experiment tractability. Probabilities around \(N\) are near zero.
2.6.2 Optimal pricing for the last period

If we let the discount factor in (2.43) approach to zero, i.e., $\rho \to 0$, we are in a situation where the future does not matter at all. This is a boundary situation similar to a last period problem. If $T$ denotes the last period, then the value function in (2.43) would be

$$J_T(N_T) = \max_{p_T} [q N_T p_T - \lambda (1 - F(p_T)) K]. \quad (2.44)$$

In order to compute the optimal price let us consider the derivative with respect to $p_T$,

$$\frac{\partial J_T(N_T)}{\partial p_T} = \frac{\partial q}{\partial p_T} N_T p_T + q N_T + \lambda f(p_T) K,$$

and thus, the first-order condition dictates that by making it equal to zero we get

$$p_T^* = - \left( \frac{\partial q}{\partial p_T} \right)^{-1} \left( q + \frac{1}{N_T} \lambda f(p_T^*) K \right), \quad (2.45)$$

which can be solved numerically by finding the fixed point solution (see Figure 2-17, where $H(p_T)$ corresponds to the RHS of (2.45)).

![Figure 2-17: Graphical solution for price $p_T^*(N_T)$ in equation (2.45)](image)

In Figure 2-17 we plot several versions of the function $H(p_T)$ for different values of $N_T$. As can be seen, $H(p_T)$ decreases\(^8\) with $N_T$, and it rapidly converges to $-q \left( \frac{\partial q}{\partial p_T} \right)^{-1}$ which is independent of other parameters such as $\lambda$, $K$, and the distribution of the reservation

---

\(^8\)In this numerical example we have considered $K = 10$, $\lambda = 1$, $F(p) = \frac{p}{10}$, and $q(p) = e^{-p/5} - \frac{1}{5}$.
prices $f(p_T)$. In this example, for a large number of customers, the optimal price converges to the value $p^*_T \to 3.13$.

Note that it is not realistic to acquire new customers in the last period, so by setting $K = 0$ we can observe that the optimal price to charge in this last period would be $p^*_T = -q \left( \frac{\partial q}{\partial p_T} \right)^{-1}$, independent of the number of customers in the system.

What we see here is a counter-intuitive result in a way. The more customers there are in the system, the lower the optimal price. If we look at the profit function in (2.44) we can observe that when there are just a few customers in the system, a high price would be optimal in order to avoid the penalty of the new customer arrivals which cost $K$ each, while suffering the lost of a higher proportion of existing customers. On the other hand, if there are too many customers in the system, that previous higher proportion would dramatically affect total profits. It will be optimal to decrease the price in order to reduce the defection proportion at the expense of allowing more new customers to arrive to the system and paying the penalty. This trade-off produces this counter-intuitive result.

2.6.3 Bellman equation

We now turn back to the infinite horizon case. By taking the limit to (2.42) as $k \to \infty$, we end up with the infinite horizon formulation so we can write down the Bellman’s system of equations for this problem as follows:

$$J(N) = \bar{l}p - \bar{w}K + \rho \sum_{j=0}^{\hat{N}} \Pr(l + w = j|N)J(j) \quad \forall N = 1, \ldots, \hat{N}$$

In matrix notation this system of equation would be

$$\rho \begin{bmatrix}
\gamma_{11} - \rho^{-1} & \gamma_{12} & \cdots & \gamma_{1\hat{N}} \\
\gamma_{21} & \gamma_{22} - \rho^{-1} & \cdots & \gamma_{2\hat{N}} \\
\vdots & \vdots & \ddots & \vdots \\
\gamma_{\hat{N}1} & \gamma_{\hat{N}2} & \cdots & \gamma_{\hat{N}\hat{N}} - \rho^{-1}
\end{bmatrix} \begin{bmatrix} J(1) \\ J(2) \\ \vdots \\ J(\hat{N}) \end{bmatrix} = \begin{bmatrix}
\lambda(p_1)K - q(p_1)p_1 \\
\lambda(p_2)K - 2q(p_2)p_2 \\
\vdots \\
\lambda(p_{\hat{N}})K - \hat{N}q(p_{\hat{N}})p_{\hat{N}}
\end{bmatrix}$$

where $\gamma_{ij} = \Pr(l + w = j|i)$, and $p_i$ is the price to charge in a certain period when at the beginning of that period there were $i$ customers in the system. The objective is to find a pricing policy which maximizes the value function at each state, i.e., to solve the following problem:

$$J^*(N) = \max_p \left[ q(p)Np - \lambda(p)K + \rho \sum_{j=0}^{\hat{N}} \Pr(l + w = j|N)J^*(j) \right].$$
2.6.4 Certainty equivalent controller (CEC)

We now turn to a suboptimal control scheme called certainty equivalent controller, which applies at each period the control (price) that would be optimal if all of the uncertain quantities (\(l_k\) and \(w_k\)) were fixed at their expected values, i.e., we replace \(l_k\) by \(\bar{l}(p_k, N_k)\) and \(w_k\) by \(\bar{w}(p_k)\). Thus, our CEC formulation is

\[
J_k(N_k) = \max_{p_k} \left[ l_k p_k - \bar{w}_k K + \rho J_{k+1}(\bar{l}_k + \bar{w}_k) \right]
\]  

which we also call the deterministic version of our stochastic DP formulation.

Figure 2-18 shows the results of a numerical experiment\(^9\) which gives us the insights to develop the rest of this section. This experiment is based on the CEC in (2.46). We started the system in period 0 with three different customer bases \((N_0 = 15, 45, 100)\). In part (a) of the figure it can be observed that regardless the initial customer base, the number of customers in the system converges to a constant (in this experiment \(N^* = 21.7\) approximately). Part (b) of the figure shows the optimal prices at each state at each point in time, which also converge to a constant \((p^* = 4.1\) approximately). Therefore, we can conclude that in steady state there is an optimal number of customers and an optimal price.

\[\begin{align*}
&\text{(a) Customer base convergence} \\
&\text{(b) Optimal price convergence}
\end{align*}\]

Figure 2-18: System convergence to steady state

When the system reaches the steady-state number of customers \(N^*\), it will remain in that state forever in a deterministic system and will remain around that state in the stochastic one. In steady-state the expected number of customers who arrive to the system

---

\(^9\)For this numerical experiment we have considered \(K = 10\), \(\lambda = 5\), \(C = 0.04\) (the constant representing the competitors offerings in the discrete choice probabilities), \(r = 0.1\) (the discount rate), \(\beta = \frac{1}{3}\) (the price response function parameter), \(F(p) = \frac{p}{10}\), and \(q(p) = e^{-p/3}/(e^{-p/3} + C)\).
will be equal to the expected number of customers who leave the system at a given period, i.e.,

$$\lambda(1 - F(p^*)) = (1 - q(p^*))N^* \quad (2.47)$$

where $p^*$ is the optimal price to charge in steady-state.

We have assumed throughout this chapter that the probability that a customer continues to the next period is non-increasing with the price, i.e., $\frac{\partial q(p)}{\partial p} \leq 0$. Considering that, we can prove that the steady-state number of customers decreases with the price. Thus, by solving for $N^*$ from (2.47) and computing its first derivative, we get:

$$\frac{\partial N^*(p)}{\partial p} = \frac{\lambda(1 - F(p))\frac{\partial q(p)}{\partial p} - \lambda f(p)(1 - q(p))}{(1 - q(p))^2}$$

which is clearly always less than zero. In other words, an increase in price yields to a decrease in the steady-state customer base.

**Proposition 17** The steady state optimal number of customers and optimal price can be obtained by solving the fixed point solution $p^*$ of the following equation, and then $N^*$ follows.

$$p^* = \frac{(q(p^*) + q(1 - \rho q))(1 - F(p^*))}{\rho q(1 - q)f(p^*)} + \left(\frac{1 - \rho q}{\rho q}\right)K$$

$$N^* = \frac{\lambda(1 - F(p^*))}{1 - q(p^*)}.$$  

Note that the steady-state price is independent of the arrival rate.

**Proof.** This proof is provided in the remaining part of this section. □

We now turn to find the steady-state pair of number of customers and price, i.e., $(N^*, p^*)$, in order to prove the above proposition. These two variables have to fulfill not only the condition in (2.47) but also the first-order condition of the following nonlinear programming model where the input parameter $N_0 = N^*$.

$$\max_p J(N_0) = \sum_{k=0}^{\infty} \rho^k (\bar{l}p - \bar{w}K)$$

$$s.t. \quad N_{k+1} = \bar{l} + \bar{w}$$

$$p \geq 0$$

where $\bar{l} = q(p)N_k$ and $\bar{w} = \lambda(1 - F(p))$. By developing the series formed by the constraint
we can find an expression which only depends upon the initial number of customers $N_0$.

$$\begin{align*}
N_1 &= qN_0 + \bar{w} \\
N_2 &= q[qN_0 + \bar{w}] + \bar{w} = q^2N_0 + \bar{w}(1 + q) \\
N_3 &= q[q^2N_0 + \bar{w}(1 + q)] + \bar{w} = q^3N_0 + \bar{w}(1 + q + q^2) \\
N_4 &= q[q^3N_0 + \bar{w}(1 + q + q^2)] + \bar{w} = q^4N_0 + \bar{w}(1 + q + q^2 + q^3).
\end{align*}$$

Thus, we can generalize the expansion of this recursion as:

$$N_k = q^kN_0 + \bar{w} \frac{1 - q^k}{1 - q}. \quad (2.48)$$

We now use (2.48) to reduce our model to an unconstrained nonlinear program as follows:

$$J(N_0) = \sum_{k=0}^{\infty} \rho^k(qNkp - \bar{w}K) = \sum_{k=0}^{\infty} \rho^k(q^{k+1}N_0 + qp\bar{w}\frac{1 - q^k}{1 - q} - \bar{w}K) = qN_0p \sum_{k=0}^{\infty} (pq)^k + \bar{w}qp \sum_{k=0}^{\infty} \left( \rho^k - (\rho q)^k \right) - \bar{w}K \sum_{k=0}^{\infty} \rho^k = N_0 \frac{q}{1 - pq} + \bar{w}qp \left[ \frac{1}{1 - \rho} - \frac{1}{1 - \rho q} \right] - \bar{w}K \frac{1 - \rho}{1 - \rho q}.$$

Then, rearranging some terms and letting $C(p) = \frac{qp}{1 - pq}$, we get the following reduced value function

$$J(N_0) = C(p) \left( N_0 + \frac{\rho}{1 - \rho} \bar{w} \right) - \frac{K}{1 - \rho} \bar{w}$$

which can be easily shown to be quasi-concave, i.e., there is a unique global optimal solution. If $p \to 0$ then the value function $J(N_0) \to \frac{-K}{1 - \rho} < 0$, and if $p \to \infty$ then $J(N_0) \to C(p)N_0$. We know that $C(p) > 0$ for all positive prices, and that $q(p)$ converges to zero quite fast, so $C(p) \to 0$ from above as $p \to \infty$. Therefore, by taking the first derivative with respect to the price $p$ we can get the desired sufficient first-order condition.

$$\frac{\partial J(N^*)}{\partial p} = \frac{\partial C(p)}{\partial p} \left( N^* + \frac{\rho}{1 - \rho} \bar{w} \right) - C(p) \frac{\rho}{1 - \rho} \lambda f(p) + \lambda f(p) \frac{K}{1 - \rho}. \quad (2.49)$$

where

$$\frac{\partial C(p)}{\partial p} = \frac{\partial q}{\partial p} p + q (1 - \rho q) \left(1 - \rho q \right)^2. \quad (2.50)$$

By replacing (2.50) in (2.49), letting $\frac{\partial J(N^*)}{\partial p} = 0$, and combining with the equilibrium
We finally get:

\[
\frac{\partial q}{\partial p} \frac{p}{1 - \rho q} + q \left(1 - \rho q\right) \left(1 - \rho q\right) N^* + \rho \bar{w} \right) \left(1 - \rho q\right) - qpp\lambda f(p) + \left(1 - \rho q\right) \lambda f(p) K = 0
\]

\[
\frac{\partial q}{\partial p} \frac{p}{1 - \rho q} + q \left(1 - \rho q\right) \lambda \left(1 - F(p)\right) \left(1 - \rho q\right) - qpp + \left(1 - \rho q\right) K = 0
\]

\[
\frac{\partial q}{\partial p} \frac{p}{\rho q \left(1 - q\right) f(p)} - \frac{q'p + q \left(1 - \rho q\right) \left(1 - F(p)\right)}{\rho q \left(1 - q\right) f(p)} + \left(\frac{1 - \rho q}{\rho q}\right) K = 0.
\]

Therefore, the optimal price \( p^* \) in steady-state satisfies the fixed point solution of the following equation:

\[
p^* = \frac{q'p^* + q \left(1 - \rho q\right) \left(1 - F(p^*)\right)}{\rho q \left(1 - q\right) f(p^*)} + \left(\frac{1 - \rho q}{\rho q}\right) K \tag{2.51}
\]

where \( q \) and \( q' \) correspond to the probability \( q(p^*) \) and its first derivative with respect to the price evaluated in \( p^* \) respectively. Note that \( p^* \) does not depend upon the customer arrival rate, what was expected.

Now, using (2.47) and \( p^* \) from (2.51) we can get the optimal number of customers in steady-state as:

\[
N^* = \frac{\lambda \left(1 - F(p^*)\right)}{1 - q(p^*)}
\]

It is worth noting that (2.51) will have solution only for certain values of \( K \). In case of more than one intersection (more than one fixed point), the optimal price will be the lowest one as shown in Figure 2-19, because it will lead to a higher \( N^* \) and therefore to higher profits. When there no longer exists a fixed point solution (for high setup costs), it means that prices are so high that \( N^* = 0 \), i.e., it is not worth acquiring new customers at the expense of \( K \) because it is too high.

**Proposition 18** The value function \( J_k(N_k) = l_k p_k - w_k K + \rho J_{k+1}(N_{k+1}) \) is linear in \( l_k \) and \( w_k \).

**Proof.** Let us prove this backward in time. We start first with the last period, for instance, period 2. It is clear that

\[
J_2(N_2) = l_2 p_2 - w_2 K
\]
Figure 2-19: $H(p)$ changes with the setup cost $K$

is a linear function in both $l_2$ and $w_2$. Now one step back and we analyze

\[
J_1(N_1) = l_1 p_1 - w_1 K + \rho J_2(N_2) \\
= l_1 p_1 - w_1 K + \rho J_2(l_1 + w_1) \\
= l_1 p_1 - w_1 K + \rho [l_2(l_1, w_1)p_2 - w_2 K].
\]

We know that the random variable $l_2$ depends upon the outcomes of the random variables $l_1$ and $w_1$. It is a binomial random variable which can take at most the value $l_1 + w_1$. Thus, we can clearly see that the domain of $(l_1, l_2)$ is convex. In that way if we consider the vectors $(l_1, w_1, l_2, w_2)^1$ and $(l_1, w_1, l_2, w_2)^2$ we have

\[
J_1(\alpha(l_1, w_1, l_2, w_2)^1 + (1 - \alpha)(l_1, w_1, l_2, w_2)^2) = (\alpha l_1^1 + (1 - \alpha) l_1^2) p_1 \\
- (\alpha w_1^1 + (1 - \alpha) w_1^2) K \\
+ \rho (\alpha l_2^1 + (1 - \alpha) l_2^2) p_2 \\
- \rho (\alpha w_2^1 + (1 - \alpha) w_2^2) K.
\]

Now, we know that $l_2^1 \in [0, l_1^1 + w_1^1]$ and $l_2^2 \in [0, l_1^2 + w_1^2]$ and so on and so forth. Thus,

\[
J_1(\alpha(l_1, w_1, l_2, w_2)^1 + (1 - \alpha)(l_1, w_1, l_2, w_2)^2) = \alpha J_1((l_1, w_1, l_2, w_2)^1) \\
+ (1 - \alpha) J_1((l_1, w_1, l_2, w_2)^2)
\]

what means that the function is linear. □

**Corollary 19** Given that the value function is linear, equality holds for Jensen’s ineq-
ity, i.e., we have that
\[ E \left[ J(l, w) \right] = J \left( E[l], E[w] \right). \]
The formulation obtained in (2.43) and the certainty equivalent one in (2.46) are equivalent.

### 2.6.5 CEC transient phase

The analysis of the transient phase while the system is reaching the steady-state is not trivial. Nevertheless, we will see that, given the properties of this problem, the DP algorithm will provide optimal policies for dealing with the system at any point in time. Still in the deterministic version, let us assume that after \( T \) periods the system reaches to steady-state, having started with \( N_0 \) customers. Thus, we can write the total expected revenues as:

\[
J(N_0) = \left( \sum_{k=1}^{T} \rho^{k-1} g(N_{k-1}, p_k, \bar{w}(p_k)) \right) + \rho^T J(N^*),
\]

If we want to maximize \( J(N_0) \) we need to find the optimal set of prices \( p_1, \ldots, p_T \) such that \( J(N_0) \) takes its maximum value. Thus, we become interested in knowing more about the function \( g(N_{k-1}, p_k, \bar{w}(p_k)) \). Our intuition is that these functions are quasi-concave in “reasonable price ranges.” In the left of Figure 2-20, we show an example\(^{10}\) of \( g(N_0, p_1, \bar{w}(p_1)) \) which is only function of \( p_1 \), and \( g(N_1, p_2, \bar{w}(p_2)) \) which is a function of both \( p_1 \) and \( p_2 \) because \( N_1 \) depends upon \( p_1 \). The right part of the figure depicts the sum of the revenue of period 1 and the discounted revenue of period 2.

If we force \( T \) to be a certain number, i.e., we want the system to be in steady-state after \( T \) periods, we have to add the constraint which ensure that the number of old customers at the beginning of period \( T + 1 \) is \( N^* \), i.e., \( q(p_T)N_{T-1} + \bar{w}(p_T) = N^* \), or the following condition after working out the recursion:

\[
N_0 \prod_{k=1}^{T} q(p_k) + \sum_{k=1}^{T} \bar{w}(p_k) \prod_{j=k+1}^{T} q(p_k) = N^*, \tag{2.52}
\]

where \( \prod_{j=T+1}^{T} q(p_k) = 1 \). For the case of the example in Figure 2-20, constraint (2.52) is function of \( p_1 \) and \( p_2 \), as it is shown in Figure 2-21. Thus, the optimization of the right function in Figure 2-20 has to be performed over the feasible set depicted in Figure 2-21. In this example the prices \( p_1 = 7.44 \) and \( p_2 = 10.16 \) are optimal and fulfill the constraint that the steady-state number of 44.3 customers is reach in period 3. That is, we start with 100 customers, then after setting \( p_1 \) we finish period 1 with 77 customers, then after

\(^{10}\)For this numerical experiment we have considered \( N_0 = 100, K = 20, \lambda = 10, C = 0.04 \) (the constant representing the competitors offerings in the discrete choice probabilities), \( r = 0.3 \) (the discount rate), \( \beta = \frac{1}{3} \) (the price response function parameter), \( F(p) = \frac{p}{100} \), and \( q(p) = e^{-p/3}/(e^{-p/3} + C) \).
Figure 2-20: Quasi-concavity of functions $g(N, p, w)$

setting $p_2$ we finish period 2 with $N^* = 44.3$.

Figure 2-21: Example of constraint (2.52) for $T = 2$.

If we do the same with higher values of $T$, we can see that we can reach our initial intuition from Remark 16. The intuition was that if we start the system with a number of customers greater than the steady-state number of customers, then we start with an optimal price less than the steady-state optimal price, and the following optimal prices are higher and higher, but never greater than $p^*$ (see Figure 2-22). This phenomenon occurs because of the parameters chosen for this example. We analyze this with more details in section 2.6.7.
Figure 2-22: Optimal prices when the system reaches steady-state in $T$ periods.

We have found by means of several computational experiments that the sequence of optimal prices are within a very relative small range, which implies that a good lower bound of $J^*(N_0)$ would be given by setting the prices of all periods to $p^*$. Thus, the lower bound would be given by

$$J_L(N_0|p^*) = \sum_{k=0}^{\infty} \rho^k g(N_k, p^*, \bar{w}(p^*))$$

$$= N_0 \frac{qp}{1-\rho q} - \bar{w} \left[ \frac{K}{1-\rho} - \frac{\rho qp}{(1-\rho)(1-\rho q)} \right]$$

$$\leq J^*(N_0).$$

### 2.6.6 Optimal fixed-price policy

There exists an interest rate $\hat{r}$ such that the optimal pricing policy corresponds to a fixed-price policy. In order to let this happen, the fixed optimal price should be given by the following optimization program:

$$\max_{\rho \geq 0} J(N_0) = \sum_{k=0}^{\infty} \rho^k (qN_k p - \bar{w} K)$$

s.t.

$$N_{k+1} = qN_k + \bar{w}$$
Figure 2-23: Interest rate for which a fixed-price policy is optimal.

By solving this in the same way we did for the proof of proposition 17, we can work out the recursion and derive the first-order condition as follows:

\[
p = \frac{\left( \frac{\partial q}{\partial p} p + q (1 - \rho q) \right) \left( (1 - \rho) N_0 + \rho \bar{w} \right)}{(1 - \rho q) q \rho \lambda f(p)} + \frac{(1 - \rho q)}{q \rho} K. \tag{2.53}
\]

The solution is a fixed point, and it can be solved numerically. The RHS of (2.53) is denoted by \( H_2(p, r) \).

In addition, we know that in steady-state the prices remain constant (in the CEC case), so the maximizing price \( p \) of \( J(N_0) \) should be equal to the steady-state optimal price \( p^* \) in (2.51). Thus, we have a system of two equations and two unknowns, \( p \) and \( \hat{r} \). Figure 2-23 illustrates the intersection of these conditions in order to determine the critical interest rate \( \hat{r} \) (or discount factor \( \hat{\rho} = 1/(1 + \hat{r}) \))\(^{11}\).

### 2.6.7 Numerical experiments

In this section we illustrate graphically the computational experiments performed in order to validate some of our findings. In this instance, we have considered a setup cost \( K = 10 \), an external arrival rate \( \lambda = 10 \) (the effective arrival rate is \( \lambda \) times the complement of the cumulative reservation price distribution), a constant representing the competitors offerings in the discrete choice probabilities \( C = 0.04 \), an interest rate \( r \) between 10% and 35%, a customer price response function parameter \( \beta = -\frac{1}{3} \), a cumulative distribution for the reservation price \( F(p) = \frac{p}{100} \), and \( q(p) = e^{-p/3}/(e^{-p/3} + C) \) as the probability that any

\[^{11}\text{In this example we have used } N_0 = 80, \text{ but note that the result is independent of the the initial number of customers } N_0.\]
particular customer will not leave the system at the current period given that the price is \( p \). The MatLab source code used for these experiments can be found in appendix C.2.

