Title
The Design of Experiential Service Processes

Permalink
https://escholarship.org/uc/item/6176q4td

Author
DAS GUPTA, APARUPA

Publication Date
2014

Peer reviewed|Thesis/dissertation
The Design of Experiential Service Processes

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Management

by

Aparupa Das Gupta
ABSTRACT OF THE DISSERTATION

The Design of Experiential Service Processes

by

Aparupa Das Gupta

Doctor of Philosophy in Management
University of California, Los Angeles, 2014

Professor Guillaume Roels, Co-chair
Professor Uday S. Karmarkar, Co-chair

For many consumer-intensive (B2C) services, delivering memorable customer experiences to maximize customer satisfaction can be a source of competitive advantage since firms that deliver outstanding customer experience achieve greater customer satisfaction, and therefore greater customer loyalty. Yet, there is no clear methodology available for service encounter design that accounts for customer behavior. This dissertation shows how experiential services should be sequenced and timed to maximize satisfaction of customers who are subject to memory decay and acclimation.

The objective of the problem addressed in this thesis is to maximize the total remembered utility of customers from their experience of the service encounter. We consider service encounters where the service provider can alter either the sequence or the duration of activities across the encounter. In particular, we consider three stylized models of service encounters, (i) variable sequence and fixed duration (VSFD), (ii) fixed sequence and variable duration (FSVD), and (iii) variable sequence and variable duration (VSVD). The thesis is organized as follows.

The first chapter introduces the motivation behind the research question and the assumptions made to model the problem. The rationale for addressing the specific behavioral phenomena of memory decay and acclimation is also explained in this chapter.

This thesis draws on research from multiple disciplines. The second chapter provides an extensive literature review, spanning the areas of psychology, behavioral economics, decision
theory, marketing, and operations management.

In chapter three, a mathematical model of customer satisfaction that includes the effect of acclimation and memory decay is introduced. Several empirical research papers have studied the satisfaction obtained from experiences by people subject to the effects of memory and adaptation. Our analytical model for total remembered utility is further validated by comparing its outcome with the reported results obtained from these experiments.

Chapter four discusses the three scenarios of service encounters mentioned above, (i) VSFD, (ii) FSVD, and (iii) VSVD. Analytical results indicate that memory decay favors positioning the highest service level near the end, whereas acclimation favors maximizing the gradient of service level. Although memory decay and acclimation lead to the same design individually, they can act as opposing forces when considered jointly. Taken, together, they maximize the gradient of service level near the end, which may result in (i) sequencing activities in a U-shaped fashion and (ii) lengthening the duration of activities with the lowest service levels.

Chapter five discusses three heuristics for sequencing and duration allocation to design the service encounter. These heuristics are based on the analytical characterization of the optimal sequence and timing of activities derived in chapter three. The performance of the heuristics is benchmarked against the true optimum for small problem instances of VSFD and FSVD. For VSVD, the heuristic is benchmarked against an upper bound.

Chapter six discusses two extensions of the service provider’s design problem, (i) endogenous service levels, and (ii) service design for a population of customers who are heterogeneous in terms of their rates of acclimation and memory decay.

In case of endogenous service levels we assume that every activity can be assigned a service level within a range, by the service provider. The sequence and duration of activities is assumed to be fixed. The results indicate that it is optimal to allocate activities towards the end of the service encounter a higher service level, and the activities near