The optimal pricing policy, i.e., the optimal price to be charged at different levels of the customer base, is the result of our dynamic programming formulation solved backward in time (value iteration algorithm). The plot on the left of Figure 2-24 shows the optimal policy at different interest rates. We can clearly observe the result presented in Remark 16 that prices increase or decrease with the customer base. The plot on the right shows the value function \( J(N|r) \), i.e., the total expected and discounted profits generated when using the optimal policy. As it was expected, profit increases with the customer base, and it also increases when the discount factor \( \rho \) increases (future profits become more important).

In terms of simulations using the optimal policies that we found, we performed thousands of them for the case when the interest rate is 10%. We can observe in Figure 2-25 that regardless the initial number of customers \( N_0 \), the system tends to move its customer base toward the steady-state or equilibrium \( N^* = 21.7 \) shown in Figure 2-18. In this figure the lowest curve represents the number of arrival customers at each period (withdrawn from a Poisson distribution), the middle curve represents the number of customers who do not leave the system (withdrawn from a Binomial distribution), and the upper and thicker curve represents the total number of customers in the system at the beginning of each period. In the case when the initial customer base is \( N_0 = 50 \), the optimal value function obtained from the backward value iteration algorithm was \( J(50) = 992.3 \). After performing several (thousands) simulations and then averaging them we got \( \bar{J}(50) = 991.8 \) with a standard deviation \( \sigma = 101 \), which is rather closed to the optimal value function.

Recalling Remark 16 and the theory shown in Section 2.6.6 about finding an optimal fixed-price policy, we can observe this phenomenon in Figure 2-26. This figure shows clearly the switching effect from two different angles. By switching effect we mean when
the optimal policy switches from an increasing optimal price with number of customers, to a decreasing optimal price with the number of customers. When the switch occurs, optimal prices are the same regardless of the number of customers in the system. In this example the switch happens when the interest rate is 24%.

As we move away from \( \hat{\rho} = 0.24 \), the fixed-price policy value function \( J_{FP}(N) \) also moves away from the optimal pricing policy value function \( J^*(N) \). Setting the interest rate at 10\%, we computed the relative difference on the value functions between the optimal pricing policy and several fixed-price policies as \( (J^*(N) - J_{FP}(N|p)) / J^*(N) \). The left plot in Figure 2-27 depicts this GAP, and it can be observed that the gap is quite small for large number of customers and for prices near the steady-state equilibrium price \( p^* = 4.1 \).
The plot in the right is a transversal cut at price 4.1 of the gap surface in the left. It shows that the gap is zero for \( N = 21.7 \), the steady-state number of customers and that it is below 1% for most of the cases.

![Figure 2-27: GAP between optimal pricing policy and fixed-price policy](image)

### 2.6.8 Single bundle system as an M/M/∞ queue

Finally, on the right end of figure 2-16 we have the case where there is no discount when adding up future profits, i.e., \( \rho \rightarrow 1 \). For that case, let us consider an M/M/∞ system where the interarrival times are independent and exponential, with parameter \( \lambda_n \) when there are \( n \) customers in the system. The service times are also independent and exponential, with parameter \( \mu_n \) when there are \( n \) customers in the system. Service times and interarrival times are independent (see Figure 2-28).

![Figure 2-28: M/M/∞ queue with state-dependent rates](image)

Using the level crossing rate balance equations \( \lambda_n \Pr(n) = (n + 1)\mu_{n+1} \Pr(n + 1) \) for the compound states \( \{0, 1, \ldots\} \) of the process \( \{N(t), t \geq 0\} \), we can get the stationary
distribution of the number of customers in the system as:

\[
\Pr(n) = \Pr(0) \prod_{k=0}^{n-1} \frac{\lambda_k}{(k+1)\mu_{k+1}}
\]  

(2.54)

where

\[
\Pr(0) = \left[ 1 + \sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{\lambda_k}{(k+1)\mu_{k+1}} \right]^{-1}.
\]  

(2.55)

Both the arrival and service rates depend upon the price, i.e., \( \lambda_n = \lambda(p_n) \) and \( \mu_n = \mu(p_n) \). The arrival rate is such that \( \lambda_0 \) is a constant, its derivative with respect to price is negative, and \( \lim_{p \to \infty} \lambda(p) = 0 \). The service rate is an increasing function of \( p \).

We can observe that this queueing system is the continuous version of our previous formulation. Now we are in a continuous time domain. We have the same arrival process, a Poisson process with arrival rate depending upon the price, and the departure process is a Poisson process with rate \( n\mu_n \), which corresponds in the discrete domain to a Binomial distribution with parameters \( n \) and \( (1 - q(p_n)) \). In this system we can compute expected profit per unit of time, which will be the limit, as \( T \) goes to infinity and the discount factor goes to 1, of the average period-profit of our DP formulation, i.e., this \( M/M/\infty \) system maximizes \( \lim_{T \to \infty} J^* / T \).

Therefore, the expected profit per unit of time could be computed as

\[
\Pi = E[\text{revenue} - \text{cost}]
= \sum_{n=0}^{\infty} (np_n - K\lambda_n) \Pr(n).
\]  

(2.56)

To characterize analytically the optimal solution for the total expected profit in (2.56) is quite difficult, mainly because of the large number of prices (decision variables) to be determined and the functional form of the queue’s steady-state probabilities.

At the right of Figure 2-29 we can observe the optimal set of prices\(^{12}\) which maximize expected profits, while at the left the steady-state probabilities are depicted. From the figure we can see that optimal prices increase with the number of customers in the system. Proving this analytically can be extremely difficult as mentioned before, but some intuition and results could be withdrawn from a simpler case.

**Case of two possible prices.** Let us assume that the prices can take only two values, a low price denoted by \( p_L \) and a high price denoted by \( p_H \). Let \( \hat{N} \) be the number of customer such that we charge \( p_L \) when there are in the system a number of customers less than or equal to \( \hat{N} \), and we charge \( p_H \) otherwise. Thus, the total expected profit is

\(^{12}\)For this example we have used almost the same parameters as in the previous ones. The departure rate \( \mu \) has been replaced by the probability of leaving the system \( (1 - q) \).
Figure 2-29: An example of an optimal pricing policy

given by

$$\Pi = \sum_{n=0}^{\hat{N}} (np_L - K\lambda_L) Pr(n) + \sum_{n=\hat{N}+1}^{\infty} (np_H - K\lambda_H) Pr(n)$$  \hspace{1cm} (2.57)$$

which can be expressed as

$$\Pi = p_L \sum_{n=0}^{\hat{N}} n Pr(n) - K\lambda_L Pr(n \leq \hat{N}) + p_H \sum_{n=\hat{N}+1}^{\infty} n Pr(n) - K\lambda_H \left(1 - Pr(n \leq \hat{N})\right)$$

$$= p_L \sum_{n=0}^{\hat{N}} n Pr(n) + p_H \sum_{n=\hat{N}+1}^{\infty} n Pr(n) + K(\lambda_H - \lambda_L) Pr(n \leq \hat{N}) - K\lambda_H$$

$$= (p_L - p_H) \sum_{n=0}^{\hat{N}} n Pr(n) + p_H \tilde{N} + K(\lambda_H - \lambda_L) Pr(n \leq \hat{N}) - K\lambda_H$$

where $\tilde{N}$ denotes the expected number of customers when the queue system is in steady-state.

Let us see the summations involved in the RHS of the profit function (2.57),

$$\sum_{n=0}^{\hat{N}} n Pr(n) = Pr(0) \sum_{n=1}^{\hat{N}} \frac{\rho_L^n}{n!} = Pr(0) \rho_L \sum_{n=0}^{\hat{N}-1} \frac{\rho_L^n}{n!}$$

$$\sum_{n=\hat{N}+1}^{\infty} n Pr(n) = Pr(0) \frac{\lambda_L \rho_L^{\hat{N}}}{\mu_L} \sum_{n=\hat{N}}^{\infty} \frac{\rho_L^n}{n!}$$

where $\rho_L = \frac{\lambda_L}{\mu_L} > 1$, $\rho_H = \frac{\lambda_H}{\mu_H} > 1$, and $\rho_L/\rho_H > 1$. In addition, the probability of an
empty system is given by

\[ \Pr(0) = \left[ \sum_{n=0}^{\hat{N}} \frac{\rho^L_n}{n!} + \sum_{n=N+1}^{\infty} \frac{\rho^H_{n-N-1}}{n!} \frac{\rho^L_{\hat{N}}}{{\mu}_H} \right]^{-1} \]

\[ = \left[ \sum_{n=0}^{\hat{N}} \frac{\rho^L_n}{n!} + \frac{\rho^L_{\hat{N}}}{{\mu}_H} \sum_{n=N+1}^{\infty} \frac{\rho^H_n}{n!} \right]^{-1}. \]

Figure 2-30 shows an example where we force a price switching when the system goes from eight to nine customers or vice-versa, i.e., \( \hat{N} = 8 \). Leaving \( p_L \) and \( p_H \) as decision variables and then solving the problem, we can see that again \( p_H > p_L \) as expected.

\[ \begin{array}{c}
\text{Number of customers} \\
\text{Steady-state probability} \\
0.00 & 0.02 & 0.04 & 0.06 & 0.08 & 0.10 & 0.12 & 0.14 & 0.16 & 0.18 \\
\end{array} \]

\[ \begin{array}{c}
\text{Number of customers} \\
\text{Optimal price} \\
0 & 0.5 & 1 & 1.5 & 2 & 2.5 & 3 \\
\end{array} \]

Figure 2-30: Example of a two-price optimal policy

### 2.6.9 Multiple bundles model formulation

Let the vector \( \mathbf{N} \) be the state variable where \( N_i \) is the number of customers in the system using bundle \( i \).

Let the vector \( \mathbf{w} \) be the random variables of the arrival process. Let assume a general Poisson arrival with rate \( \lambda \), which is split into \( B + 1 \) sub processes by the mean of choice probabilities \( \alpha_j \) as in the one-step decision process model. Thus, the arrival rates of the split processes will be \( \lambda \alpha_j(p) \), where \( p \) is the \( B \)-dimensional vector of prices. For the case of a two-step decision process, the arrival rates would be given by \( \lambda \pi_j(p) (1 - F(p_j)) \) where \( F(\cdot) \) is the cumulative distribution of the reservation prices.

The net profit of a particular period is given by

\[ g(\mathbf{N}, p, l, \mathbf{w}) = \sum_{j=1}^{B} l_j p_j - \sum_{j=1}^{B} w_j K_j. \]
Thus, the DP algorithm would be
\[
J_k(N_k) = \max_{p_k} E_{l,w} \{ g(N_k, p_k, l_k, w_k) + \rho J_{k+1}(l_k + w_k) \}.
\]

By taking expectation and considering the fact that we are modeling an infinite horizon problem, we can write down the Bellman’s system of equations as follow:
\[
J(N) = g(N, p, \bar{l}, \bar{w}) + \rho \sum_{k_1=1}^{N_1} \cdots \sum_{k_B=1}^{N_B} \Pr(k_1, \ldots, k_B) J(k_1, \ldots, k_B).
\] (2.58)

Since the problem becomes more and more complicated and difficult to be solved as \( B \) grows, we focus our analysis in the case of only two bundles. Thus, the Bellman’s system of equations in (2.58) for the case \( B = 2 \) is given by:
\[
J(N_1, N_2) = [q(p_1)N_1p_1 + q(p_2)N_2p_2] - [w_1K_1 + w_2K_2] + \rho \sum_{k_1=1}^{N_1} \sum_{k_2=1}^{N_2} \Pr(k_1, k_2) J(k_1, k_2).
\]

### 2.7 Extensions and further research

In this section we present some natural extensions of the previous models for further research. The resulting models will be even more challenging, and there is no doubt that the results (building blocks) found in this chapter will help address them.

#### 2.7.1 Upselling opportunities

Sometimes it may be critical to give customers incentives to stay in the system longer. Jain and Singh [JS02] point out that it is believed that long lifetime customers are more profitable. To this regard, the upselling opportunities for the company could be significantly higher. According to our model, in order to increase the customer’s lifetime, lower prices have to be offered, so the new trade-off we face now is to reduce prices in order to increase upselling opportunities of other company’s products. The question now is by how much should the prices be decreased in order to maximize total expected profits?

#### 2.7.2 Down payment

An interesting extension is the case when the setup cost \( K \) is paid partially when signing the contract, i.e., the customer pays a down payment. Let \( \alpha \) be the proportion of the setup cost to be charged through the monthly fees, i.e., \((1 - \alpha)K\) corresponds to the down payment. What is the optimal value of \( \alpha \)? What are the marketing implications of doing that?
2.7.3 Current and new customers

Our model considers just the customers who are joining the system at any particular point in time, so it makes pricing decisions without considering current contracts of customers who joined the system in previous periods. If the model suggests a monthly payment much lower than the one paid by old customers for the same bundle, it is very likely that overall profits (from old and new customers) are not being maximized. Thus, taking this into account would lead to a not so low price but profits would be maximized.

![Diagram showing customers in the system](image)

Figure 2-31: Customers in the system

Figure 2-31 shows for example that at the beginning of period 2 there are four old customers besides the six new ones. So, if these four old customers were under a contract with a monthly payment of $10 and our model suggests a monthly payment of $7 for the six new customers, it is likely that the old customers’ contracts will have to be modified incurring in a high cost. This cost is the one that has to be considered when optimizing the profit function. Thus, for instance, an optimal price would be $8 instead. This extra dollar will offset the cost of reducing the old contracts’ monthly payments.

The simplest case of this problem has been studied in the previous section, where a single bundle is offered to customers and prices can fluctuate up and down without any constraint. More realistic settings include markdown constraints and multiple bundles.
2.7.4 Marketing expenditure

Companies have limited budgets for marketing expenditures. The purpose of this budget is to increase customer demand as much as possible. Because of this, we say that \( N_i \), the number of potential customers in segment \( i \), is a function of the marketing expenditure, i.e., the more we spend in marketing, the more potential customers will be in the system (nondecreasing function).

Let \( E \) be the marketing budget allocated by the company to promote all its products or bundles over all segments, and let \( e_i \) be a decision variable representing the marketing expenditure allocated to segment \( i \). Thus, \( N_i(e_i) \) represents the potential number of customers in segment \( i \) when \( e_i \) is spent in marketing (see Figure 2-32 where \( c_{i0} \) represents the number of potential customers when there is no marketing expenditure in that segment). Despite knowing that \( N_i \) is mainly affected by \( e_i \), it can also be indirectly affected by marketing in other segments. However, for the sake of simplicity we assume it only depends on \( e_i \).

\[
N_i(e_i) = c_{i0} + \text{function of } e_i \quad \text{(nondecreasing function)}
\]

\[
\sum_{i=1}^{S} e_i \leq E,
\]

where \( e = (e_1, ..., e_S) \) represents the vector of marketing expenditure in the different segments.

![Figure 2-32: Potential customers and marketing expenditure in segment i](image)

Let us reformulate the bundling problem including the above issue of marketing expenditure. The new formulation is:

\[
\text{P) } \quad \max_{p \geq 0, e \geq 0} \sum_{i=1}^{S} \sum_{j=1}^{B} N_i(e_i)(1 - \Psi_{ij}(p_j))\pi_{ij}(X, p)V_{ij}(p_j)
\]

subject to

\[
\text{P) } \quad \sum_{i=1}^{S} e_i \leq E,
\]

where \( e = (e_1, ..., e_S) \) represents the vector of marketing expenditure in the different segments.
2.7.5 Limited supply

How much supply of each product or bundle does the company have? Certainly this issue is very important because our previous formulations can select prices that unnecessarily capture more demand for a specific bundle than the available supply, yielding to lower profits. To avoid this happening, the following constraint should be added to the model:

$$\sum_{i=1}^{S} N_i(\epsilon_i)(1 - \Psi_{ij}(p_j))\pi_{ij}(X,p) \leq I_j, \quad \forall j = 1, \ldots, B,$$

where $I_j$ represents the available supply of bundle $j$.

2.7.6 Dynamics of the system

A dynamic model similar to the one developed in Bitran and Schilkrut [BS00] could be built here in order to understand what drives customer lifetime in the system. Customers arrive to the system and at every period they experience a service delivery with a certain quality level. This experience will affect the probability distribution of continuing or leaving the service. There are some exogenous factors (randomness of the economy, e.g., bankruptcy) that can also make a customer leave. Figure 2-33 shows the dynamics of the system.

![Figure 2-33: Dynamics of the system](image)

As we have mentioned before, the level of prices will influence the time a particular customer spends in the service. Here, we say that the quality of service a customer experiences every period will also influence lifetime.

2.8 Summary

This chapter addresses the problem of pricing bundles with high setup costs. The bundles considered contain an initial component which is delivered as soon as new customers buy
the bundle and a service component that is delivered on a continuous basis until the customer leaves the system. The bundle is paid through monthly fees, which include a cost component (to cover the expenses incurred by the company in providing the bundle) and a profit component.

This problem was analyzed through three models:

1. A one-step choice model that maximizes the total expected profits by computing the optimal prices that should be charged to a set of customers who sign a contract today. A nonlinear programming model is formulated.

2. A model analogous to (1) with a two-step choice process. In this model we also consider a nonlinear profit per bundle and customers are assumed to be random utility maximizers as well.

3. A third model considers the evolution of the customer base through time. We formulate the corresponding dynamic programming model.

The analysis shows that the higher the price of the bundle, the higher the cost component of the monthly fee. Furthermore, the analysis yields a closed-form expression for total profit per customer per bundle. This closed-form expression proves to be an essential building block in the optimization problem of finding the optimal prices because it is equivalent to an otherwise untractable summation. The total profit per customer per bundle can be expressed as a function of the moment generating function of the customer lifetime distribution.

The expected profit per customer is an increasing function of price. This result was proved in the case when the customer lifetime distribution is geometric. This is a significant result because it takes into account the competing dynamics that increasing the monthly fee simultaneously increases monthly revenue per customer and lowers the customer lifetime. Yet, an increase in price always results in higher profit regardless of the original price level. This result can then be incorporated into a more general model (including external demand), where a very high price will lower the probability that a given bundle will be bought.

The analysis of the one-step decision process shows that optimal prices fulfill an insightful first-order condition. More precisely, profit per customer per bundle minus its derivative with respect to price is constant for all bundles. This condition allows for the development of an algorithm that iteratively determines the optimal prices. The algorithm is very efficient because the first-order condition establishes that the price of one bundle automatically determines the price of all others. Therefore, the search space over which the optimization occurs can be reduced to feasible prices for a single bundle.

The first-order conditions for the two-step decision process are similar to those obtained for the one-step process. In this case, the product of the probability of purchase
with the total expected profit per customer per bundle (denoted \( R_j \)), minus its derivative with respect to price (denoted \( R'_j \)), is constant for all bundles. The functions \((R_j - R'_j)\) are concave with respect to price. Therefore, the search process is not as easy as the previous one since each profit level can correspond to two different prices. Nevertheless, an algorithm can still be implemented. Since the \((R_j - R'_j)\) functions are concave, we can easily find the maximum level achieved for each bundle. The lowest such level will be the highest possible level that can be achieved by the optimal total expected profit. Therefore, this profit level and its unique corresponding price can be used to find an upper bound for the overall profit. Using this price and the first-order conditions we can then find the optimal price and profit level for other bundles, thus deriving a lower bound for profit. In this way, the search space for the optimal profit level is bounded from above and below, and the search space for the optimal solution is well-defined. Iterations occur inside these bounds, which were experimentally found to be very small.

From the dynamic formulation of the customer base evolution, we found that for any given set of parameters, the optimal prices stabilize the number of customers in the system over time. For example, it is never optimal to have too many customers because this means that the current price is too low and more profit could be obtained by increasing prices. In other words, if the rate of defection is smaller than arrival rate, the number of clients always increases over time. By increasing prices, the rate of defection increases and the arrival rate decreases. Our intuition backed by numerical experiments was that steady-state number of customers in the system always converges to a constant. The optimal pricing policy depends upon the discount factor. We found that for discount factors below a known threshold, the optimal pricing policy is such that the price decreases with the number of customers. Conversely, for discount factors above this threshold the price increases with the number of customers. The threshold is well characterized analytically.

Comparing our dynamic programming stochastic formulation with the certainty equivalent counterpart we found an equivalence using Jensen’s inequality, because it turns out that the profit function is linear in the number of customers who continue in the system and in the number of new customers. Numerical experiments suggest that a single fixed-price policy is very nearly optimal when the system is loaded, which implies that offering multiple prices can at best capture second-order increases in profit.
Chapter 3

Bundle Pricing and Composition under Competition

This chapter addresses the problem of how to determine the composition and the price of a bundle so as to maximize total expected profits. The setting is a high-tech company in a highly competitive environment. Bundles are built from a set of components that must meet certain technical constraints. The company’s objective is to build a bundle and offer it in a market where it will compete with other substitute products. Consumers will purchase the bundle that maximizes their utility after examining all available bundles. Since consumer choice can vary depending on the purchase occasion, the decision process is modeled through a random utility model. Therefore, the selection of a bundle’s components and its price is made in light of the bundles with which it will be competing in the market and the uncertainty in the consumer choice process. The optimal decision can be found by solving a nonlinear mixed integer program. We propose two alternative solution methods to determine the optimal composition for the bundle and the price at which it should be offered.

3.1 Problem definition

This chapter addresses the problem of composing and pricing fully substitute bundles. The decision maker is a company operating in a competitive environment. Customers are heterogeneous and belong to different market segments. Therefore, companies must offer a product line that covers different price ranges in order to increase their market share (c.f. “Benefit Segmentation” as explained in Urban and Hauser [UH93, chapter 11]). The objective of introducing new products is to meet the needs of customers. A line of products is more effective than a single product in accomplishing this objective because not all customers have the same needs.

The decision maker’s objective is to maximize profits by determining the optimal prices
and composition of bundles formed from a given set of components. In the sequence of events, price ranges are first defined according to customers’ needs and willingness to pay. Then, one bundle is offered in each price range (this assumption can be relaxed). Finally, customers make their selection in one of the price ranges.

Feasible bundles must meet two types of constraints:

1. **Technological feasibility:** Combining components is not a random procedure. There are technical specifications and requirements which every bundle should meet. Components are grouped by type. Every bundle should fulfill constraints of the type: “must have exactly one of those,” “must have at least one of those,” “could have one of those,” “could have multiple of those,” etc.

2. **Competitive feasibility:** The composition and price of the competitors’ bundles are not controllable by the decision maker. The optimization model must take into account the competition. Decision makers with an unfavorable cost structure should not offer bundles very similar to those of their competitors. Conversely, decision makers with a favorable cost structure may want to offer bundles similar to those of their competitors, but at lower price ranges.

In addition to these feasibility constraints, we can find other constraints which will help narrow the set of feasible bundles for each price range. For example, if \( p^l \) and \( p^u \) are the lower and upper price limits of a certain range, a bundle with cost higher than \( p^u \) does not need to be considered in this subset.

![Figure 3-1: Composition of a feasible bundle](image)

Figure 3-1 shows how feasible bundles are constructed. The bundle contains a predefined set of “basic components,” and a set of “flexible components.” One component
is chosen from each of the \( m \) sets of components. Any problem's structure can fit into this framework by applying two rules. (1) In the case of constraints such as “must have exactly two components of Set 2,” we need to replicate that set (i.e., to create Set 2’) and change the constraint to “must have exactly one of Set 2 and one of Set 2’.” (2) If it is not necessary to include one component from a certain set (e.g., Set \( k \)), then the Set \( k \) is replaced by Set \( k’ = (\text{Set } k) \cup \{\text{no component}\} \).

A major issue of interest in the bundling literature is modeling a customer’s demand across different bundles. Customers’ willingness to pay (reservation price) varies across bundles because each bundle has a different level of attractiveness for each type of customer. For instance, if we compare a laptop with CD-RW drive and the same laptop with CD-ROM drive, customers are willing to pay more money for the first one (due to the higher attractiveness that a CD-RW has over a CD-ROM). In this study we propose to build an “attractiveness factor” for each bundle, which we denote by \( I_j \). The value of the attractiveness factor is the weighted sum of the individual attractiveness factor of each of the bundle’s components. For instance, Urban and Hauser [UH93, Chapter 10] use the term “importance weights” to denote the vector of components’ attractiveness, and they discuss issues relating to measuring these weights.

Where do the components’ attractiveness factors come from? Companies are surveying customers all of the time in several ways. Companies directly interview customers in brick and mortar stores, mailings, or enclose short surveys in products sold, offer free services for customers who respond to web-based surveys, etc. This information is used to estimate the relative level of attractiveness that each component has for the customers. This intermediate factor can be used instead of asking customers what they are willing to pay for each component. This methodology has the advantage of avoiding biased answers due to conflict of interests. Since consumer preferences can change over time, companies are continuously learning about their customers’ component evaluation. Urban and Hauser [UH93] point out that direct measures, preference regression, and conjoint analysis are good techniques in determining how bundle attractiveness changes as a result of bundle design modifications.

In what follows we present mathematical models that determine the bundle to be offered in each price-range, and find their optimal prices. The objective is to maximize total expected profits. Different approaches are discussed. In §3.2 we address the case where customer choice is based exclusively on the reservation price distribution, without taking into account the competitor’s offerings. In §3.3 we introduce customer choice based on utility maximization and solve both the pricing and composition problems. In §3.4 we extend the model from §3.3 to the case where there are multiple bundles offered by the company. We also propose a model which includes customers who are willing to pay an extra fee in order to customize their own bundles.
3.2 A model based on reservation price

We first consider a simple metric that allows us to find tractable ways to solve the price bundling problem. Instead of measuring the bundle’s attractiveness with a vector of components’ attractiveness (that would be the ideal case), we compute a weighted sum of all vector’s components into a single number which we refer to as “the bundle’s attractiveness factor.” Note that by doing this, some information is lost, in this case, the relative attractiveness of individual components.

Before presenting the model, we introduce the following notation:

- $\Omega$: set of all feasible bundles for any price-range
- $\Omega_k$: set of feasible bundles for price-range $k$ ($\Omega_k \subseteq \Omega$)
- $p_k$: price of bundle offered in price-range $k$
- $p^l_k$: lower limit of price-range $k$
- $p^u_k$: upper limit of price-range $k$
- $S_j$: set of components from which only one will be part of the bundle
- $c_j$: cost of bundle $j$ (sum of all components costs plus cost of basic components)
- $I_j$: attractiveness-to-customer factor of bundle $j$
- $\psi(p|I_j)$: probability density function for the reservation price of bundle $j$ given its attractiveness-to-customer
- $K$: number of price ranges in the market
- $B$: number of feasible bundles, i.e., cardinality of $\Omega$
- $B_k$: number of feasible bundles for price-range $k$, i.e., cardinality of $\Omega_k$

For the sake of simplicity we avoid mathematical notation at the component level. Thus, we present the following definitions.

**Definition 20** The attractiveness factor of bundle $j$ is given by

$$I_j = \sum_{k \in \{\text{components in bundle } j\}} (\text{attractiveness}_k \cdot \text{weight}_k)$$

**Definition 21** The cost of bundle $j$ is given by

$$c_j = (\text{cost of basic components}) + \sum (\text{cost of all components in bundle } j)$$

Let us assume that there is no relation among all optimal bundles and their prices across the different price ranges. In this case, we could solve an independent problem for each price range. This assumption could seem very strong, but it provides us with a tractable procedure to solve this problem. In a nutshell, what we do here is to enumerate all possible bundles a priori and then to evaluate one by one using a simple optimization
problem. By doing this we can avoid a formulation involving too many binary variables (which help compose feasible bundles), that really makes problems very difficult to solve.

**Definition 22** \( \Psi(\cdot|I_j) \) is the cumulative distribution function (cmf) of the reservation price for bundle \( j \), defined as \( (1-\Psi(p_j|I_j)) = \int_{p_j}^{\infty} \psi(r|I_j)dr \), where \( \psi(\cdot|I_j) \) is the probability density function of the reservation price for bundle \( j \) given its level of attractiveness.

The expression in Definition 22 allows us to compute the proportion of customers willing to pay a certain price for a bundle with a given attractiveness level. This will help us when computing the expected profit later on.

We assume that the higher the attractiveness factor of a bundle the higher the mean of the reservation price pdf. This assumption is depicted in Figure 3-2, where \( I_1 < I_2 < \cdots < I_{10} \). In the example of Figure 3-2 the shape of all curves is the same, i.e., \( I_j \) only impacts the mean of the pdf. In a more general situation, it could also impact the variance and the tails.

As mentioned before, our objective is to find for each price range an optimal bundle and its price while fulfilling some constraints and maximizing a company’s profits. Figure 3-3 shows a reservation price pdf for each price range \([p_l^k, p_u^k] \) and its optimal price \( p^*_k \).

Figure 3-3: Reservation price pdf for offered bundles in the different price-ranges
Let us build first the set \( \Omega \) of all feasible bundles. The set \( \Omega \) does not simply contain all possible combinations of components, it involves the fulfillment of several constraints (explained in Section 3.1). First, every bundle has to fulfill some technical requirements. There are \( q \) sets of components, and exactly one component has to be picked from every set in order to be part of the bundle (note that the sets can include the component “no component” which obviously has zero cost). Second, in case of an unfavorable cost structure it is most likely that the model would not choose competitors’ offerings in \( \Omega^C \), although these bundles will be part of all feasible bundles (\( \Omega^C \subseteq \Omega \)). By doing this, customers will not be able to compare a company’s offerings with those of the competition.

The next step is to build the individual bundle set \( \Omega_k \) for every price range, which of course will be subsets of \( \Omega \), i.e., \( \Omega_k \subseteq \Omega \) for all \( k = 1, \ldots, K \). Let \( \delta_a \) be the threshold for proportion of customers willing to pay less than the bundle cost, and let \( \delta_b \) be the threshold for proportion of customers willing to pay less than the lower limit of a price-range. The idea of dealing with these smaller sets of bundles is to reduce the time of solving each optimization problem. For instance, bundles with a high attractiveness level will probably have high costs so they will not be an optimal solution for a low-price range. Conversely, bundles with a low attractiveness level will not be attractive for any customer in a high-price range. Constraints (3.1a) and (3.1c) prevent the above situations from happening, and they help to reduce the computational burden. The solution found could be suboptimal, but it is not very likely to happen. Constraint (3.1a) ensures that for bundle \( j \) to be part of \( \Omega_k \) its cost has to be lower than the upper limit of price-range \( k \). Constraint (3.1b) makes sure that items with extremely high costs with respect to the demand, are excluded from the set \( \Omega \). Finally, constraint (3.1c) enforces that a large proportion (higher than \( 1 - \delta_b \)) of potential customers are willing to pay a price in range \( k \).

\[
\begin{align*}
c_j & \leq p_k^u \quad \text{(3.1a)} \\
\Psi(c_j|I_j) & \leq \delta_a \quad \text{(3.1b)} \\
\Psi(p_k^l|I_j) & \leq \delta_b \quad \text{(3.1c)}
\end{align*}
\]

If bundle \( j \) is offered in the price range \( k \), then the profit per bundle sold is \( (p_k - c_j) \). The total expected profit would be this expression times the proportion of customers willing to pay the price \( p_k \), i.e., \( (1 - \Psi(p_k|I_j))(p_k - c_j) \). This profit can be visualized in Figure 3-4, where it is represented by the shaded area. Since only one bundle will be offered in a particular price range, we would like this bundle to be the one with the highest total expected profit, that is, the one with the highest shaded area in the figure.

In this way, we can formulate an optimization model for finding the optimal bundle.
in a particular price-range. This would be given by

$$\max \max_{p_k^l \leq p_k \leq p_k^u} \{ (1 - \Psi(p_k | I_j))(p_k - c_j) \},$$

or equivalently, by finding the maximum of $B_k$ individual optimization problems we have

$$\max_{j=1,...,B_k} \{ \max_{p_{kj}^l \leq p_{kj} \leq p_{kj}^u} (1 - \Psi(p_{kj} | I_j))(p_{kj} - c_j) \},$$

where $p_{kj}$ is the decision variable of the optimization problem for bundle $j$, and where the price domain is a compact set ($p_k^l \leq p_k \leq p_k^u$).

Finally, the $K$ solutions obtained from solving these optimization problems have to be compared. In case the same bundle is selected as the optimal one in more than one price-range, the solutions and their profit contribution have to be compared. The same bundle cannot be offered in two different price ranges.

### 3.3 A model based on customer choice

In this section we consider competitors’ bundles in an explicit way. These bundles are not under our control, so we consider them as given information. Each bundle is characterized by its attractiveness $I_j$ and its price. Our control is the ability to build a new bundle (characterized also by its cost $c_j$) which should maximize expected profits. Customers choose among all offered bundles and pick the one that maximizes their utility.

We will concentrate on the case of a unique price-range, because there is no coupling constraint that would force to optimize across all price ranges. Thus, separable problems could be solved in order to find solutions for each customers’ segment.

Customers have to choose among the company’s bundle (bundle 1), the competitors’ bundles (bundle 2 to n which belong to $\Omega^C$), or simply not buy anything (bundle 0), as
in Figure 3-5. They will choose the alternative which maximizes their utility. As we have mentioned, bundle 1 is the only one under our control, in terms of selecting the bundle and setting its price.

![Diagram of choices](image)

**Figure 3-5: Controllable and non-controllable choices**

As it is common practice in the literature, we use a Multinomial Logit model (MNL) to assess customers’ choice probabilities, hence, the demand. From an analytical perspective this model provides nice closed form solutions. This model assumes that a customer’s utility is a random variable, and that it is the sum of a deterministic term and a stochastic term, as follows:

\[
U_i = \alpha^T I_i + \beta p_i + \varepsilon_i \quad \forall i = 0, ..., n
\]

where \( \alpha^T = (\alpha_1, ..., \alpha_m) > 0 \) is a vector of weights, \( \beta < 0 \) is a scalar, and \( \varepsilon_i \) is a random perturbation that accounts for factors such as unobserved attributes, imperfect information, or measurement errors that make the customer unable to determine the exact utility of a product upon initial examination. The vector \( I_i \) is the relative attractiveness of each component. In order to be consistent with the notation used in the literature, let \( V_i \) be the deterministic term of the utility, so \( U_i = V_i + \varepsilon_i \). MNL assumes that if \( \varepsilon_i \) are i.i.d. with double exponential or Gumbel distribution, then it can be shown that the probability of choosing bundle 1 is given by

\[
q = \frac{e^{V_1}}{e^{V_1} + \gamma}
\]  

(3.2)

where \( \gamma = e^{V_0} + \sum_{i \in \Omega \setminus \{1\}} e^{V_i} \), \( V_0 \) is the non-purchase deterministic utility. Note that \( q > 0 \) for any possible bundle. For a further explanation of this model refer to section 2.4.1 of this dissertation. Basically, we group probabilities all together but the one of choosing bundle 1, i.e., a customer will choose bundle 1 with probability \( q \) and will choose something else with probability \( (1 - q) \). We do that because the company has control and
perceive profits only from bundle 1, and has no control and gets zero profits from other choices (see Figure 3-5). Thus, the total expected profit for the company offering bundle 1 would be given by

\[ \Pi = q(p - c) + (1 - q)0 \]
\[ = q(p - c) \]

where \( c \) is the cost of bundle 1.

Let us include in the model the binary variables that will define the optimal bundle composition. Let the 0-1 matrix \( \mathbf{x} \) be the selected bundle, where its element \( x_{ij} \) is equal to 1 if component \( i \) in \( S_j \) is part of the bundle, and 0 otherwise. Thus, the nonlinear mixed-integer programming formulation for the bundling problem (BP) is given by

\[ \max_{p, \mathbf{x}} q(p, \mathbf{x})(p - c(\mathbf{x})) \]
subject to

\[ q = \frac{e^{\alpha I_1 + \beta p}}{e^{\alpha I_1 + \beta p} + \gamma} \]
\[ I_1 = (I_1, \ldots, I_m) \]
\[ I_j = \sum_{i \in S_j} I_{ij} x_{ij} \quad \forall j = 1, \ldots, m \]
\[ c = \sum_{j=1}^{m} \sum_{i \in S_j} c_{ij} x_{ij} \]
\[ 1 = \sum_{i \in S_j} x_{ij} \quad \forall j = 1, \ldots, m \]
\[ x_{ij} \in \{0, 1\} \quad \forall i, j \]
\[ p \geq 0 \]

Besides the decision variables’ non-negative and integer constraints, the only constraint of this problem is (3.4) which represent the fact that only one component is selected from each of the \( m \) sets \( S_j \). It has been already proved in [DK93] that non-linear programming models for product line selection and pricing problems such as (BP) are NP hard.

### 3.3.1 Bundle pricing optimization

Given the difficulty of solving this problem, we propose exploring first an enumerative approach, exploring every single possible bundle and ending up with the one which maximizes total expected profits. The optimization problem for finding the optimal price for a given bundle with attractiveness \( I \) and cost \( c \) is

\[ \max_{p \geq 0} \Pi = \frac{e^{\alpha I + \beta p}}{e^{\alpha I + \beta p} + \gamma}(p - c), \]
where $l$ and $c$ are determined by the given bundle.

In what follows we compute and characterize the solution of this optimization problem, showing that it is a maximum and analyzing the behavior of the expected profit with respect to the cost. Let us compute the first derivative of the objective function in (3.5).

$$\frac{\partial \Pi}{\partial p} = \frac{\beta e^{\alpha l + \beta p}(e^{\alpha l + \beta p} + \gamma) - e^{\alpha l + \beta p} \beta e^{\alpha l + \beta p}}{(e^{\alpha l + \beta p} + \gamma)^2} (p - c) + \frac{e^{\alpha l + \beta p}}{(e^{\alpha l + \beta p} + \gamma)}$$

$$= \frac{\beta \gamma e^{\alpha l + \beta p}}{(e^{\alpha l + \beta p} + \gamma)^2} (p - c) + \frac{e^{\alpha l + \beta p}}{(e^{\alpha l + \beta p} + \gamma)}$$

$$= \beta q (1 - q) (p - c) + q$$

$$= q(\beta (1 - q)(p - c) + 1)$$

The first order condition $\frac{\partial \Pi}{\partial p} = 0$ suggests that either $q = 0$ or $(\beta (1 - q)(p - c) + 1) = 0$. Having $q = 0$ means to have an infinite price which is not a feasible solution. Hence, we explore the latter condition.

$$q = \frac{\beta (p - c)}{\beta (p - c)} + 1$$

$$\frac{e^{\alpha l + \beta p}}{(e^{\alpha l + \beta p} + \gamma)} = \frac{\beta (p - c)}{\beta (p - c)} + 1$$

$$\beta (p - c) e^{\alpha l + \beta p} = (e^{\alpha l + \beta p} + \gamma) (\beta (p - c) + 1)$$

$$e^{\alpha l + \beta p} = -\gamma \beta (p - c) - \gamma$$

$$\alpha l + \beta p = \ln(-\gamma \beta (p - c) - \gamma)$$

**Proposition 23** From the first order condition (3.7) we can get the optimal price $p^*$ in closed form as follows:

$$p^* = c - \frac{1}{\beta}(1 + W(\frac{1}{\gamma} e^{\alpha l + \beta c - 1}))$$

where $W(\cdot)$ is the Lambert W function\(^1\). Note that the optimal price is always higher than the cost ($\beta < 0$ and $W(z) \geq 0$ for all positive $z$), hence profits are always positive (trivial result).

**Proof.** By replacing (3.8) in (3.7) we can easily prove the identity. Let the auxiliary

\(^1\)The Lambert W function is the solution of $W(z)e^{W(z)} = z$. Definitions and properties of this function that are relevant for our case are presented in Appendix B.2. Refer to Corless et al. [CGH+93] for a thorough analysis and understanding of this special function.
variable $z$ be equal to $\frac{1}{\gamma}e^{\alpha^T 1 + \beta c - 1}$, thus

\[
\alpha^T 1 + \beta \left( c - \frac{1}{\beta} - \frac{1}{\beta} W(z) \right) = \ln \left( -\gamma \beta \left( \left( c - \frac{1}{\beta} - \frac{1}{\beta} W(z) \right) - c \right) - \gamma \right)
\]

\[
\alpha^T 1 + \beta c - 1 - W(z) = \ln \left( -\gamma (\beta c - 1 - W(z) - \beta c) - \gamma \right)
\]

\[
\alpha^T 1 + \beta c - 1 - W(z) = \ln (\gamma W(z))
\]

\[
\alpha^T 1 + \beta c - 1 - W(z) = \ln (\gamma) + \ln(W(z))
\]

Then, from the definition of the Lambert W function $W(z)e^{W(z)} = z$, we can derive that \(\ln(W(z)) = \ln(z) - W(z)\). Replacing it into the identity we have

\[
\alpha^T 1 + \beta c - 1 - W(z) = \ln (\gamma) + \ln(W(z))
\]

\[
\alpha^T 1 + \beta c - 1 = \ln (\gamma) + \ln(\frac{1}{\gamma}e^{\alpha^T 1 + \beta c - 1})
\]

\[
\alpha^T 1 + \beta c - 1 = \ln (\gamma) + \ln(e^{\alpha^T 1 + \beta c - 1}) - \ln(\gamma)
\]

\[
\alpha^T 1 + \beta c - 1 = \alpha^T 1 + \beta c - 1
\]

thus, concluding this proof. \(\blacksquare\)

Next, we compute the second derivative of (3.5) in order to characterize the solution $p^*$ as a maximum point. We start from (3.6),

\[
\frac{\partial^2 \Pi}{\partial p^2} = \frac{\partial}{\partial p} (\beta (1 - q)(p - c)q + q)
\]

\[
= \beta (1 - q)(p - c)q' + \beta (1 - q)q - \beta q'(p - c)q + q'
\]

\[
= \beta (p - c)q'(1 - 2q) + \beta (1 - q)q + q'
\]

where $q' = \frac{\partial q}{\partial p} = \beta (1 - q)q$, so

\[
\frac{\partial^2 \Pi}{\partial p^2} = q(1 - q)(1 - 2q)(p - c)\beta^2 + 2q(1 - q)\beta. \quad (3.9)
\]

From the first order condition in (3.6) we know that there are two candidates or extreme points, one is $p^*$ in (3.8) and the other is a very high price, let us say $p^* \to \infty$. In both cases the objective function has a zero first derivative, so we need to know where the function is concave and where it is convex. In order to be convex we need $\frac{\partial^2 \Pi}{\partial p^2} \geq 0$, so from (3.9)

\[
q(1 - q)(1 - 2q)(p - c)\beta^2 + 2q(1 - q)\beta \geq 0
\]

\[
(1 - 2q)(p - c)\beta^2 \geq -2\beta
\]

\[
\frac{-\beta}{2}(1 - 2q)(p - c) \geq 1
\]
Figure 3-6: Expected profit function shape

which holds when $p^* \to \infty$ (remember $q > 0$ for all possible bundles). Then, $p^*$ in (3.8) would be a maximum since the objective function is concave in that neighborhood (see Figure 3-6).

**Corollary 24** The optimal expected profit $\Pi^*$ can be reduced to

$$\Pi^* = -\frac{1}{\beta} W\left(\frac{1}{\gamma} e^{\alpha T + \beta c - 1}\right)$$

(3.10)

or equivalently $\Pi^* = p^* - c + \frac{1}{\beta}$.

**Proof.** Starting form the profit function in (3.5) we have that the optimal profit is given by

$$\Pi^* = e^{\alpha T + \beta p^*} \left(\frac{1}{\gamma} (p^* - c) + \frac{1}{\gamma}\right)$$

where by replacing the optimal price $p^*$ from (3.8) we get

$$\Pi^* = e^{\alpha T + \beta c - W(z) - 1} \left(\frac{1}{\beta} W(z) + \frac{1}{\beta}\right)$$

where $z = K/\gamma$ and $K = e^{\alpha T + \beta c - 1}$. Then,

$$\Pi^* = -\frac{1}{\beta} (W(z) + 1) \frac{Ke^{-W(z)}}{Ke^{-W(z)} + \gamma}$$

$$= -\frac{1}{\beta} (W(z) + 1) \frac{ze^{-W(z)}}{ze^{-W(z)} + 1}$$
and by definition of the Lambert W function we finally get

\[ \Pi^* = -\frac{1}{\beta} (W(z) + 1) \frac{W(z)}{W(z) + 1} \]

\[ = -\frac{1}{\beta} W(z) \]

\[ = -\frac{1}{\beta} W(\frac{1}{\gamma}e^{\alpha I + \beta c - 1}) \]

concluding this proof. ■

We now analyze how the optimal expected profit changes with variations in the cost and price. We present first the following definition from [Ath98] which will be used in this analysis, and then we establish a proposition.

**Definition 25** Let \( \Pi(p, c) : \mathbb{R}^2 \to \mathbb{R} \). If \( \Pi \) is twice continuously differentiable, then \( \Pi \) has **increasing differences** if and only if for all \( (p, c) \), \( \frac{\partial^2 \Pi}{\partial p \partial c} \) \( (p, c) \geq 0 \), i.e., it holds if the incremental returns \( \Pi(p'', c) - \Pi(p', c) \) are nondecreasing in \( c \) for all \( p'' > p' \).

**Proposition 26** The expected profit \( \Pi(p, c) \) satisfies increasing differences.

**Proof.** This proof is straightforward, we just take derivative of (3.6) and show that

\[ \frac{\partial}{\partial c} \left( \frac{\partial \Pi}{\partial p} \right) = \frac{\partial}{\partial c} (q(\beta(1-q)(p-c)+1)) \]

\[ \frac{\partial^2 \Pi}{\partial c \partial p} = -\beta q(1-q) \]

\( \frac{\partial^2 \Pi}{\partial c \partial p} \geq 0 \) for all values of \( (p, c) \). From Definition 25 we can see that \( \Pi(p, c) \) satisfies increasing differences (see Figure 3-7 for a graphical interpretation). ■

It is also interesting to study how the optimal price in (3.8) changes with the cost \( c \). Figure 3-8 depicts this phenomenon. We can see that it converges to a constant distance from the cost (45-degree line), i.e. \( \frac{\partial p^*}{\partial c} = \frac{1}{1+W(\frac{1}{\gamma}e^{\alpha I + \beta c - 1})} \to 1 \) when \( c \) gets large, and \( \lim_{c \to \infty} (p^* - c) = -1/\beta \). Actually this constant gap is asymptotically reached very fast.

The optimal expected profit in (3.10) increases with changes in \( \beta \) (remember that \( \beta < 0 \)), i.e., the price has less negative impact on the demand. This can be seen in the following expression where \( z = \frac{1}{\gamma}e^{\alpha I + \beta c - 1} > 0 \).

\[ \frac{\partial \Pi^*}{\partial \beta} = \frac{1}{\beta^2} W(z) - \frac{1}{\beta} \frac{\partial W(z)}{\partial z} \frac{\partial z}{\partial \beta} \]

From Appendix B.2.1 we can get the derivative of \( W(z) \) and it is easy to prove that
\[
\frac{\partial z}{\partial \beta} = cz.
\]
Thus, we have
\[
\frac{\partial \Pi^*}{\partial \beta} = \frac{1}{\beta^2} W(z) - \frac{c}{\beta} \frac{W(z)}{1 + W(z)} \geq 0 \quad \forall z
\]
On the other hand, the optimal expected profit decreases with \( \gamma \), which means that when the utility from choosing competitors’ bundles increases then the profits will decrease. From (3.10) again we have that
\[
\frac{\partial \Pi^*}{\partial \gamma} = -\frac{1}{\beta} \frac{W(z)}{1 + W(z)} - \frac{z}{\beta} \frac{W(z)}{1 + W(z)}
\]
\[
= -\frac{1}{\beta} \frac{W(z)}{1 + W(z)} \gamma
\]
which is negative for all values of \( z \).

### 3.3.2 Bundle composition optimization

We now turn to a solution method of (BP) which does not enumerate all possible bundle configurations. Given that we have solved in closed-form the pricing optimization problem given a bundle configuration, we can now optimize \( \Pi^* \) over \( x \), the bundle configuration. As we mentioned before this is a NP-hard problem, which makes things more challenging.

We start by assessing the impact of the cost in the profit function. We summarize this result in the following proposition.

**Proposition 27** The profit function \( \Pi(p, c) \) is decreasing in the cost \( c \).
Proof. \( \frac{\partial \Pi}{\partial c} = -q < 0 \) for all values of \( p \).

Using proposition 27 we can detect dominance among the different bundles that could be offered by the company. At any given level of weighted attractiveness \( \alpha^\top l \), there is going to be a subset of bundles, where it is clear that the bundle with the lowest cost will dominate over the others. Thus, all dominated bundles can be excluded from \( \Omega \), reducing its cardinality dramatically. We summarize this idea in the following proposition which is very useful in our next analysis.

**Proposition 28 (Pareto Bundles)** If \( \alpha^\top l_1 \leq \alpha^\top l_2 \) and \( c_1 \geq c_2 \), bundle 2 is more profitable (it dominates over) than bundle 1.

**Proof.** At any cost level \( c_1 = c_2 \), the higher the attractiveness level the higher the profit. On the other hand, at the same weighted attractiveness level \( \alpha^\top l_1 = \alpha^\top l_2 \), the lower the cost the higher the profit (see proposition 27).

Proposition 28 allows us to find the Pareto frontier of all possible bundles, i.e., the set of all those bundles that cannot be dominated by other bundles in the set. Basically, for every level of attractiveness there is only one bundle that dominates the others in that level (the one with the lowest cost). Figure 3-9a shows all possible bundles in a specific instance\(^2\) of the problem, while Figure 3-9b shows just the dominant bundles (Pareto

---

\(^2\)For this instance of the problem, we consider 5 sets of components (e.g. processor, screen, hard drive, memory, and CD drive) with 3 choices in each one. The weights (\( \alpha \)) for the attractiveness factors are all equal to 1, and the \( \beta \) factor in the price response function is \(-0.007\). The \( \gamma \) factor which accounts for the competitors bundles is set equal to \(10 \times 10^{12}\). The attractiveness level for each component range from 4 to 9, so the range for the bundles will be from 20 to 45 (since \( \alpha = 1 \)).
frontier). Note that in this example there are two bundles which dominate within their attractiveness level but they are dominated by others in the Pareto set.

![Figure 3-9: Dominant (Pareto) Bundles](image)

Thus, the optimization problem to determine each Pareto bundle for each level of attractiveness \( \upsilon \) is given by

\[
(CSP_\upsilon) \quad \min_x \quad \sum_{j=1}^{m} \sum_{i \in S_j} c_{ij} x_{ij} \\
\text{subject to} \\
1 = \sum_{i \in S_j} x_{ij} \quad \forall j = 1, \ldots, m \\
\upsilon \leq \alpha^\top I \\
I^\top = (I_1, \ldots, I_m) \\
I_j = \sum_{i \in S_j} I_{ij} x_{ij} \quad \forall j = 1, \ldots, m \\
x_{ij} \in \{0, 1\} \quad \forall i, j
\]

The problem \((CSP_\upsilon)\) to determine the Pareto set corresponds to a constrained shortest path problem (due to the constraint in (3.11)) which has been extensively analyzed in the literature (e.g., White [Whi82], Warburton [War87], Beasley and Christofides [BC89], and Papadimitriou and Yannakakis [PY00]). Ahuja et al. [AMO93] show that the constrained shortest path problem is NP-complete by transforming the knapsack problem to it. Figure 3-10 depicts the network behind our CSP problem where each set of arcs between any two nodes are the components of a particular set from which only one is selected.

Given that we are dealing with an acyclic network, the Pareto frontier denoted by \( \Omega^* \) can be determined by means of the following pseudopolynomial-time algorithm.
Algorithm \textit{pareto_frontier};
begin
    i:=0;
    j:=1;
    A:=\{(0,0)\};
    while \(j \leq m\) do
        begin
            B:=set of pairs \((\alpha I, c)\) corresponding to arcs \((i, j)\);
            A:=\textit{cross_sum}(A, B);
            A:=\textit{keep_dominants}(A);
            i:=i+1;
            j:=j+1;
        end;
    pareto_frontier:=A;
end;

The subroutine \textit{cross-sum}(A, B) returns as an output the set of all possible sums among elements of \(A\) and \(B\), while the subroutine \textit{keep-dominants}(A) eliminates from the set \(A\) all those elements which are dominated by other elements in \(A\). The overall complexity of the \textit{pareto-frontier} algorithm is \(O(nm^2\hat{I} \log(nm\hat{I}))\) or \(O^*(nm^2\hat{I})\), where \(m\) is the number of sets, \(n\) is the maximum number of components of a set, and \(\hat{I}\) is the maximum level of attractiveness of any component. This is because the algorithm has \(m\) stages where in each one a set of at most \(nm\hat{I}\) elements has to be sorted and cleaned from non-Pareto elements which has complexity \(O(nm\hat{I} \log(nm\hat{I}))\).

Moreover, let us consider that the objective function \(\Pi^*\) in (3.10) is a function of the bundle’s attractiveness level \(I = \alpha^\top I\) and the bundle’s cost \(c\), since every possible bundle is characterized by these two measures.

\textbf{Proposition 29} The gradient of the objective function \(\Pi^*(I, c)\) is a constant direction times a variable magnitude, and it is given by

\[
\nabla \Pi^*(I, c) = \left( -\frac{1}{\beta}, -1 \right) q^*.
\]
**Proof.** It is quite straightforward. To determine the gradient of $\Pi^*(I, c)$ we take first derivatives as follows:

\[
\frac{\partial \Pi^*(I, c)}{\partial I} = -\frac{1}{\beta} \frac{\partial W \left( \frac{1}{\gamma} e^{I+\beta c-1} \right)}{\partial I} = \frac{-1}{\beta} \frac{W(\frac{1}{\gamma} e^{I+\beta c-1})}{W(\frac{1}{\gamma} e^{I+\beta c-1}) + 1}
\]

\[
\frac{\partial \Pi^*(I, c)}{\partial c} = -\frac{1}{\beta} \frac{\partial W \left( \frac{1}{\gamma} e^{I+\beta c-1} \right)}{\partial I} \frac{\beta}{\gamma} e^{I+\beta c-1} = -\frac{W(\frac{1}{\gamma} e^{I+\beta c-1})}{W(\frac{1}{\gamma} e^{I+\beta c-1}) + 1}
\]

Therefore,

\[
\nabla \Pi^*(I, c) = \left( -\frac{1}{\beta}, -1 \right) q^*
\]

where $q^* = \frac{W(\frac{1}{\gamma} e^{I+\beta c-1})}{W(\frac{1}{\gamma} e^{I+\beta c-1}) + 1}$ from the proof of Corollary 24. ■

---

**Figure 3-11: Pre-optimized objective function**

The result in Proposition 29 is quite interesting. We can visualize it in Figure 3-11. With this proposition our problem is reduced to the maximization of a linear function over the bundles in the Pareto set $\Omega^*$, i.e.,

\[
\max_{(I, c) \in \Omega^*} \frac{-1}{\beta} I - c.
\]

Figure 3-12 shows the optimal bundle in the example we have been using through this section.
3.4 Extensions and further research

3.4.1 Control over multiple bundles

An interesting extension would be to consider the case when the company has to price and compose multiple bundles, which will compete alongside competitors’ products. By extending the problem formulation in (3.3) to the multiple bundles case we get the following formulation:

\[
\begin{align*}
\text{(BP)} & \quad \max_{(p_1, \ldots, p_B, x_1, \ldots, x_B)} \sum_{k=1}^{B} q_k(p_k, x_k)(p_k - c_k(x_k)) \\
\text{subject to} & \quad (3.12)
\end{align*}
\]

\[
\begin{align*}
q_k &= \frac{e^{\alpha \mathbf{l}_k^\top + \beta p_k}}{\gamma + \sum_{l=1}^{B} e^{\alpha \mathbf{l}_l^\top + \beta p_l}} \\
\mathbf{l}_k^\top &= (I_1^k, \ldots, I_m^k) \quad \forall k = 1, \ldots, B \\
I_j^k &= \sum_{i \in S_j} I_{ij} x_{ijk} \quad \forall j = 1, \ldots, m; \quad \forall k = 1, \ldots, B \\
c_k &= \sum_{j=1}^{m} \sum_{i \in S_j} c_{ij} x_{ijk} \\
1 &= \sum_{i \in S_j} x_{ijk} \quad \forall j = 1, \ldots, m; \quad \forall k = 1, \ldots, B \\
x_{ijk} &\in \{0, 1\} \quad \forall i, j, k \\
p_k &\geq 0 \quad \forall k = 1, \ldots, B
\end{align*}
\]
Zhang et al. [ZBGT02] (an unpublished document from our sponsor company) have developed an interesting and curious result with regard to the optimal profit margins (given a set of bundles). Actually, this is similar to a more general case analyzed in the previous chapter of this dissertation. This result is formalized in the following proposition.

Before that, we state the objective function of the pricing optimization problem as follows:

\[
\Pi = \max_{(p_1, \ldots, p_B) \geq 0} \sum_{k=1}^{B} \frac{e^{\alpha^T I_k + \beta p_k}}{\gamma + \sum_{l=1}^{B} e^{\alpha^T I_l + \beta p_l}} (p_k - c_k(x_k))
\]  

(3.14)

**Proposition 30** In optimality, all profit margins are the same for all bundles, i.e., \( p_k - c_k = p_l - c_l \) for all \( l \neq k \).

**Proof.** Let us compute the first derivative of the profit function in (3.14) as

\[
\frac{\partial \Pi}{\partial p_k} = \frac{e^{\alpha^T I_k + \beta p_k} \left[ (1 + \beta p_k - \beta c_k) \left( \gamma + \sum_{l=1}^{B} e^{\alpha^T I_l + \beta p_l} \right) - \beta \sum_{l=1}^{B} (p_l - c_l) e^{\alpha^T I_l + \beta p_l} \right]}{\left( \gamma + \sum_{l=1}^{B} e^{\alpha^T I_l + \beta p_l} \right)^2},
\]

for all \( k = 1, \ldots, B \). By making the above derivative equal to zero we get the first order condition of this problem, i.e., \( \frac{\partial \Pi}{\partial p_k} = 0 \) yields

\[
(1 + \beta p_k - \beta c_k) \left( \gamma + \sum_{l=1}^{B} e^{\alpha^T I_l + \beta p_l} \right) - \beta \sum_{l=1}^{B} (p_l - c_l) e^{\alpha^T I_l + \beta p_l} = 0.
\]

(3.15)

The analog condition for bundle \( j \) is

\[
(1 + \beta p_j - \beta c_j) \left( \gamma + \sum_{l=1}^{B} e^{\alpha^T I_l + \beta p_l} \right) - \beta \sum_{l=1}^{B} (p_l - c_l) e^{\alpha^T I_l + \beta p_l} = 0.
\]

(3.16)

By subtracting (3.16) from (3.15) and dividing by \( \beta \), we obtain

\[
(p_k - c_k - (p_j - c_j)) \left( \gamma + \sum_{l=1}^{B} e^{\alpha^T I_l + \beta p_l} \right) = 0.
\]

Since \( \gamma + \sum_{l=1}^{B} e^{\alpha^T I_l + \beta p_l} > 0 \) the proposition follows. \( \blacksquare \)

Moreover, we can use our methodology from the single bundle case to express the prices in closed-form solution. Let us formalize these closed-forms in the following proposition.

**Proposition 31** From the first order conditions in (3.15) we can get the optimal prices \( p_k^* \) in closed-form as follows:

\[
p_k^* = c_k - \frac{1}{\beta} \left( 1 + W \left( \frac{1}{\gamma} \sum_{l=1}^{B} e^{\alpha^T I_l + \beta c_l - 1} \right) \right) \quad \forall k = 1, \ldots, B
\]

(3.17)

where \( W(\cdot) \) is the Lambert W function\(^3\). Moreover, these conditions are sufficient and the

---

\(^3\)The Lambert W function is the solution of \( W(z)e^{W(z)} = z \). Definitions and properties of this function

107
optimal prices are unique.

**Proof.** In order to prove this proposition we have to find an expression of the form of $W(z)e^{W(z)} = z$, the Lambert function definition. Let $r$ be the common profit margin, i.e., $r = p_k - c_k$ for all $k$. Now from (3.15) we have

$$(1 + \beta r) \left( \gamma + \sum_{l=1}^{B} e^{\alpha^T l_1 + \beta r + \beta c_l} \right) - \beta \sum_{l=1}^{B} r e^{\alpha^T l_1 + \beta r + \beta c_l} = 0.$$  

After a few rearrangements, we obtain

$$
(1 + \beta r) \gamma + \sum_{l=1}^{B} e^{\alpha^T l_1 + \beta r + \beta c_l} = 0 \\
(1 + \beta r) \gamma + e^{\beta r} \sum_{l=1}^{B} e^{\alpha^T l_1 + \beta c_l} = 0 \\
(1 + \beta r) \gamma + e^{\beta r} C = 0,
$$

where $C = \sum_{l=1}^{B} e^{\alpha^T l_1 + \beta c_l}$ is constant for a given set of bundles. Using the functional form of the Lambert function definition, we derive

$$
(-1 - \beta r) \gamma = e^{\beta r} C \\
(-1 - \beta r) e^{-\beta r} = \frac{C}{\gamma} \\
(-1 - \beta r) e^{-(-1 - \beta r)} = \frac{Ce^{-1}}{\gamma}.
$$

Thus, we can infer that $W(z) = (-1 - \beta r)$ and $z = \frac{Ce^{-1}}{\gamma}$. Then, solving for the optimal prices, we get

$$
-1 - \beta r = W\left(\frac{1}{\gamma} Ce^{-1}\right) \\
r = \frac{-1}{\beta} \left(1 + W\left(\frac{1}{\gamma} Ce^{-1}\right)\right) \\
p^*_k = c_k - \frac{1}{\beta} \left(1 + W\left(\frac{1}{\gamma} \sum_{l=1}^{B} e^{\alpha^T l_1 + \beta c_l - 1}\right)\right) \quad \forall k = 1, ..., B.
$$

With regard to uniqueness, we can easily verify for the case of linear utility functions, that the objective function is the sum of $B$ unimodal or quasiconcave functions. Therefore, a unique finite price is guaranteed for each of the $B$ bundles.

The composition optimization part (solve for $x_k$) of this multiple bundle case is quite complex. Note that the optimal bundle of the single-bundle case will not necessarily be part of the optimal set in this multiple-bundle case.

---

*that are relevant for our case are presented in Appendix B.2. Refer to Corless et al. [CGH+95] for a thorough analysis and understanding of this special function.*
3.4.2 Pre-configured bundles with rebates

In this model we consider the setting when customers arrive at a store and can decide whether to buy the bundle or to customize their own. As is common when using a bundling strategy, a discount is associated with buying the pre-configured bundle in order to provide incentives to the customers. Figure 3-13 on the left depicts the reservation price probability distribution at every possible attractiveness level. We assume that bundles are ordered, i.e., a customer who wants a bundle with attractiveness $I'$ will also be happy with any bundle whose attractiveness is greater than or equal to $I'$.

![Figure 3-13: Profit maximization with rebate](image)

The question now is how to find a bundle composition, bundle price and rebate amount such that profits are maximized. Here the selected bundle will have a saving of $A$ over its cost per unit $c_b$, due to economies of scale. The company wonders if it is profitable to transfer part of these savings, let us say $\alpha A$ (where $0 \leq \alpha \leq 1$), to customers who select this pre-configured bundle through a rebate or through giving an additional product with the bundle (i.e., a laptop plus a free printer). On the right of Figure 3-13 we can see all Pareto bundles plotted at their regular cost (without considering economies of scale). The shaded area with diagonal lines represents all customers who need a bundle with an attractiveness level less than or equal to $I^*$ and also are willing to pay $p^*$.

The profit margin per customer ($PMC$) would be given by

\[
PMC = p - (c_b - A) - \alpha A = p - c_b + (1 - \alpha)A,
\]

and the market share ($MS$) is computed by aggregating all customers who meet the following criteria: those who are willing to pay $p$ and their needs are being fulfilled buy
the selected bundle. Thus, the market share would be given by

\[ MS = \sum_{k=0}^{I} \int_{p}^{\infty} f(\eta, k) d\eta. \]

Therefore, the total expected profits are

\[ \Pi = (p - c_b + (1 - \alpha)A) \sum_{k=0}^{I} \int_{p}^{\infty} f(\eta, k) d\eta, \]

which will be maximized with respect to \( p \), \( I \), and \( \alpha \).

### 3.4.3 Regular bundle and an add-on component

The fact that some customers look for additional accessories, such as printers, makes a natural division among the customer population. Given a price for the bundle and the extra fee for getting the accessory, there are going to be three groups of customers: 

(i) those who do not buy anything; 
(ii) those who buy the bundle; and 
(iii) those who consider the bundle with the extra accessory. Thus, these groups will be characterized by reservation prices, which is their willingness to pay. Customers with reservation prices lower than the bundle price \( p \) do not buy anything, those with reservation prices between the price \( p \) and the price plus the accessory fee \( (p + K) \) buy the bundle, and those with reservation prices higher than the price plus the accessory fee buy either the bundle or the customized product (see Figure 3-14). We assume that the distribution of the reservation prices is invariant with time because the selling period considered here is the short life cycle of these kinds of products (e.g., 3 months for laptops).

![Figure 3-14: Arrival process - Bundle vs. Bundle with accessory](image)

From Figure 3-14, we can realize that we split those customers willing to pay more
than the price $p$ plus the accessory fee $K$ using the probabilities $\beta$ and $(1-\beta)$. In the remainder of the chapter, we assume that $\beta$ is constant.

We also assume that the customers’ arrival is described by a Poisson process with an arrival rate that is a function of general purchasing patterns rather than a function of the specific price $p$ of the bundle under consideration. Since after each customer arrives he/she will be part of one of the three groups mentioned above depending upon the probabilities in Figure 3-14, it follows that each group has its own independent Poisson arrival process$^4$. Thus, the arrival rates for not buying anything is denoted by $\lambda_A$, for buying the bundle is denoted by $\lambda_B$, and for buying the bundle plus the accessory is denoted by $\lambda_C$, and are given by the following expressions:

$$
\begin{align*}
\lambda_A &= \lambda F(p) \\
\lambda_B &= \lambda(1-\beta)(1-F(p+K)) - \lambda F(p) + \lambda \beta \\
\lambda_C &= \lambda(1-\beta)(1-F(p+K)).
\end{align*}
$$

Given that the expected number of customers is the arrival rate of the corresponding process (this is true in the case of Poisson processes), the total expected revenue is given by

$$
\Pi = p\lambda_B + (p + K)\lambda_C
$$

Rewriting the profit function in terms of $\beta$ yields a linear function which decreases with $\beta$, i.e., $\Pi = a - b\beta$ where $a = p\lambda(1-F(p)) + K\lambda(1-F(p+K))$ and $b = K\lambda(1-F(p+K))$. This result suggests that it is never optimal for customers willing to pay for the bundle plus the accessory to choose only the bundle.

Given a price $p$, what would be the optimal value of the accessory fee $K$ in order to maximize the expected profits $\Pi$?

Taking derivative to $\Pi$ with respect to $K$ yields the first order conditions:

$$
\frac{\partial \Pi}{\partial K} = \frac{\partial}{\partial K}(p(\lambda_B + \lambda_C) + K\lambda_C) = \frac{\partial}{\partial K}(p\lambda(1-F(p)) + K\lambda(1-\beta)(1-F(p+K))) = \lambda(1-\beta)(1-F(p+K)) - K\lambda(1-\beta)f(p+K).
$$

Letting $\frac{\partial \Pi}{\partial K} = 0$ we find that the optimal accessory fee $K^*$ is

$$
\lambda(1-\beta)(1-F(p+K^*)) - K^*\lambda(1-\beta)f(p+K^*) = 0
$$

$$
(1-F(p+K^*)) - K^*f(p+K^*) = 0
$$

$^4$A classical result in probability theory is that the two split or resulting processes are independent despite the fact that the split processes arise from the same original Poisson process.
as long as $\lambda \neq 0$ and $\beta \neq 1$. Thus, rearranging some terms

$$K^* = \frac{1 - F(p + K^*)}{f(p + K^*)}.$$  \hfill (3.18)

The expression obtained in (3.18) could be solved analytically for some special cases of a few distributions of the reservation prices, but certainly can be easily solved numerically (fixed point solution).

Next, consider the first order conditions with respect to the price:

$$\frac{\partial \Pi}{\partial p} = \frac{\partial}{\partial p} \left( p(\lambda_B + \lambda_C) + K\lambda_C \right)$$

$$= \lambda(1 - F(p)) - p\lambda f(p) - K\lambda(1 - \beta)f(p + K) = 0,$$

which implies that

$$p^* = \frac{1 - F(p^*)}{f(p^*)} - \frac{K(1 - \beta)f(p^* + K)}{f(p^*)}$$  \hfill (3.19)

Taking second derivatives is not helpful because it is not clear if the derivatives of the density functions are positive or negative. We can find a solution by iterating numerically with equations (3.18) and (3.19).

### 3.5 Summary

How does a company determine the composition and the price of a bundle so as to maximize total expected profits? The setting in this research is a high-tech company in a highly competitive environment. Bundles are built from a set of components that must meet certain technical constraints. The company’s objective is to build a bundle and offer it in a market where it will compete with other substitute products. Consumers will purchase the bundle that maximizes their utility after examining all available bundles. Therefore, the selection of a bundle’s components and its price is made in light of the bundles with which it will be competing in the market and the uncertainty in the consumer choice process.

The optimal decision can be found by solving a nonlinear mixed integer programming model. These models have been proven to be NP hard. Demand is a function of the bundle price and a vector of attributes, and it is modeled using a multinomial logit model (MNL). The profit per sold bundle is the difference between the price and the bundle cost. Each bundle is characterized by its cost and its attractiveness level (computed as a weighted average of the individual component’s attractiveness).

If demand is modeled with MNL, then, for a given bundle, the objective function is unimodal (quasi-concave). In this case, there is a unique optimal price which can be determined in closed-form. The optimal marginal profit per customer converges to $\frac{-1}{\beta}$,
where $\beta$ is the price response parameter (a negative constant) as the cost goes to infinity. This convergence is very fast due to the MNL.

This model can be extended to the case where there are multiple bundles, in which case the result is that the profit (price minus cost) is constant across all bundles. This result is analogous to the optimality conditions found in the previous chapter for the one-step and the two-step decision process models.

The fact that we found a closed-form solution for the optimal price is very important. This result can be substituted into the original formulation. A pre-optimized model can then be solved in order to find the optimal bundle configuration.

We prove that for each attractiveness level, the bundle with the smallest cost dominates all other bundles. This allows us to find a Pareto frontier, which is a set composed of the dominating bundle for each attractiveness level. This frontier can be determined with an efficient algorithm that we specified. It turns out that the profit function has a constant gradient’s direction, which means that the problem of finding the optimal composition reduces to a solvable integer program with a linear objective function being optimized over this Pareto frontier set.
Chapter 4

Selling Bandwidth Capacity Bundles on a Network

This chapter addresses the problems of capacity allocation and pricing in telecommunication networks. The networks analyzed consist of several regions whose connecting links have a fixed bandwidth capacity. Suppose that point A is connected to point B through link (AB), and point B is also connected to point C through link (BC). The capacity allocation problem consists of deciding how much bandwidth to allocate to consumers who wish to transfer data between A and C and how much bandwidth to allocate to those who wish to transfer data between A and B or B and C. Since transferring data between A and C involves allocating capacity between (AB) and (BC), this problem is intimately related to that of making bundling decisions. Should companies sell capacity in (AB) and (BC) separately or as a bundle?

The pricing problem comes from the fact that bandwidth is an intangible commodity and has to be sold in advance. This is a problem of yield management. Companies must make an initial investment in infrastructure, and what is not sold by the time of delivery is lost. The decision problem faced by managers in this industry is analogous to the problem faced in the airline industry, where an empty seat at the time of a flight’s departure is lost revenue.

Suppose that a network owner is selling capacity to be delivered $T+1$ periods from now. The decision problem is to decide how much to sell from the present period to period $T$. In this chapter we consider the case where the network owner faces an external demand for each market or product (capacity between two cities) that can be fulfilled by assigning capacity of the corresponding arc (direct link) or by assigning capacity of an alternative path (indirect link), incurring a higher cost. This problem differs from classical routing problems faced in the transportation industry in that customers do not distinguish between routings. They only care about communicating between two specific cities and the Quality of Service (QoS). The goal is to maximize the network owner’s revenues while deciding on each period whether to bundle links (for a demand from node
4.1 Capacity allocation with known demand

We start analyzing revenue and profit maximization in the competitive market case, where the company is a price taker, and then we move to the monopolist case. Profit maximization takes into account the operating costs in the objective function. It should impact the decision of using alternative paths to fulfill demand when a certain link is scarce, as long as the variable cost is significant with respect to the price. If we are dealing with very high fixed costs and negligible variable costs, it does not make sense to deal with a profit maximization model. In the competitive case, prices are exogenous rather than in the monopolistic case when prices are decision variables.

In order to gain insights into the real problem, it would be interesting to analyze the deterministic case of simple networks. In this section we consider the cases of an inline network, and a one-loop network.

Let us consider the following notation:

\( T \): Number of selling periods
\( C_{ij0} \): Initial available capacity in arc \((i, j)\)
\( C_{ijt} \): Available capacity in arc \((i, j)\) in period \(t\)
\( P_{ijt} \): Price of one unit of capacity in arc \((i, j)\) in period \(t\)
\( D_{ijt} \): Units of arc \((i, j)\) capacity demanded in period \(t\) (can be 0 or 1)
\( x_{ijt} \): Units of arc \((i, j)\) capacity sold in period \(t\)
\( y_{ijt} \): Units of arc \((i, j)\) capacity sold in period \(t\) (satisfied using an alternative path)

Since we are dealing with integer variables, MIP formulations will not be trivial to solve for real size problems due to the large number of variables. For that reason it is important to figure out whether our formulations fulfill the unimodularity property or not. Let’s review that property first.

### 4.1.1 Unimodularity property

In this section we show the relationships between the integrality property (Definition 32) and certain integrality results (Definition 33) in linear programming.
Definition 32 A nonempty polyhedron $P = \{ x \in \mathbb{R}^n : Ax \leq b \}$ with $\text{rank}(A) = n$ is integral if and only if all of its extreme points are integral.

Proving polyhedra are integral is often a difficult task. Proofs usually look for matching problems. This has led to the development of various sufficient conditions for integrality, stating that if $A$ and $b$ satisfy certain properties, then we know $\{ x : Ax \leq b \}$ is integral. In terms of the number of applications, perhaps the most successful of these integrality conditions is the one developed by Hoffman and Kruskal [HK56], involving the determinants of submatrices of $A$ (conditions stated in definition 33 and theorem 34).

Definition 33 An $m \times n$ integral matrix $A$ is **totally unimodular** (TU) if all of its square submatrices have determinant 0, 1, or -1. In particular, every entry in a totally unimodular matrix is 0, 1, or -1.

Theorem 34 (Hoffman-Kruskal Theorem) Let $A$ be an $m \times n$ integral matrix. Then the polyhedron defined by $Ax \leq b$, $x \geq 0$ is integral for every integral vector $b \in \mathbb{R}^m$ if and only if $A$ is totally unimodular.

Proof. See Cook et al. [CCPS98] theorem 6.25

The following two propositions are good characterizations of the totally unimodular property.

Proposition 35 The following statements are equivalent.

1. $A$ is TU.
2. The transpose of $A$ is TU.
3. $(A, I)$ is TU.
4. A matrix obtained by deleting a row (column) of $A$ is TU.
5. A matrix obtained by multiplying a row (column) of $A$ by $-1$ is TU.
6. A matrix obtained by interchanging two rows (columns) of $A$ is TU.
7. A matrix obtained by duplicating columns (rows) of $A$ is TU.
8. A matrix obtained by a pivot operation on $A$ is TU.


Proposition 36 If $A$ is TU, if $b$, $b'$, $d$, and $d'$ are integral, and if $P(b, b', d, d') = \{ x \in \mathbb{R}^n : b' \leq Ax \leq b, d' \leq x \leq d \}$ is not empty, then $P(b, b', d, d')$ is an integral polyhedron.
Proof. See Nemhauser and Wolsey [NW88] generalization of Proposition 2.2.

**Definition 37** An $m \times n$ $(0,1)$ matrix $A$ is called an interval matrix if in each column the 1’s appear consecutively; that is, if $a_{ij} = a_{kj} = 1$ and $k > i + 1$, then $a_{lj} = 1$ for all $l$ with $i < l < k$.

**Corollary 38** Interval matrices are TU.

Proof. See Nemhauser and Wolsey [NW88] Corollary 2.10.

These last two statements (Definition 37 and Corollary 38) are very important in order to characterize the solutions of our model formulations.

### 4.1.2 Inline network case

We first start with the so-called inline network, which is also called “linear network” in Bertsimas and Popescu [BP01]. In this type of network there is only one possible path between any pair of nodes, and every node has at most two adjacent nodes. Figure 4-1 depicts the structure of the network. Since both the demand (orders) and the prices are given, the problem reduces to the decision of which orders will be accepted.

![Inline Network of n-1 links](image)

Figure 4-1: Inline Network of n-1 links

A good starting point for getting insights is to consider the 4-node inline network $G(N, L_I)$ where $N = \{1, 2, 3, 4\}$ is the set of nodes and $L_I = \{(1, 2); (2, 3); (3, 4)\}$ is the set of all arcs where inventory is available. Let $L_D = \{(1, 2); (2, 3); (3, 4); (1, 3); (2, 4); (1, 4)\}$ be the set of all pairs of cities for which there is demand for the available bandwidth capacity between them ($L_I \subset L_D$). The formulation of this problem is the following:

$$\max_x \sum_{(i,j) \in L_D} \sum_{t=1}^T x_{ijt} P_{ijt} D_{ijt}$$
subject to the following constraints:

\[
\sum_{t=1}^{T} (x_{12t} + x_{13t} + x_{14t}) \leq C_{12}
\]

\[
\sum_{t=1}^{T} (x_{23t} + x_{13t} + x_{24t} + x_{14t}) \leq C_{23}
\]

\[
\sum_{t=1}^{T} (x_{34t} + x_{24t} + x_{14t}) \leq C_{34}
\]

\[x_{ijt} \in \{0, 1\} \quad \forall (i, j) \in \mathcal{L} \text{ and } t = 1, ..., T\]

The objective function represents total revenues, and each of the first three constraints takes care of the available capacity in each link in \(\mathcal{L}\). This problem seems to be simple, but its combinatorial nature (last constraint) makes it infeasible to be solved for a large \(T\). However, the following proposition shows that a related problem can be solved in order to find the optimal integer solution.

**Proposition 39** Matrix \(A\) in the previous formulation is totally unimodular, so any feasible solution to the problem is integral. By solving the linear programming relaxation the optimal solution is attained.

**Proof.** Matrix \(A\) (now \(A_3\)), in this 4-node case and for \(T = 2\) would be represented by

\[
A_3 = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

and vector \(x = [x_{121}, x_{122}, x_{231}, x_{232}, x_{341}, x_{342}, x_{131}, x_{132}, x_{241}, x_{242}, x_{141}, x_{142}]\), which represents the decision variables of the problem, shows the order of columns in \(A\).

From proposition (35) part 7 we have that if \(A_3\) is TU, then the matrix that results from eliminating all duplicating columns would also be TU. In addition, from the same proposition, part 3, we have that by eliminating the columns that form an identity, total unimodularity is preserved. Finally, we just need to prove that matrix \(\tilde{A}_3\) is TU.

\[
\tilde{A}_3 = \begin{bmatrix}
1 & 0 & 1 \\
1 & 1 & 1 \\
0 & 1 & 1
\end{bmatrix}
\]

From definition 37, it follows that \(\tilde{A}_3\) is an interval matrix. Using corollary 38, this matrix is also TU. Therefore, matrix \(A_3\) is TU.

Since network capacities \(C_{ij}\) are integers, using the Hoffman-Kruskal theorem (34) we can now conclude that the solution of a linear relaxation problem (by relaxing integer
constraints) is integral.

The generalization of this representation is straightforward. The following proposition generalizes the previous result for a \((n+1)\)-node inline network case.

**Proposition 40** Matrix \(A_n\) is TU. Therefore by solving the linear relaxation problem an integral solution is attained.

**Proof.** Let \(I_n\) be the \(n\)-dimensional identity matrix, and let \(0_n\) be a null vector whose components are equal to zero. Thus, matrix \(A_n\) for one single period has the following structure:

\[
A_n = \left[ I_n \left| \begin{array}{cccc}
I_{n-1} & 0_{n-1} & \cdots & 0_{n-1} \\
0_{n-1} & I_{n-1} & \cdots & 0_{n-1} \\
\vdots & \vdots & \ddots & \vdots \\
0_{n-1} & 0_{n-1} & \cdots & I_{n-1} \\
\end{array} \right| \right]
\]

where every expression between bars \(|·|\) represents an \(n \times i\) matrix, where \(i\) goes from \(n\) to 1. It is easy to see that the total number of columns in matrix \(A_n\) is \(\frac{n(n+1)}{2}\).

By applying the same propositions used in the previous proof, and by noticing that \(A_n\) is an interval matrix, we conclude this proof.

According to proposition (40), the deterministic \((n+1)\)-node inline network problem can be solved by finding the optimal solution to the following linear programming relaxation problem:

\[
\max \sum_{t=1}^{T} x_t \cdot r_t \quad \text{where} \quad x_t \text{ and } r_t \text{ are vectors for all } t, \text{ and each component of } r_t \text{ represents the product } P_{ijt}D_{ijt} \text{ component, subject to the constraints } \sum_{t=1}^{T} A_n x_t \leq C \\
\text{and } 0 \leq x_{ijt} \leq 1 \forall (i, j) \in L_D \text{ and } t = 1, \ldots, T \text{ where } C \text{ is the vector containing the} \\
\text{capacity of each link in } L_I.
\]

**4.1.3 One-loop network case**

In what follows we explain a second case where the network configuration consists of \(n\) connected nodes forming a single loop (i.e., a ring). This is the so-called one-loop network (see Figure 4-2).

![Figure 4-2: One-loop Network of \(n\) links](image)

This type of network has an additional and very relevant characteristic. Each demand can be fulfilled by two different routes or paths in the network, e.g., demand from point
1 to point $n$ can be fulfilled by the single link $(n, 1)$ or by the alternative path $(1, 2) - (2, 3) - ... - (n - 1, n)$. This feature makes this yield management problem very different from those in the literature. Here, even if there is no available capacity between two points, there is a chance of fulfilling a particular request using an alternative path.

In order to understand the structure of this case, let us consider a 3-node network $G(N, L_I)$ such that the arcs form a triangle, i.e., a network of three cities with a certain bandwidth capacity between each pair of cities, where $N = \{1, 2, 3\}$ are the nodes and $L_I = \{(1, 2); (2, 3); (3, 1)\}$ are the arcs or links where capacity is available. Note that in this particular case of $n = 3$, $L_I = L_D$.

It is important that prices are set in such a way that no arbitrage opportunity is available, i.e. for every link there is no alternative path with a lower price than the direct link. For that reason prices must satisfy some arbitrage-free constraints (network triangular inequalities). For case $n = 3$ they would be:

$$P_{12t} \leq P_{23t} + P_{31t} \forall t = 1, ..., T$$

$$P_{23t} \leq P_{12t} + P_{31t} \forall t = 1, ..., T$$

$$P_{31t} \leq P_{12t} + P_{23t} \forall t = 1, ..., T$$

Let us start by writing the linear programming model that determines when and where to sell the available capacity for the $(T + 1)^{th}$ period while maximizing revenues. Let $V_t$ be the revenue in period $t$, so the formulation would be

$$\max_{x, y} \sum_{t=1}^{T} V_t$$

subject to

$$V_t = [(x_{12t} + y_{12t})P_{12t}] + [(x_{23t} + y_{23t})P_{23t}] + [(x_{31t} + y_{31t})P_{31t}] \forall t$$

$$x_{12t} + y_{23t} + y_{31t} \leq C_{12t-1} \forall t = 1, ..., T$$

$$y_{12t} + x_{23t} + y_{31t} \leq C_{23t-1} \forall t = 1, ..., T$$

$$y_{12t} + y_{23t} + x_{31t} \leq C_{31t-1} \forall t = 1, ..., T$$

$$x_{ijt}, y_{ijt} \in \{0, 1\} \forall (i, j) \in L_I \text{ and } t = 1, ..., T$$

After some straightforward algebra, combining equations (4.2) and (4.4) we end up
with the following linear programming formulation in standard form. For notation simplicity we replace the initial capacity of arc $k, C_{k0}$, with $C_k$.

$$\max_{x,y} \sum_{t=1}^{T} V_t(x,y)$$

subject to

$$V_t(x,y) = \sum_{(i,j) \in L_D} [(x_{ijt} + y_{ijt}) P_{ijt}] \text{ for } t = 1, ..., T$$

$$\sum_{t=1}^{T} (x_{12t} + y_{23t} + y_{31t}) = 0 \quad C_{12}$$

$$\sum_{t=1}^{T} (y_{12t} + x_{23t} + y_{31t}) = 0 \quad C_{23}$$

$$\sum_{t=1}^{T} (y_{12t} + y_{23t} + x_{31t}) = 0 \quad C_{31}$$

$$x_{ijt} + y_{ijt} = D_{ijt} \quad \forall (i,j) \in L_I \text{ and } t = 1, ..., T \quad (4.5)$$

$$x_{ijt}, y_{ijt} \in \{0, 1\} \quad \forall (i,j) \in L_I \text{ and } t = 1, ..., T \quad (4.6)$$

The corresponding matrix $A_3$ from the previous formulation for just one period ($T = 1$) is given by

$$A_3 = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix}.$$ 

Unfortunately, matrix $A_3$ is not totally unimodular. Therefore, in this case, we cannot evade the combinatorial complexity for large scale problems as we did in the inline case. Given that in this case we are forced to write explicitly the demand constraints (4.5), we could allow demands to be greater than one\(^1\) so decision variables $x_{ijt}$ and $y_{ijt}$ will be integers instead of binary variables. This reduces the dimension of the problem significantly but the combinatorial complexity still remains.

4.1.4 Extension: Star network case

A third type of network topology could be analyzed. This is the case of the so-called *star network* (see Figure 4-3). As we can see in the figure, this network has a special feature similar to one in the inline case. It is that between any two pair of nodes there is only one possible path. Again, the number of all possible demands is given by $\frac{n(n-1)}{2}$, where $n$

---

\(^1\)In the inline network case demands $D_{ijt}$ were 0 or 1 and there were no alternative paths, so the demands constraints (4.5) were implicit in the objective function.
is the number of nodes.

![Star Network of (n – 1) links](image)

**Figure 4-3: Star Network of (n – 1) links**

### 4.2 Uncertain demand

In this section we concentrate on the monopolistic case where prices are endogenous, capacities are known, and the duration of the contract is fixed, e.g., one month. We first analyze the spot pricing problem for the inline and one-loop network cases, and then, we solve the multi-period pricing model for the same two cases. For the sake of simplicity and mathematical tractability the idea is to work with the smallest number of cities (i.e., 3 nodes).

![Network topologies for 3 nodes](image)

**Figure 4-4: Network topologies for 3 nodes**

The owner of the network faces three different demands. There is demand for bandwidth between nodes 1 and 2, between nodes 2 and 3, and between nodes 1 and 3. Since the products (the bandwidth between cities) are not substitutes, we assume that there is no correlation among demands. We also assume that customers do not know the network topology. In the 3-node case there are only two possible topologies which are shown in Figure 4-4. It would be very interesting to compare these two cases in order to see the difference and to be able to recommend a network expansion or reduction depending on the situation. This comparison will allow us to realize whether it is worth having direct capacity between nodes 1 and 3.
4.2.1 Capacity allocation

In this section we consider stochastic demands denoted by $\tilde{w}_{ij}$, for each pair of nodes $(i, j)$ in our bandwidth network. Prices $P_{ij}$ are fixed and do not change over time.

Before establishing the allocation algorithm, we first define the “Marginal Contribution” concept for the inline network.

**Definition 41** The Marginal Contribution $\Delta_{ij}$ of accepting a request for a unit of bandwidth between node $i$ and node $j$, corresponds to the expected or “actual” price of bandwidth between $i$ and $j$, and it is computed as follows:

$$\Delta_{ij}(x_{ij}) = P_{ij} \left( E[\tilde{w}_{ij} \min\{x_{ij} + 1, \tilde{w}_{ij}\}] - E[\tilde{w}_{ij} \min\{x_{ij}, \tilde{w}_{ij}\}] \right)$$

where $x_{ij}$ is the already reserved capacity for demand between $i$ and $j$.

**Proposition 42** The marginal contribution $\Delta_{ij}$ is proportional to the probability that the demand exceed the current reserved capacity $x_{ij}$, i.e.,

$$\Delta_{ij}(x_{ij}) = P_{ij} \Pr(\tilde{w}_{ij} \geq x_{ij} + 1).$$

**Proof.** Starting from Definition 41, by taking the expected values the proof is quite straightforward. Let $\varphi_{\tilde{w}_{ij}}(k)$ be the probability that the demand $\tilde{w}_{ij}$ is exactly $k$, i.e., $\varphi_{\tilde{w}_{ij}}(k) = \Pr(\tilde{w}_{ij} = k)$.

$$\Delta_{ij}(x_{ij}) = P_{ij} \left( E[\tilde{w}_{ij} \min\{x_{ij} + 1, \tilde{w}_{ij}\}] - E[\tilde{w}_{ij} \min\{x_{ij}, \tilde{w}_{ij}\}] \right)$$

$$= P_{ij} \left( \sum_{k=0}^{x_{ij}+1} k \varphi_{\tilde{w}_{ij}}(k) + (x_{ij} + 1) \sum_{k=x_{ij}+2}^{\infty} \varphi_{\tilde{w}_{ij}}(k) \right)$$

$$- \sum_{k=0}^{x_{ij}} k \varphi_{\tilde{w}_{ij}}(k) - x_{ij} \sum_{k=x_{ij}+1}^{\infty} \varphi_{\tilde{w}_{ij}}(k) \right)$$

$$= P_{ij} \left( \varphi_{\tilde{w}_{ij}}(x_{ij} + 1) + \sum_{k=x_{ij}+2}^{\infty} \varphi_{\tilde{w}_{ij}}(k) \right)$$

$$= P_{ij} \Pr(\tilde{w}_{ij} \geq x_{ij} + 1)$$

concluding this proof. ■

**Algorithm 43** For the inline network case with 3 nodes, there exist 3 different demands and 2 capacities.

1. Set $x_{ij} = 0$
2. Check capacity constraints. If infeasible undo the last allocation and STOP.

\[ x_{12} + x_{13} \leq C_{12} \]
\[ x_{23} + x_{13} \leq C_{23} \]

3. Compute the marginal contributions \( \Delta_{ij}(x_{ij}) \) for each \((i, j) \in L_D\)

4. If \( \Delta_{13} > \Delta_{12} + \Delta_{23} \) then \( x_{13} = x_{13} + 1 \) and go to step 2

else if \( \Delta_{12} > \Delta_{23} \) then \( x_{12} = x_{12} + 1 \) and go to step 2

else \( x_{23} = x_{23} + 1 \) and go to step 2

This algorithm will allocate optimally any given capacity for any given set of prices.

4.2.2 Pricing

In these models we consider a single period selling horizon, where at the beginning of that period the seller knows its available capacity on each link of the network. The seller has to set a price for each product (bandwidth capacity between two nodes), which will impact the current customer arrivals during the selling period. At the end of the period the seller delivers bandwidth to the committed customers and collects the revenues. Thus, there are trade-offs between prices and demands; low prices will generate high demands (probably exceeding capacity), but the revenues generated might be not high enough. On the other hand, high prices will generate low demands, but revenues might be higher.

In order to get this insight we analyze cases of networks with three nodes, where there are three different demands, i.e., \( L_D = \{(1, 2); (2, 3); (3, 1)\}\). The idea is to compare both the inline and the one-loop cases in order to determine when it is more convenient to have one or the other topology.

**Inline network**

In the inline network case \( L_I = \{(1, 2); (2, 3)\}\), i.e., there is bandwidth capacity in only 2 links (see Figure 4-4a). The mathematical formulation would be

\[
\max \sum_{(i,j) \in L_D} P_{ij} E_{\tilde{d}_{ij}} \min \{x_{ij}, \tilde{d}_{ij}\} \\
\text{subject to} \quad x_{12} + x_{31} \leq C_{12} \tag{4.8a} \\
\quad x_{23} + x_{31} \leq C_{23} \tag{4.8b} \\
\quad x_{ij}, P_{ij} \geq 0 \quad \forall (i, j) \in L_D \tag{4.8c}
\]
where \( E_{\tilde{d}_{ij}}[\cdot] \) is the expected value with respect to the random variable \( \tilde{d}_{ij} \) which represents the demand for product \((i, j)\).

**Modeling the demand as a discrete random variable** Demand for each product follows a Poisson process with rate \( \lambda_{ij} = \lambda^e_{ij}(1 - F_{ij}(P_{ij})) \) (modeled as in Bitran and Mondschein [BM97]), where \( \lambda^e_{ij} \) is the external customer arrival rate when everybody is willing to buy, and \( F_{ij}(\cdot) \) is the cumulative probability distribution of the reservation prices. The \( pmf \) of the random variable for the demand \( \tilde{d}_{ij} \) is given by

\[
Pr(\tilde{d}_{ij} = k) = \pi_{ij}(k) = \frac{(\lambda_{ij})^k e^{-\lambda_{ij}}}{k!} \quad \forall (i, j) \in L_D, \quad k = 0, 1, 2, ...
\]

**Proposition 44** The objective function in (4.7) is increasing in \( x = (x_{12}, x_{23}, x_{31}) \).

**Proof.** If that is true, then each of the three components of (4.7) is increasing in each of the three components of \( x \) respectively, i.e., \( P_{ij} \left( \sum_{k=0}^{\infty} \min\{x_{ij}, k\} \pi_{ij}(k) \right) \) should increase with \( x_{ij} \). This proof is actually quite straightforward. Consider the following inequality:

\[
P_{ij} \left( \sum_{k=0}^{x_{ij}} k \pi_{ij}(k) + x_{ij} \sum_{k=x_{ij}+1}^{\infty} \pi_{ij}(k) \right) < P_{ij} \left( \sum_{k=0}^{x_{ij}+1} k \pi_{ij}(k) + (x_{ij} + 1) \sum_{k=x_{ij}+2}^{\infty} \pi_{ij}(k) \right)
\]

subtracting \( \sum_{k=0}^{x_{ij}} k \pi_{ij}(k) \) in both sides

\[
x_{ij} \sum_{k=x_{ij}+1}^{\infty} \pi_{ij}(k) < (x_{ij} + 1) \pi_{ij}(x_{ij} + 1) + (x_{ij} + 1) \sum_{k=x_{ij}+2}^{\infty} \pi_{ij}(k)
\]

\[
x_{ij} \sum_{k=x_{ij}+1}^{\infty} \pi_{ij}(k) < (x_{ij} + 1) \sum_{k=x_{ij}+1}^{\infty} \pi_{ij}(k)
\]

\[
x_{ij} < x_{ij} + 1
\]

which proves this proposition. ■

Now the challenge is to analyze this function in terms of the prices. For a given feasible vector of capacity allocation \( \hat{x} = (\hat{x}_{12}, \hat{x}_{23}, \hat{x}_{31}) \), we can compute the first and second order derivatives of the objective function in (4.7) with respect to the prices.
Rewriting the function in (4.7) we have

\[
J(P | \hat{x}) = \sum_{(i,j) \in L_D} J_{ij}(P_{ij} | \hat{x}_{ij}) = \sum_{(i,j) \in L_D} P_{ij} \left( \sum_{k=0}^{\infty} \min\{\hat{x}_{ij}, k\} \pi_{ij}(k) \right)
\]

\[
= \sum_{(i,j) \in L_D} P_{ij} \left( \sum_{k=0}^{\infty} k \pi_{ij}(k) + \hat{x}_{ij} \sum_{k=\hat{x}_{ij}+1}^{\infty} \pi_{ij}(k) \right)
\]

so

\[
\frac{\partial J(P | \hat{x})}{\partial P_{ij}} = \frac{\partial J_{ij}(P_{ij} | \hat{x}_{ij})}{\partial P_{ij}} = J'_{ij}(\hat{x}_{ij}) \quad \forall (i, j) \in L_D.
\]

Then,

\[
J'_{ij}(\hat{x}_{ij}) = \sum_{k=0}^{\hat{x}_{ij}} k \pi_{ij}(k) + \hat{x}_{ij} \sum_{k=\hat{x}_{ij}+1}^{\infty} \pi_{ij}(k) + P_{ij} \sum_{k=0}^{\hat{x}_{ij}} k \pi'_{ij}(k) + P_{ij} \hat{x}_{ij} \sum_{k=\hat{x}_{ij}+1}^{\infty} \pi'_{ij}(k) \quad (4.9)
\]

where

\[
\pi'_{ij}(k) = \frac{k(\lambda_{ij})^{k-1} e^{-\lambda_{ij}} \lambda'_{ij}}{k!} - \frac{(\lambda_{ij})^{k} e^{-\lambda_{ij}} \lambda'_{ij}}{k!}
\]

\[
= \lambda'_{ij} \left( \frac{(\lambda_{ij})^{k-1} e^{-\lambda_{ij}}}{(k-1)!} - \frac{(\lambda_{ij})^{k} e^{-\lambda_{ij}}}{k!} \right)
\]

\[
= \lambda'_{ij} (\pi_{ij}(k-1) - \pi_{ij}(k)) \quad (4.10)
\]

and

\[
\lambda'_{ij} = \frac{\partial}{\partial P_{ij}} (\lambda_{ij}^e (1 - F_{ij}(P_{ij}))) = -\lambda_{ij}^e f_{ij}(P_{ij}).
\]

After replacing (4.10) into (4.9) and doing some algebra we get

\[
J'_{ij}(\hat{x}_{ij}) = \sum_{k=0}^{\hat{x}_{ij}} k \pi_{ij}(k) + \hat{x}_{ij} \sum_{k=\hat{x}_{ij}+1}^{\infty} \pi_{ij}(k) + P_{ij} \lambda'_{ij} \sum_{k=0}^{\hat{x}_{ij}-1} \pi_{ij}(k). \quad (4.11)
\]
Now, taking derivatives again with respect to $P_{ij}$ and doing some substitutions we get

$$J''_{ij}(\hat{x}_{ij}) = \sum_{k=0}^{\hat{x}_{ij}} k \pi'_{ij}(k) + \hat{x}_{ij} \sum_{k=\hat{x}_{ij}+1}^{\infty} \pi'_{ij}(k) + \lambda_{ij} \sum_{k=0}^{\hat{x}_{ij}-1} \pi_{ij}(k)$$

$$+ P_{ij} \lambda'_{ij} \sum_{k=0}^{\hat{x}_{ij}-1} \pi_{ij}(k) + P_{ij} \lambda''_{ij} \sum_{k=0}^{\hat{x}_{ij}-1} \pi_{ij}(k)$$

$$= \lambda'_{ij} \sum_{k=0}^{\hat{x}_{ij}-1} \pi_{ij}(k) - \lambda'_{ij} \hat{x}_{ij} \pi_{ij}(\hat{x}_{ij}) + \lambda'_{ij} \pi_{ij}(\hat{x}_{ij})$$

$$+ (\lambda'_{ij} + P_{ij} \lambda''_{ij}) \sum_{k=0}^{\hat{x}_{ij}-1} \pi_{ij}(k) + P_{ij} \lambda'_{ij} \lambda'_{ij} (-\pi_{ij}(\hat{x}_{ij} - 1))$$

$$= (2\lambda'_{ij} + P_{ij} \lambda''_{ij}) \sum_{k=0}^{\hat{x}_{ij}-1} \pi_{ij}(k) - P_{ij} (\lambda'_{ij})^2 \pi_{ij}(\hat{x}_{ij} - 1) \quad (4.12)$$

It can be observed in (4.12) that the function $J_{ij}(P_{ij}|\hat{x}_{ij})$ is not concave for every $P_{ij}$, because we cannot determine the sign$^2$ of $\lambda''_{ij}$. If $\lambda''_{ij}$ were always non-positive (as it is in the uniform distribution case), we could say that the function is concave and therefore, there exists a unique maximum.

**Definition 45** A unimodal function has one mode (usually a maximum, but it could be a minimum, depending on context). If $f$ is defined on the interval $[a, b]$, and $x^*$ is its mode, then, $f$ strictly increases from $a$ to $x^*$ and strictly decreases from $x^*$ to $b$ (reverse the monotonicity on each side of $x^*$ if the mode is a minimum). (For line search methods, like Fibonacci, the mode could occur in an interval, $[a^*, b^*]$, where $f$ strictly increases from $a$ to $a^*$, is constant (at its global max value) on $[a^*, b^*]$, then strictly decreases on $[b^*, b]$.)

In order to get some intuition about the shape of the functions $J_{ij}(P_{ij}|\hat{x}_{ij})$ we plot them using different probability distribution functions for the reservation prices (see Table 4.1).

It can be observed that for a given capacity $\hat{x}$, the function $J_{ij}(P_{ij}, \hat{x})$ is always unimodal, that is, there is a unique price that maximizes the revenues. Figure 4-5 depicts the function $J_{ij}(P_{ij}, x_{ij})$ for the case when the reservation price distribution is exponential with parameter 0.3. Two views are presented in order to see more clearly how the function increases with $x_{ij}$ and that it is unimodal in $P_{ij}$.

**Proposition 46** The function $J_{ij}(P_{ij}, x_{ij})$ is unimodal for a given $\hat{x}_{ij}$.

**Proposition 47** The optimal price $P_{ij}$ decreases with $x_{ij}$, as can be seen in Figure 4-5.

$^2$The second derivative of the effective customer arrival rate with respect to the price is given by $\lambda''_{ij} = -\lambda'_{ij} f'_{ij}(P_{ij})$
<table>
<thead>
<tr>
<th>Distribution</th>
<th>c.d.f. $F(P)$</th>
<th>$J_{ij}(P, 5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform ($a = 0, b = 100$)</td>
<td>$P / (b - a)$</td>
<td><img src="image" alt="Uniform CDF" /></td>
</tr>
<tr>
<td>Exponential ($b = .3$)</td>
<td>$1 - e^{-bP}$</td>
<td><img src="image" alt="Exponential CDF" /></td>
</tr>
<tr>
<td>Gamma ($a = 1, b = 1$)</td>
<td>$1 - (1 + P)e^{-P}$</td>
<td><img src="image" alt="Gamma CDF" /></td>
</tr>
<tr>
<td>Weibull ($a = 1, b = 2$)</td>
<td>$1 - e^{-P^2}$</td>
<td><img src="image" alt="Weibull CDF" /></td>
</tr>
</tbody>
</table>

Table 4.1: Function $J(\cdot)$ under different PDF for the reservation prices
Proof. Using the first order condition in (4.11) we have that

\[ P_{ij} = \frac{E_{\tilde{d}_{ij}}[\min\{x_{ij}, \tilde{d}_{ij}\}]}{\lambda_{ij} f_{ij}(P_{ij}) \Pr(\tilde{d}_{ij} < x_{ij})}. \]

We can see that the price cannot be obtained in closed form since it corresponds to the fixed point solution of the above expression. By performing several computational experiments for different pdf for the reservation prices, we proved that \( P_{ij} \) decreases when we increase \( x_{ij} \).

Modeling the demand as a continuous random variable Let \( h_{ij}(\cdot) \) be the probability density function of the demand for product \((i, j)\), and \( h'_{ij}(\cdot) = \frac{\partial}{\partial P_{ij}} h_{ij}(\cdot) \) the first partial derivative of the demand pdf with respect to the price.

In that case, the objective function of our mathematical formulation would be

\[
\max_{(i,j) \in L_D} \int_0^{\infty} \min\{x_{ij}, \xi\} h_{ij}(\xi) d\xi.
\]
The Lagrangian function of this modified formulation would be

\[ L = \sum_{(i,j) \in L_D} P_{ij} \left( \int_{0}^{x_{ij}} \xi h_{ij}(\xi) d\xi + \int_{x_{ij}}^{\infty} x_{ij} h_{ij}(\xi) d\xi \right) - \mu_1 (x_{12} + x_{31} - C_{12}) - \mu_2 (x_{23} + x_{31} - C_{23}) \]  \hspace{1cm} (4.13)

where \( \mu_1 \) and \( \mu_2 \) are the Lagrange multipliers for the constraints.

In what follows, we write down the first and second order derivatives of the Lagrangian function in (4.13). Hence, for every \((i, j)\) in \(L_D\), we have:

\[ \frac{\partial L}{\partial x_{ij}} = P_{ij} \left( \int_{0}^{x_{ij}} \xi h_{ij}(\xi) d\xi - x_{ij} h_{ij}(x_{ij}) \right) - \mu_i \hspace{1cm} \text{where } \mu_3 = \mu_1 + \mu_2 \]  \hspace{1cm} (4.15a)

\[ \frac{\partial L}{\partial P_{ij}} = \int_{0}^{x_{ij}} \xi h_{ij}(\xi) d\xi + \int_{x_{ij}}^{\infty} x_{ij} h_{ij}(\xi) d\xi \]  \hspace{1cm} (4.15b)

\[ + P_{ij} \left( \int_{0}^{x_{ij}} \xi h_{ij}'(\xi) d\xi + \int_{x_{ij}}^{\infty} x_{ij} h_{ij}'(\xi) d\xi \right) \]  \hspace{1cm} (4.15c)

\[ \frac{\partial^2 L}{\partial x_{ij}^2} = P_{ij} \left( -2h_{ij}(x_{ij}) - x_{ij} h_{ij}'(x_{ij}) \right) \]  \hspace{1cm} (4.16a)

\[ \frac{\partial^2 L}{\partial P_{ij}^2} = 2 \int_{0}^{x_{ij}} \xi h_{ij}'(\xi) d\xi + 2 \int_{x_{ij}}^{\infty} x_{ij} h_{ij}'(\xi) d\xi \]  \hspace{1cm} (4.16b)

\[ + P_{ij} \left( \int_{0}^{x_{ij}} \xi h_{ij}''(\xi) d\xi + \int_{x_{ij}}^{\infty} x_{ij} h_{ij}''(\xi) d\xi \right) \]  \hspace{1cm} (4.16c)

(4.16d)

(4.16e)
\[ \frac{\partial^2 L}{\partial P_{ij} \partial x_{ij}} = \int_{x_{ij}}^{\infty} h_{ij}(\xi) d\xi - x_{ij} h_{ij}(x_{ij}) \]

\[ + P_{ij} \left( \int_{x_{ij}}^{\infty} h'_{ij}(\xi) d\xi - x_{ij} h'_{ij}(x_{ij}) \right) \]

\[ \frac{\partial^2 L}{\partial P_{ij} \partial x_{i'j'j}} = 0 \quad \forall (i, j) \in LD, \forall (i', j') \in LD, \text{ such that } (i, j) \neq (i', j') \]

The Hessian matrix of \( L \) is given by

\[
H = \begin{bmatrix}
    a_1 & 0 & 0 & b_1 & 0 & 0 \\
    0 & a_2 & 0 & 0 & b_2 & 0 \\
    0 & 0 & a_3 & 0 & 0 & b_3 \\
    b_1 & 0 & 0 & c_1 & 0 & 0 \\
    0 & b_2 & 0 & 0 & c_2 & 0 \\
    0 & 0 & b_3 & 0 & 0 & c_3 \\
\end{bmatrix}
\]

where \( a_i = \frac{\partial^2 L}{\partial x_{ij}^2} \), \( b_i = \frac{\partial^2 L}{\partial P_{ij} \partial x_{ij}} \), and \( c_i = \frac{\partial^2 L}{\partial P_{ij}^2} \) for all \( i = 1, 2, 3 \).

The determinant (\( \Delta_6 \)) and subdeterminants (\( \Delta_5, \ldots, \Delta_1 \)) of this Hessian matrix are

\[
\Delta_6 = (c_1 a_1 - b_1^2) (c_2 a_2 - b_2^2) (c_3 a_3 - b_3^2) \\
\Delta_5 = c_1 (c_2 a_2 - b_2^2) (c_3 a_3 - b_3^2) \\
\Delta_4 = c_1 c_2 (c_3 a_3 - b_3^2) \\
\Delta_3 = c_1 c_2 c_3 \\
\Delta_2 = c_2 c_3 \\
\Delta_1 = c_3
\]

which will help us determine concavity of objective function.

In order to have concavity, we need \( \Delta_6, \Delta_4, \Delta_2 \geq 0 \) and \( \Delta_5, \Delta_3, \Delta_1 \leq 0 \), necessary conditions to have a negative semi-definite Hessian matrix \( H \). The last condition, \( \Delta_1 \leq 0 \), implies that \( c_3 \leq 0 \), and this, together with \( \Delta_2 \geq 0 \), implies that also \( c_2 \leq 0 \), and so on and so forth. Continuing in that way up to \( \Delta_6 \), we conclude that the following necessary conditions are needed to be in the presence of concavity:

\[
c_i \leq 0 \quad \forall i = 1, 2, 3 \\
\frac{b_i^2}{c_i} \leq a_i \quad \forall i = 1, 2, 3
\]

The KKT (Karush-Kuhn-Tucker) necessary optimality conditions for the prices and
capacity allocations are given by

\[ \frac{\partial L}{\partial x_{ij}} \leq 0 \quad \forall (i, j) \in L_D \]

\[ x_{ij}^* \left( \frac{\partial L}{\partial x_{ij}} \right) = 0 \quad \forall (i, j) \in L_D \]

\[ \frac{\partial L}{\partial P_{ij}} \leq 0 \quad \forall (i, j) \in L_D \]

\[ P_{ij}^* \left( \frac{\partial L}{\partial P_{ij}} \right) = 0 \quad \forall (i, j) \in L_D \]

\[ x_{12}^* + x_{31}^* - C_{12} \leq 0 \]

\[ \mu_1(x_{12}^* + x_{31}^* - C_{12}) = 0 \]

\[ x_{23}^* + x_{31}^* - C_{23} \leq 0 \]

\[ \mu_2(x_{23}^* + x_{31}^* - C_{23}) = 0 \]

\[ x_{ij}^*, P_{ij}^* \geq 0 \quad \forall (i, j) \in L_D \]

\[ \mu_1, \mu_2 \geq 0 \]

One-loop network

In that case \( L_I = \{(1, 2); (2, 3); (3, 1)\} \), i.e., there is a bandwidth capacity in all 3 links (see Figure 4-4b). The mathematical formulation would be

\[
\max_{(i, j) \in L_D} \sum P_{ij} \tilde{E}_{d_{ij}} [\min\{x_{ij} + y_{ij}, \tilde{d}_{ij}\}]
\]

subject to

\[ x_{12} + y_{23} + y_{31} \leq C_{12} \]

\[ y_{12} + x_{23} + y_{31} \leq C_{23} \]

\[ y_{12} + y_{23} + x_{31} \leq C_{31} \]

\[ x_{ij}, y_{ij}, P_{ij} \geq 0 \quad \forall (i, j) \in L_D \]

where \( \tilde{E}_{d_{ij}} [\cdot] \) is the expected value with respect to the random variable \( \tilde{d}_{ij} \) which represents the demand for product \((i, j)\). Note that if \( C_{31} = 0 \) we end up in the previous inline case.

4.3 Extensions

The number of extensions to this problem is quite large. This is a research area that is still at its infancy. In order to prioritize the different types of models that we can develop for this yield management bandwidth capacity problem, a taxonomy of models is presented (see Figure 4-6). There are other dimensions such as variable-duration contracts and
endogenous/exogenous prices that we leave aside for the moment. In this chapter, we just concentrated on the monopolistic case where prices are endogenous and the duration of the contract is fixed (e.g. one month).

<table>
<thead>
<tr>
<th>Network Topology</th>
<th>Spot Pricing Model</th>
<th>Multi-period Pricing Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certain Capacity</td>
<td>Uncertain Capacity</td>
<td>Certain Capacity</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 4-6: Stochastic formulation taxonomy

A natural extension would be to analyze the 4-node case, where we add an extra topology “the star network” to the comparison (see Figure 4-7), as it was mentioned in section 4.1.4. With this extra node, all cases now are much more difficult to be solved.

Figure 4-7: Network topologies for 4 nodes

An extension of Dhadwal [Dha01] is another promising research that could be done. He studied the case of a single link under uncertain capacity. We could extend his problem to the case of an inline network starting by analyzing the 2-link or 3-node inline case.

There are some other issues that we intend to explore in future research that will certainly enrich these models. They are:

- Loyal or frequent customers: Some capacity has to be reserved for future periods in order to be able to fulfill a frequent customer’s request for bandwidth.
• Over-commitments: There is uncertainty in customer usage, since some customers will not use 100% of their allocated capacity. Based on that, we could sell more than the total capacity existing in the network.

• Network reliability: When there is more than one possible path from city $i$ to city $j$, the network connection will be more reliable, because a link failure will not be an obstacle to deliver bandwidth between any two nodes.

4.4 Summary

Suppose that a network owner is selling bandwidth capacity to be delivered $T + 1$ periods from now on. The decision problem is to decide how much to sell from the present period to period $T$. We consider the case where the network owner faces an external demand for each market or product (capacity between two cities) that can be fulfilled by assigning capacity of the corresponding arc (direct link) or by assigning capacity of an alternative path (indirect link), incurring a higher cost. This problem differs from classical routing problems faced in the transportation industry in that customers do not distinguish between routings. They only care about communicating between two specific cities and the Quality of Service (QoS). The goal is to maximize the network owner’s revenues while deciding on each period whether to bundle links (for a demand from node $i$ to node $j$ when there is no direct link $(i, j)$) or to leave them separated.

We built models for comparing two 3-node network topologies, an inline network (with two arcs) and a single-loop network (with three arcs). We first studied two deterministic models for bandwidth capacity allocation when demand and prices are known. Then, we considered static models with stochastic demand and unknown prices for the two network topologies. Finally, dynamic models establish selling and allocation policies based on the available capacity and the remaining time of the selling horizon.

We first considered the deterministic case in which demand and prices are known. In this problem, the integer program that determines optimal capacity allocation in an inline network is unimodular. This means that its linear programming relaxation will yield the optimal integer solution. Unfortunately, the unimodularity property does not extend to the case of networks with loops, in which case the IP has to be solved.

In the case of stochastic demand, the profit function is unimodal (quasiconcave) for a given capacity allocation. Therefore, there is a unique optimal price which can be found. This price is characterized by a fixed point expression, and has the property that the higher the capacity allocation, the smaller the optimal price. Since demand is stochastic, the higher the allocation, the closer we are to the expected demand. If capacity is higher than demand, the price must be decreased to attract more clients. This result has been verified numerically with many reservation price distributions.
Chapter 5

Concluding Remarks

This dissertation analyzed three separate but related problems concerning bundles in high-tech industries. This chapter brings together the insights from these analyses.

The first problem analyzed in this dissertation concerns pricing bundles with high setup costs. The bundles considered contain an initial component that is delivered as soon as new customers buy the bundle and a service component that is delivered on a continuous basis. The bundle is paid through monthly fees, which include a cost component (to cover the expenses incurred by the company in providing the bundle) and a profit component. The analysis in Chapter 2 showed that the higher the price of the bundle, the higher the cost component of the monthly fee. Furthermore, the analysis yielded a closed-form expression for total profit per customer per bundle. This closed-form expression proved to be an essential building block in the optimization problem of finding the optimal prices because it is equivalent to an otherwise untractable summation. The total profit per customer per bundle can be expressed as a function of the moment generating function of the customer lifetime distribution.

The expected profit per customer is an increasing function of price. This result was proven for the case when the customer lifetime distribution is geometric. This is a significant result because it takes into account the competing dynamics that increasing the monthly fee simultaneously increases monthly revenue per customer and lowers the customer lifetime. Yet, an increase in price always results in higher profit regardless of the original price level. This result can then be incorporated into a more general model (including external demand), where a very high price will lower the probability that a given bundle will be bought.

The analysis of the one-step decision process shows that optimal prices fulfill an insightful first-order condition. More precisely, profit per customer per bundle minus its derivative with respect to its price is constant for all bundles. This condition allows for the development of an algorithm that iteratively determines the optimal prices. The algorithm is very efficient because the first order condition establishes that the price of one bundle automatically determines the price of all others. Therefore, the search space over
which the optimization occurs can be reduced to feasible prices for a single bundle.

The first-order conditions for the two-step decision process are analogous to those obtained for the one-step process. In this case, the product of the probability of purchase with the total expected profit per customer per bundle (denoted $R_j$), minus its derivative with respect to price (denoted $R'_j$), is constant for all bundles. The functions $(R_j - R'_j)$ are concave with respect to price. Therefore, the search process is not as easy as the previous one since each profit level can correspond to two different prices. Nevertheless, an algorithm can still be implemented. Since the $(R_j - R'_j)$ functions are concave, we can easily find the maximum level achieved for each bundle. The lowest such level will be the highest possible level that can be achieved by the optimal total expected profit. Therefore, this profit level and its unique corresponding price can be used to find an upper bound for the overall profit. Using this price and the first order conditions we can then find the optimal price and profit level for other bundles, thus deriving a lower bound for profit. In this way, the search space for the optimal profit level is bounded from above and below and the search space for the optimal solution is well-defined. Iterations occur inside these bounds, which were experimentally found to be very small.

From the dynamic formulation of the customer base evolution, we found that for any given set of parameters, the optimal prices stabilize the number of customers in the system over time. For example, it is never optimal to have too many customers because this means that the current price is too low and more profit could be obtained by increasing prices. In other words, if the rate of defection is smaller than arrival rate, the number of clients always increases over time. By increasing prices, the rate of defection increases and the arrival rate decreases. Our intuition backed by numerical experiments was that steady-state number of customers in the system always converges to a constant. The optimal pricing policy depends upon the discount factor. We found that for discount factors below a known threshold, the optimal pricing policy is such that the price decreases with the number of customers. Conversely, for discount factors above this threshold the price increases with the number of customers. The threshold is well characterized analytically.

Comparing our dynamic programming stochastic formulation with the certainty equivalent counterpart we found an equivalence using Jensen’s inequality because the profit function is linear in the number of customers who continue in the system and in the number of new customers. Numerical experiments suggest that a single fixed-price policy is nearly optimal when the system is loaded, and this implies that offering multiple prices can at best capture second-order increases in profit.

We now turn to the results obtained in Chapter 3, where we address the problem of how to determine the composition and price of a bundle so as to maximize total expected profits. The optimal decision can be found by solving a nonlinear mixed integer programming model. These models have been proven to be NP hard. Demand is a function of the bundle price and a vector of attributes, and it is modeled using a multinomial logit model (MNL). The profit per sold bundle is the difference between the price and the bundle
cost. Each bundle is characterized by its cost and its attractiveness level (computed as a weighted average of the individual component’s attractiveness).

If demand is modeled with MNL, then for a given bundle the objective function is unimodal (quasi-concave). In this case, there is a unique optimal price which can be determined in closed-form. The optimal marginal profit per customer converges to \( \frac{1}{\beta} \) where \( \beta \) is the price response parameter (a negative constant) as the cost goes to infinity. This convergence is very fast due to the MNL.

This model can be extended to the case when there are multiple bundles, in which the result is that the profit (price minus cost) is constant across all bundles. This result is analogous to the optimality conditions found in chapter 2 for the one-step and the two-step decision process models.

The fact that we found a closed-form solution for the optimal price is very important. This result can be substituted into the original formulation. A pre-optimized model can then be solved in order to find the optimal bundle configuration.

We prove that for each attractiveness level, the bundle with the smallest cost dominates all other bundles. This allows us to find a Pareto frontier which is a set composed by the dominating bundle for each attractiveness level. This frontier can be determined with an efficient algorithm that we have specified. It turns out that the profit function has a constant gradient’s direction, which means that the problem of finding the optimal composition reduces to a solvable integer program with a linear objective function being optimized over this Pareto frontier set.

We now turn to the results from Chapter 4 that analyzed the problem of bundles in networks. We built models for comparing two 3-node network topologies, an inline network (with two arcs) and a single-loop network (with three arcs). We first studied two deterministic models for bandwidth capacity allocation when demand and prices are known. Then, we considered static models with stochastic demand and unknown prices for the two network topologies. Finally, dynamic models establish selling and allocation policies based on the available capacity and the remaining time of the selling horizon.

We first considered the deterministic case in which demand and prices are known. In this problem, the integer program that determines optimal capacity allocation in an inline network is unimodular. This means that its linear programming relaxation will yield the optimal integer solution. Unfortunately, the unimodularity property does not extend to the case of networks with loops, in which case the IP has to be solved.

In the case of stochastic demand, the profit function is unimodal (quasiconcave) for a given capacity allocation. Therefore, there is a unique optimal price which can be found. This price is characterized by a fixed point expression and has the property that the higher the capacity allocation, the smaller the optimal price. Since demand is stochastic, the higher the allocation, the closer we are to the expected demand. If capacity is higher than demand the price must be decreased to attract more clients. This result has been verified numerically with many reservation price distributions.
Finally, bundling is not always “the strategy” to use, but when it is, it can be powerful in helping companies increase their profits. This dissertation provides a clear contribution to the Bundling Operations Research literature, and increases companies’ awareness about the importance of bundling.
Appendix A

Bundle Pricing Overview

The design of customized services is a dynamic process. The service interface should be a personalized interactive communication/selling channel where the information stream and the service/product come in the same package. During each interaction the customer provides information (e.g., by answering a few questions) that is used to dynamically design and deliver the most suitable service offerings.

The objective of delivering personalized services is to induce customer loyalty. Customers receiving high quality customized services face barriers when they consider switching to competitors. These barriers, the size of which depends on the amount of information available, the quality of the information, and the way in which that information is used, work as deterrents to changes.

![Service Platform Diagram]

Figure A-1: Service Platform

Highly customized services can be made possible through a service platform that provides the basic offerings common to most small and medium businesses. This platform is complemented by specific offerings for a given business (customer) that will shape the final service to be delivered. These specific offerings will depend upon business characteristics, such as industry, management style, and customer needs (see Figure A-1).

The final purchase of e-services and products will be a bundle consisting of many
offerings that will meet the customers’ needs and give them the most value. Consequently, the issue of bundling becomes extremely important in deciding how to combine and price the offerings. HP wants to maximize its revenue and help their customers choose bundles that best fulfill their needs. These objectives can be reached by understanding the notion of bundling, bundling strategies, and decision models for bundling.

A.1 Motivation

Bundling is a widely used price and design strategy. For example, when buying a new car, customers can purchase such options as power windows, power seats, a sunroof, or they can purchase a “luxury package” in which these options are sold as a bundle. Bundling makes sense when customers have heterogeneous demands, and when the company cannot price discriminate. In other words bundling involves offering special prices to buyers purchasing the main items plus one or more auxiliary items. This is widely used in industrial marketing when complementary products and services exist. In order to be effective, bundling requires that true complementary relationships exist.

How can a company decide whether to bundle its products and determine the profit-maximizing prices? Most companies do not know their customers’ reservation prices (the maximum amount of money that customers are willing to pay). However, by conducting market surveys, they may be able to estimate the distribution of reservation prices, and then use this information to design a pricing strategy. The following example illustrates this concept and the mechanism behind it.

A company needs to decide whether to bundle two services or not. Table A.1 shows the reservation prices of each customer for each service.

<table>
<thead>
<tr>
<th>Customer</th>
<th>Service A</th>
<th>Service B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$12,000</td>
<td>$3,000</td>
</tr>
<tr>
<td>2</td>
<td>$10,000</td>
<td>$4,000</td>
</tr>
</tbody>
</table>

Table A.1: Customers’ reservation prices

If the company does not bundle, it can earn only $26,000 since $P_A = $10,000 and $P_B = $3,000 are the optimal prices that maximize the revenues (both customers will purchase both services). On the other hand, if the two services are bundled, they can earn $28,000 since $P_{AB} = $14,000 makes the bundle attractive for both customers. In economic terms, the consumer surplus from the highly valued service A is transferred to the less valued service B ($1,000 is transferred from service A to service B in customer 1). In that case, with heterogeneous demand, the company is better off by bundling.
A.2 Notion of Price Bundling

Guiltinan and Paul [Gui87] define price bundling as the practice of marketing several products or services together for a special price. Similarly, Fuerderer [FHW99] states that price bundling is the practice of collecting goods or services in a package and selling them at a discounted package price. As the previous example shows, bundling allows the transfer of excess consumer surplus from one product to another by adding reservation prices (willingness to pay). It has become a widespread sales practice in many production or service-oriented industries.

According to Simon [SR95], bundling plays an increasingly important role in many industries, and some companies even build their business strategies on bundling. A renowned case is Microsoft. By smartly combining its application software into the “Office” bundle, Microsoft increased the market share of Access and PowerPoint when it bundled these two less attractive components with the more attractive components Word and Excel. Simon also mentions that bundling is particularly popular in the service sector. Other examples are vacation packages (airline ticket, hotel accommodation plus rent-a-car), insurance packages, restaurant menus (appetizer, entrée, dessert), and telecommunication packages (local, long-distance, internet access, cellular phone).

Fuerderer [FHW99] indicates that cross-industry bundling can also be observed. As one of many examples, German Lufthansa Card in alliance with credit institutes Eurocard and VISA, German telecommunication giant Telekom, and many hotel and car rental companies. This multi-purpose card includes not only air transportation and credit card services, but also a telephone chip, accident insurance, and special hotel and car rental conditions.

Bundling is an implicit way to price discriminate since it transfers consumer surplus to the producer (see example in previous section). The most well known types of price discrimination are the following:

- **First-Degree Price Discrimination:** It is the practice of charging each customer his or her reservation price; i.e., the maximum price that a customer is willing to pay for each unit bought.

- **Second-Degree Price Discrimination:** It works by charging customers different prices for different quantities or “blocks” of the same good or service.

- **Third-Degree Price Discrimination:** It involves the division of consumers into two groups, with separate demand curves for each group. The optimal prices and quantities are such that the marginal revenue from each group is the same and equal to marginal cost.

- **Two-Part tariff:** The two-part tariff is related to price discrimination and provides another means of extracting consumer surplus. It requires consumers to pay a fee
(T) up front for the right to buy a product. Consumers then pay an additional fee (P) for each unit of the product they wish to consume (Figure A-2a shows the fees for homogeneous consumers, while A-2b shows the fees for heterogeneous consumers). The classic example is an amusement park. You pay an admission fee to enter, and you also pay for each ride you go on.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>MC/SQ</th>
<th>P</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T: entry fee</th>
<th>P: usage fee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantity</th>
<th>MC/SQ</th>
<th>P</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) One consumer  
(b) Two consumers

Figure A-2: Fixed and variable fees for the two-part tariff

A.3 Bundling Strategies

Making decisions about product/service bundling includes the evaluation of each possible form to implement it. The most traditional strategies distinguished throughout the literature are presented below.

a) **No Bundling**: Products or services are offered and priced individually. (Also called separate pricing or pure component pricing.)

b) **Pure Bundling**: Products or services are only offered in bundles, they cannot be bought individually. This strategy usually is the one that yields the lowest profits, as opposed to mixed bundling which yields the highest profits.

c) **Mixed Bundling**: This type of bundling is a mix of the previous two. Technically, most companies employ mixed price bundling: buyers are given the choice of either buying products in a package or buying the products individually. In two-product bundles, buyers who place a low value on one of the two products will avoid the bundle. However, the economic incentive of a lower price on one item will lead to additional sales of both products to some buyers who otherwise would buy only one. When complementary relationships are very strong, the effects of the special price are even greater.

1. There are two approaches to accomplish mixed price bundling:
• **Mixed leader**: the price of a lead product is discounted on condition that a second product be purchased, e.g., a ski resort may lower ski rental rates to customers who buy skiing lessons at the regular price.

• **Mixed joint**: two or more products or services are offered for a single package price, e.g., home entertainment centers (packaging electronic audio and video products into a total system), or vacation packages (airline, car rental, and hotel accommodations).

The next two programs, extracted from Guiltinan [Gui87], shows characteristics of successful mixed price bundling programs:

**Mixed leader programs**

1. Demand for the lead product should be price-elastic.

2. Complementarity is based on the leader being enhanced by the other product(s) or on convenience.

3. If the objective is to cross-sell complements to regular customers; leader is the lower margin product (so that the lost profit from the price reduction is minimized); volume for the leader exceeds that of other products.

**Mixed joint programs**

1. Demand for the total package is price-elastic.

2. Complementarity is bi-directional (each product in the bundle enhances the value of the other) or is based on convenience.

3. If the objective is to cross-sell complements to regular customers, the various products in the bundle are approximately equal in volume and in profit margins so that sales gains from regular purchasers of any product are about equal.

**d) Premium Bundling**

As in mixed bundling, sellers price discriminate by offering products both separately and as bundles. However, bundles are sold at a premium (rather than at a discount) relative to the prices charged for the individual components.

According to Cready [Cre91], this is possible when individual products alone offer little benefit. Implementing a premium bundling strategy requires that the seller prevent component purchasers from purchasing more than one of the two bundle components at component prices.

In general, sellers operating in service markets should find it relatively easy to implement premium bundling strategies because of their requisite knowledge of customers,
while those selling products to large numbers of unknown customers may find it difficult to implement such strategies.

e) Other types of Bundling

Tie-in sales: The buyer of the main product (tying good) agrees to buy one or several complementary goods (tied goods), which are necessary to use the tying good, exclusively from the same supplier. A well-known example is printers. They require ink cartridges or toners for their operation.

Add-on bundling: It is similar to tie-in sales, but here the “add-on” product will not be sold unless the lead product is purchased. A car wash (lead service) bundled with extra wax (add-on service), is a good example of add-on bundling. The extra wax cannot be bought without buying the car wash service.

Cross couponing: It is often used to introduce new products and/or increase the sales of weak products by linking them with established products in the company’s product line. For example, Diet Coke with a coupon for buying a new diet-minute maid juice.

Reservation price is the main issue when bundling. The optimal strategy and the bundle itself depend on the distribution of the customers’ reservation prices. According to Simon [FHW99], if reservation prices are high for one product and low for the other product, separate pricing tends to be optimal. If reservation prices are relatively high for both products, pure bundling is recommendable. And if we have a combination of both customer groups, i.e., those with “extreme” preferences and those with “balanced” preferences, mixed bundling is probably the best pricing strategy.

A.4 Why Bundling?

Simon and Wuebker [FHW99] and Carlton et al. [CW98] point out several reasons for the use of bundling. An exhaustive list is presented below:

Extension of monopoly power or preserve a monopoly position: A firm with monopoly power in one market can use the leverage provided by this power to foreclose sales in, and thereby monopolize, a second market. As Carlton et al. [CW98] explains, if the primary (monopolist) and the complementary product are used in fixed proportions, then the monopolist of the primary product can achieve a virtual tie by setting a high price on the primary product and a very low price (say zero) on the complementary product. This situation achieves a virtual tie since alternative producers of the complementary product cannot operate profitably given the very low price charged for this product by the monopolist.

Price discrimination: Bundling works as an implicit price discrimination tool because it allows sellers to extract more consumer surplus from buyers.

Reduction in complexity costs: Bundling reduces the wide variety of options, which leads to a reduction in complexity costs (e.g., the automobile industry)
Reduction in transaction costs: Buyers avoid the transaction costs of contracting with several firms. Hence, they save time and information costs (e.g., the telecommunication industry)

Barriers to entry: Bundling can help to lock customers, and hence deny them to competitors. Switching costs may be increased, acting as a barrier to new entrants.

Economies of scope and scale: The incentive for quantity discounts arises from scale economies, whereas bundling yields scope economies.

A.5 Economic Research

Bundling has been the focus of attention of scholars in economics for over two decades. Traditionally, economic studies present formal analyses of bundling issues.

The most important building block in the bundling literature is the work by Adams and Yellen [AY76], where they introduce a two-dimensional graphical framework for analyzing the effect of bundling as a price discrimination tool. They examine a multiproduct monopolist with two products, independent and additive consumer valuations and linear “unit demand” for these goods (i.e. consumers buy either zero or one unit, zero marginal utility of a second unit). They consider additive variable costs and no fixed costs associated with each commodity. Comparing unbundled sales to pure bundling and mixed bundling, they demonstrate that each pricing strategy can be advantageous. The two major factors determining the profitability of either strategy are the level of cost and the distribution in the reservation price space. A strongly negative correlation of reservation prices, for instance, is shown to favor bundling strategies as opposed to single pricing. For a symmetric demand distribution and a relatively high cost level, a single pricing strategy may be preferred.

In “Commodity Bundling” [DC84], Dansby and Conrad drop the assumption of additive reservation prices. They examine the case, where a bundle may either contain an unwanted component (value reducing) or provide additional value beyond the aggregated values of the individual items (value enhancing). They conclude that this diversity in consumers’ bundle preferences can be an additional incentive for companies to bundle, even if the company has no monopoly power.

Schmalensee [Sch84] conducts a more formal analysis and develops numerical criteria on which a bundling strategy turns out to be more profitable. He assumes that buyers’ reservation prices follow a bivariate Gaussian distribution. Stating the profit function and the according problem, he shows that there is no explicit representation of an optimal price. Thus, he obtains profitability conditions in either bundling case by numerical analysis. He extends the result of Adams and Yellen [AY76] for pure bundling and demonstrates that a negative correlation in consumers’ reservation prices is not a necessary condition for pure bundling to be more profitable than unbundled sales. Actually, pure bundling is shown to be always more profitable if reservation prices have a positive correlation and
the cost level is low enough. Furthermore, he proves that mixed bundling is not more profitable than pure bundling in the case of perfect positive correlation of reservation prices. As in the pure bundling case, he suggests that mixed bundling is more profitable than unbundled sales if the cost level is low enough. Yet, he claims that his analysis of an optimal mixed bundling strategy is not complete.

Cready [Cre91] addresses profitability conditions for premium bundling. In the case of overall positively correlated reservation prices, and a strong negative correlation among consumers with relatively low reservation prices, premium bundling can be more profitable than unbundled sales, pure, or mixed bundling. The major practical problem, of course, is how to prevent customers from self-bundling the bundle components. Although he mentions couponing and rebates as potential tools to employ premium bundling, realizations of this practice, as for collectible items (stamps, coins), seem to be rare.

Salinger [Sal95] extends the results of Schmalensee [Sch84] by developing a graphical framework to analyze the profitability and welfare implications of bundling two goods, primarily in the context of independent linear demand functions. He finds that bundling two goods tends to be profitable when consumer valuations are negatively correlated and high relative to marginal costs.

A.6 Consumer Behavior Research

There is extensive research in product bundling in the economics literature. However, given the macro orientation of this research and its limited applicability to marketing practice, marketing academics have recently started to aggressively pursue research in product bundling. The focus of some of the consumer behavior studies, such as those by Yadav [Yad94] and Yadav and Monroe [YM93], has been to understand the process by which consumers evaluate product bundles. In particular, the latter study examines the transaction value of a bundle with a focus on a customer’s perception of savings in a bundle price. They found that consumers’ perceptions of total savings on a bundle is a sum of (a) perceived saving on individual items if purchased separately, and (b) the additional savings when the items in a bundle are bought together as a set.

Suri and Monroe [FHW99] extend this line of research in consumer behavior by understanding the effects of a contextual factor, such as consumers’ prior purchase intentions on their evaluation of product bundles. They conducted an experiment to validate the following hypothesis: “Given that a consumer has prior intentions to buy one or none of the items in a two item bundle, the total transaction value (perceptions of saving in a transaction) for the bundle will be influenced by the bundle transaction value with little or no impact on the item’s transaction value” (p. 182). The results of this experiment showed that when subjects had prior intentions to purchase either one or none of the items in a two-item bundle, the total transaction value of the bundle was significantly influenced by the bundle transaction value only. This result is consistent with that from
Yadav and Monroe’s [YM93] study, which showed that while evaluating the overall savings on a bundle the savings on the bundle itself had a greater impact than the savings on individual items.

According to Wuebker, Mahajan, and Yadav [FHW99], bundling strategies are often accompanied by promotions on individual items. That is, firms promote their products individually as well as offer them as part of a package at a special price. The company’s motivation for price-promoting the individual items is to enhance the sale of these individual items. This research examines the extent to which bundling can stimulate demand under specific promotion conditions. Moreover, it shows that increased promotion activity on the individual items can significantly lower buyers’ evaluations of bundle offers featuring the individual items. Therefore, it appears that buyers tend to buy the bundle when the price advantage of the single items over the bundle seems to be low and when it seems to be difficult to find these individual items on sale.

How do buyers evaluate a bundle of items? The most common approach in earlier economic analyses was to start with the additivity assumption (the overall utility of a bundle equals the sum of the bundle items individual utilities). The restrictiveness of this additivity assumption was later recognized and removed. However, the question of how buyers’ evaluation processes may account for additivity has remained largely unexplored. In [Yad94], Yadav provides insights about the anchoring and adjustment heuristic in the context of bundle evaluation. In this experiment, subjects examined items in decreasing order of perceived importance, making insufficient upward or downward adjustments to form the overall bundle evaluation. When faced with an excellent anchor and moderate add-on items, subjects readily adjusted the overall bundle evaluation downward. However, the tendency to adjust upward was considerably less when the anchor was poor and the add-on items were moderate. Furthermore, the detrimental effect of moderate items on excellent anchors was more pronounced than their enhancing effect on poor anchors. This result is compatible with the well-known “prospect theory” of Kahneman and Tversky [KT79] that the impact of perceived losses is greater than the impact of perceived gains.

Finally, Herrmann, Huber, and Coulter [FHW99] examine the individual and combined effects of four bundle factors on product and service purchase intention. Their results show that price discount and complementarity of bundle components appear to be key drivers of purchase intention for both the automobile and automotive service bundles. Moreover, the interaction between the two factors suggests that there is a threshold level of discounting necessary to affect purchase intention. On average, a 20% discount results in the greatest purchase intention, but care must be given to assessing the value of a 10% discount. In some cases, the 10% discount will save money because it generates the same level of purchase intention as a 20% discount. In general, their findings emphasize the need to attend not only to individual bundle variable effects but also the interaction effects.
A.7 Decision Models Research

a) Enumerative Approaches

Using these approaches all or subsets of possible bundle configurations are enumerated, and a test price is assigned to every bundle. Then, the profit for each configuration is computed in order to evaluate them.

Advantages:

- Enumerative models are widely used in industry.
- Elaborated methods are rather unknown.
- The use of sophisticated methods is not cost efficient for non-complex situations.
- Digital data processing allows for directed simulation methods.

Disadvantages:

- A total enumeration of bundle configurations is rather infeasible.
- Partial enumeration leads to suboptimal results with unpredictable quality of the solution.

b) Optimization Approaches

First Model

Only a few years ago, Hanson and Martin [HM90] presented the first mathematical programming formulation to determine the profit maximizing bundle configurations and prices, without explicitly considering the full range of feasible problem solutions.

The model developed is from the company’s perspective, and the objective function represents the company’s profit. The behavior of the customer as maximizers of their surplus is treated in the model as constraints on the firm’s objective function. Customers are grouped in segments according to their reservation prices for the bundles.

Below are presented the main elements of this model: its design, variables and parameters, and a formal mathematical formulation.

Decision variables:

- $p_i$ = price of bundle $i$
- $q_{ki}$ = 1 if customer $k$ selects bundle $i$, 0 otherwise
- $s_k$ = consumer surplus of customers in segment $k$

Some parameters:

- $R_{ki}$ = reservation price of segment $k$ for bundle $i$
- $c_{ki}$ = cost of supplying one customer of segment $k$ with bundle $i$
The output of their model is an optimal bundle for each segment of customers and the optimal price to charge for it, in order to maximize revenues (company’s objective) while maximizing consumer surplus (customer’s objective). The assumptions considered by the model are:

- The benefit of a duplicate component is zero, and component resale is not possible.
- A single profit maximizing company determines the price of every bundle.
- Given these prices, every customer maximizes his consumer surplus (reservation price minus the product price).
- All customer segments face the same prices.
- The reservation prices of all customer segments are known for all bundles.
- There is free disposal for unwanted components. It means that the reservation price of bundle \( j \) is greater or equal to the reservation price of bundle \( i \), if bundle \( j \) contains bundle \( i \).
- Marginal costs of a bundle are subadditive.
- Customers have zero assembly transaction costs for creating bundles from separately offered subbundles.

Second Model
Three years later Venkatesh and Mahajan [VM93] published a work where two dimensions in the consumer decision-making process are considered, in order to determine optimal prices for components and for a given bundle. They modeled a problem faced by the entertainment industry (music/dance performances), which frequently makes bundling decisions at the end-consumer level. The bundle in this case is a series of performances organized over a certain time horizon. The temporal dispersion of performances adds uncertainty to the consumer’s ability to spare the time to attend any performance. With that in mind, a prospective viewer’s decision to attend a performance is related to two key independent resources, time and money, which are assumed independent (this model is appropriate for situations in which there is no significant relationship between the two dimensions). Now, given a set of \( n \) performances, the seller faces a crucial and interesting decision problem of pricing the single ticket and/or the season ticket, which maximizes profits.

Some notation used in this model:
\[ M = \text{market size} \]
\[ n = \text{number of performances} \]
\[ p_s = \text{optimal price of a single ticket} \]

\[ p_b = \text{optimal price of the season ticket} \]

\[ E_j = \text{expenditure to organize the } j\text{th performance.} \]

\[ f(z) = \text{pdf of the likely number of performances potential consumers can attend, assuming that price is not a constraint.} \]

\[ g(p_s) = \text{pdf of performancewise reservation price, assuming that the consumers are willing to spare the time to attend.} \]

For viewers who expect to spare the time to attend \( i \) performances, the likely number of single tickets sold across all \( n \) performances is given by

\[ NT_i = iM \int_i^{i+1} f(z)dz \int_{p_s}^{\infty} g(p)dp. \]

The authors provide a mathematical formulation for the profit function separately for pure components, pure bundling, and mixed bundling strategies. Under each strategy of the first two strategies, they found expressions that characterize the optimal prices, and for the third one a numerical search needs to be performed in order to identify the simultaneous optimal values of \( p_s \) and \( p_b \).

With this approach they answer four issues across the three different strategies in an integrative way: (a) optimal prices of bundles and/or their components; (b) corresponding profit levels; (c) sources of revenues as well as gains and losses in revenue if the seller changes from one bundling strategy to another; and (d) price sensitivity of profits.

A limitation in this model is that they do not capture supply side constraints, such as the capacity of the auditorium or stadium that may alter the optimal solutions. Yield management approaches need to be considered when designing the model so that selling-to-capacity factor is taken into account.

Third Model

Fuerderer et al. [FHW99] proposed another approach in modeling bundle pricing. In this study, they demonstrate how a single company can optimize the design and the price of its product line under uncertainty in reservation prices and consumer choice behavior considering both volume-dependent and variant-dependent costs. Furthermore, they assess two bundling strategies: pure bundling and mixed bundling. They developed a model based on the following assumptions:

- Reservation prices for segment \( i \) for bundle \( j \) is assumed normally distributed.

- Demand correlation (linear) of bundles \( j \) and \( j' \) for segment \( i \) do not depend on the decision variables.

- The utility of bundle \( j \) for segment \( i \) only depends on the realized consumer surplus.
\((e_{ij})\) and the supply of other substitutive bundles:

\[
\varepsilon_{ij}(p_j) = \int_{p_j}^{\infty} (p - p_j) \varphi_{ij}(p) dp
\]

where \(\varphi_{ij}(p)\) is the reservation price density function.

- Intraline competition for segment \(i\) and bundle \(j\) can be approximated by a modified a-choice pattern, i.e.,

\[
\pi_{ij}(A, p, y, \alpha_i) = \frac{\varepsilon_{ij} y_j}{\sum_{k=1}^{m} \rho_{ijk} e^{\alpha_i} y_k}.
\]

- No customer undertakes a purchase resulting in a negative consumer surplus for each bundle.
- All customer segments face the same price for bundles.

Based on these assumptions they formulated a stochastic bundling problem. In the profit function, variable and fixed costs for each bundle and segment are considered. The probability that a customer from segment \(i\) buys the bundle \(j\) at price \(p_j\) is computed as \(\pi_{ij}(A, p, y, \alpha_i) \Phi_{ij}(p_j)\). Due to this non-linearity, the objective is a mixed-integer non-linear function. The authors therefore propose a solution method based on the decomposition of the stochastic bundling problem in an integer master problem and a non-linear sub problem.
Appendix B

Mathematical Appendix

B.1 Gumbel Distribution Properties

The following properties were extracted from Ben-Akiva and Lerman [BAL85]. If random variable $x$ distributes Gumbel with parameters $\eta$ (location) and $\mu$ (positive scale), then its cumulative distribution function is given by

$$F(x) = e^{-e^{-\mu(x-\eta)}}, \mu > 0$$

and the corresponding probability density function is given by

$$f(x) = \mu e^{-\mu(x-\eta)} e^{-e^{-\mu(x-\eta)}}, \mu > 0.$$ 

Some properties of this double exponential distribution:

1. The mode is $\eta$.
2. The mean is $\eta + \frac{2}{\mu}$, where $\gamma$ is the Euler constant ($\sim 0.577$).
3. The variance is $\pi^2/6\mu^2$.
4. If $\varepsilon \sim G(\eta, \mu)$ then for scalars $\alpha > 0$ and $\nu$, $(\alpha \varepsilon + \nu) \sim G(\alpha \eta + \nu, \mu/\alpha)$.
5. If $\varepsilon_1$ and $\varepsilon_2$ are i.i.d. $G(\eta_1, \mu)$ and $G(\eta_2, \mu)$ respectively, then:

   (a) $\varepsilon^* = \varepsilon_1 - \varepsilon_2$ distributes logistically with cdf $F(\varepsilon^*) = \left(1 + e^{\mu(\eta_2 - \eta_1 - \varepsilon^*)}\right)^{-1}$,
   (b) $\max(\varepsilon_1, \varepsilon_2) \sim G\left(\frac{1}{\mu} \ln(e^{\mu \eta_1} + e^{\mu \eta_2}), \mu\right)$,
   (c) $\max(\varepsilon_1, ..., \varepsilon_J) \sim G\left(\frac{1}{\mu} \ln \sum_{j=1}^{J} e^{\mu \eta_j}, \mu\right)$.

Figure B-1 depicts the shape of the standard Gumbel distribution’s probability density function in (a) and cumulative probability function in (b).
B.2 Lambert W Function

B.2.1 Properties and approximations

The Lambert $W(z)$ function is the solution of

$$W(z)e^{W(z)} = z \tag{B.1}$$

Figure B-2 shows that $W(z)$ is a monotonic increasing function. Thus, its derivative is positive for all values of $z$. Taking the derivative of (B.1) with respect $z$ yields

$$\frac{\partial W(z)}{\partial z} = e^{-W(z)} - \frac{\partial W(z)}{\partial z}ze^{-W(z)}$$

$$= \frac{e^{-W(z)}}{1 + ze^{-W(z)}}$$

$$= \frac{1}{e^{W(z)} + z}$$

or equivalently

$$\frac{\partial W(z)}{\partial z} = \frac{W(z)/z}{W(z) + 1} \tag{B.2}$$

The Lambert W function allows the following analytical representation of a power tower such as $h(x) = x^{x^{x^{\ddots}}}$, where $x^{x^x} = x^{(x^x)}$:

$$h(x) = -\frac{W(-\ln x)}{\ln x}$$
The Lambert function has the following series expansion

\[ W(z) = \sum_{n=1}^{\infty} \left( \frac{(-n)^{n-1}}{n!} z^n \right) \]

which cannot be used for practical numerical computation due to oscillation between large positive and negative values for real \( z \geq 0.4 \). Hence, in order to compute \( W(z) \), Corless et al. \[CGH^{+}95, \text{equation 4.18}\] state the following asymptotic formula which yields reasonably accurate results:

\[ W(z) = \log z - \log \log z + \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} c_{km} (\log \log z)^m (\log z)^{-k-m} \]

which is absolutely convergent and can be rearranged into the form

\[
W(z) = L_1 - L_2 + \frac{L_2}{L_1} + \frac{L_2 (-2 + L_2)}{2L_1^2} + \frac{L_2(6 - 9L_2 + 2L_2^2)}{6L_1^3} \\
+ \frac{L_2 (-12 + 36L_2 - 22L_2^2 + 3L_2^3)}{12L_1^4} + O\left(\left\{ \frac{L_2}{L_1}\right\}^5\right)
\]

where \( L_1 = \log z \) and \( L_2 = \log \log z \). The last term is less than 0.007 for all values of \( z \).

Another good, simple, and accurate approximation (see Figure B-3) of \( W(z) \) in (B.1) can be found in http://www.desy.de/~t00fri/qcdins/texhtml/lambertw/, which is the following:

\[
W(z) \approx \begin{cases} 
0.665(1 + 0.0195 \ln(z + 1)) \ln(z + 1) + 0.04 & : 0 \leq z \leq 250 \\
\ln(z - 4) - (1 - \frac{1}{\ln(z)}) \ln(\ln(z)) & : 250 \leq z 
\end{cases} \quad (B.3)
\]
Figure B-3: Lambert approximation error, where $\tilde{W}(z)$ denotes the approximation in (B.3)

### B.2.2 C code for Lambert W function, principal branch

```c
/* Lambert W function.
* Was ~/C/LambertW.c written K M Briggs Keith dot Briggs at bt dot com 97 May 21.
* Revised KMB 97 Nov 20; 98 Feb 11, Nov 24, Dec 28; 99 Jan 13; 00 Feb 23; 01 Apr 09
* Computes Lambert W function, principal branch.
* See LambertW1.c for -1 branch.
* Returned value W(z) satisfies W(z)*exp(W(z))=z
* test data...
* W(1) = 0.5671432904097838730
* W(2) = 0.8526055020137254914
* W(20) = 2.2050032780240599705
* To solve (a+b*R)*exp(-c*R)-d=0 for R, use
* R=-(b*W(-exp(-a*c/b)/b*d*c)+a*c)/b/c
* Test:
gcc -DTESTW LambertW.c -o LambertW -lm & & LambertW
* Library:
gcc -O3 -c LambertW.c
*/
#include <math.h>
#include <stdio.h>
double LambertW(const double z);
```

157
const int dbgW=0;
double LambertW(const double z) {
    int i;
    const double eps=4.0e-16, em1=0.3678794411714423215955237701614608;
    double p, e, t, w;
    if (dbgW) fprintf(stderr,"LambertW: z=%g \
");
    if (z <= em1 || isinf(z) || isnan(z)) {
        fprintf(stderr,"LambertW: bad argument %g, exiting. \
"); exit(1);
    }
    if (0.0==z) return 0.0;
    if (z < -em1) { // series near -em1 in sqrt(q)
        double q=z+em1, r=sqrt(q), q2=q*q, q3=q2*q;
        return
        -1.0
        +2.331643981597124203363536062168*r
        -1.812187885639363490240191647568*q
        +3.0668590105063191289314892704*r*q2
        -4.175335600258177138854984177460*q3
        +5.858023729874774148815053846119*r*q3
        -8.401032217523977370984161688514*q3*q; // error approx 1e-16
    }
    /* initial approx for iteration... */
    if (z < 1.0) { /* series near 0 */
        p=sqrt(2.0*(2.7182818284590452353602874713526625*z+1.0));
        w=1.0+p*(1.0+p*(-0.333333333333333333333333333333333+0.15277777777777777777777));
    } else
        w=log(z); /* asymptotic */
    if (z > 3.0) w-=log(w); /* useful? */
    for (i=0; i < 10; i++) { /* Halley iteration */
        e=exp(w);
        t=w*e-z;
        p=w+1.0;
        t/=exp(p-0.5*(p+1.0)*t/p);
        w -= t;
        if (fabs(t) < eps*(1.0+fabs(w))) return w; /* rel-abs error */
    }
    /* should never get here */
    fprintf(stderr,"LambertW: No convergence at z=%g, exiting. \
");
}
```c
exit(1);

#ifdef TESTW
/* test program... */
int main() {
    int i;
    double z, w, err;
    for (i = 0; i < 100; i++) {
        z = i / 100.0 - 0.3678794411714423215955;  // example values
        w = LambertW(z);
        err = exp(w) - z / w;
        printf("W(%.4f)=%.16f, check: exp(W(z)) - z/W(z)=%.e\n", z, w, err);
    }
    return 0;
}
#endif

#ifdef INTW
int main() {
    int i, n = 1000;
    double w, z, s = 0, err;
    for (i = 1; i <= n; i++) {
        z = i / (double)n;
        w = LambertW(1 / z) / (1 + z);
        s += w;
        printf("%.8f %.8f \n", z, w);
    }
    fprintf(stderr, "%.8f \n", exp(s / n / log(2)));
    return 0;
}
#endif
```
Appendix C

Computational Experiments Code

C.1 Bundle bounds source code

Some experiments were performed using AMPL as the modeling language to specify our model, and LOQO\(^1\) as the solver to find local optimal solutions. Here we present the source code in AMPL used for our problem.

```ampl
# Price Bundling eServices for SMBs
# Objective: Nonlinear
# Constraints: linear
# Reads the parameters from the file bundling.dat and bund_gen.dat
# General parameters
reset;
param n default 2; # Number of bundles
param m default 1; # Number of customer segments
param r default 0.008; # Discount rate per period
param T default 36; # Contract Length
param gen_info {1..1, 1..4};
data bund_gen.dat;
let n := gen_info[1,1];
let m := gen_info[1,2];
let r := gen_info[1,3];
let T := gen_info[1,4];

param maxp {1..n} default 0; # Maximum price in the uniform distribution
param Unp {1..n} default 0; # Utility of nonprice attributes

\(^1\)LOQO is a system for solving smooth constrained optimization problems, developed by Robert J. Vanderbei, Department of Operations Research and Financial Engineering at Princeton University.
```
# Utility of price

# Service cost per period

# Initial setup cost

# Minimum probability of leaving at any period

# parameter for q(p)

# parameter for q(p)

data bundles.dat;

let {j in 1..n} maxp[j] := bundinfo[j,1];
let {j in 1..n} Unp[j] := bundinfo[j,2];
let {j in 1..n} Up[j] := bundinfo[j,3];
let {j in 1..n} S[j] := bundinfo[j,4];
let {j in 1..n} K[j] := bundinfo[j,5];
let {j in 1..n} qmin[j] := bundinfo[j,6];
let {j in 1..n} a[j] := bundinfo[j,7];
let {j in 1..n} b[j] := bundinfo[j,8];

# Decision variables

var p {j in 1..n} >= S[j];
# Probability of leaving the system given prices

var q {j in 1..n} = qmin[j] * ( 1 + a[j]*( 1 - exp(-( (p[j]-S[j])/b[j] ) )));

# Customers willing to pay the price

var WillPay {j in 1..n} = 1 - ((p[j]-S[j])/(maxp[j]-S[j]));
# Utilities

var U {j in 1..n} = Unp[j] + Up[j] * (p[j]-S[j]);
# Choice probabilities

var ChProb {j in 1..n} = exp(U[j]) / sum{k in 1..n} exp(U[k]);
# Per-customer profit

var PCprofit {j in 1..n} = (((p[j]-S[j])/(q[j]+r)) * (1 - ((1-q[j])/(1+r))^(T)) - K[j];

maximize objective:

sum {j in 1..n} WillPay[j] * ChProb[j] * PCprofit[j];

subject to PriceRange {j in 1..n}: p[j] <= maxp[j];

option solver loqo;
option loqo_options ''verbose=2 timing=1'';
solve > answbund;
C.2 Pricing policy source code

The following source code was used in MatLab to develop optimal pricing policies for the dynamic programming problem addressed in Chapter 2.

```matlab
% *** Convol_Matrix.m ***
% Creates the matrix pr with all convolutions for all states and prices.

pr = zeros(d1+1,d2,d3+1);
for n=0:d3
    for i=1:d2
        pr(:,i,n+1)=PrCust2(d1,n,q(i/100),lambda(i/100));
    end
end
save('pr.mat','pr')

% *** Initialization.m ***
% Parameters initialization

clear;
K=10;
rho=1/(1+0.1);
d1=100; % Number of states (customers)
d2=1000; % Number of price point between 0 and 10
J_old = zeros(1,d1+1);
J_new = zeros(1,d1+1);
p_new = zeros(1,d1+1);

% *** Policy_Gen.m ***
% Policy generation

for i=1:200
    for n=0:d1
```
aux = (g(n, [1:1000]/100, K) + rho*E_J(J_old, n, [1:1000], pr));
[j, p] = max(aux);
J_new(n+1) = J;
p_new(n+1) = p/100;
end
J_old = J_new;
end

% *** Fixed_Policy_Gen.m ***
% Fixed-price Policy generation
% J =[];
for p=530:10:640
  J_old = zeros(1, d1+1);
  J_new = zeros(1, d1+1);
  for i=1:100
    for n=0:d1
      aux = (g(n, [1:1000]/100, K) + rho*E_J(J_old, n, [1:1000], pr));
      J_new(n+1) = aux(p)';
      p_new(n+1) = p/100;
    end
    J_old = J_new;
  end
  J = [J ; J_old];
end

% *** E_J.m ***
% Expected future return when n customers and pp/100 is the price
% function result = E_J(J, n, pp, pr);
% result = J*pr(:, pp, n+1);

% *** g.m ***
% Expected revenue of current period when price p and N customers
% function result = g(N, p, K);
% result = ((q(p)*N) .* p) - (lambda(p)*K);
% *** lambda.m ***
% Computes the effective arrival rate for a price vector p.
% External demand rate is 5, and a Uniform[0,10] is assumed
% as reservation price distribution.
% function result = lambda(p);
result = 5*(ones(1,1000)-(p/10));

% *** PrCust2.m ***
% Convolution between Poisson and Binomial pmf.
% It generates a vector of dimension d1+1
% function result = PrCust2(d1,N,q,LAMBDA);
m_poiss = tril( ones(d1+1) * diag(poisspdf([0:d1],LAMBDA)) );
m_bino = binopdf(flipud(hankel([d1:-1:0])))',N,q);
result = diag( m_poiss * m_bino );

% *** q.m ***
% Computes the probability that a customer will stay in the system
% one more period for a price vector p. We assume a logit form.
% function result = q(p);
result = exp(-p/3)./(exp(-p/3)+(0.04*ones(1,1000)));

% *** Simulation.m ***
% Using a pricing policy, simulates the system.
cust_evol = [];
prof=[[];
for r=1:1
profit = 0;
N = 5;
for i=0:100
w = poissrnd(lambda_ind(p_new(N+1)));
l = binornd(N,q_ind(p_new(N+1)));
profit = profit + (rho^i)*( l*p_new(N+1) - w*K );
cust_evol = [cust_evol [N;l;w]];
N = l+w;
end
prof = [prof profit];
end
plot(cust_evol')
profit;
Bibliography


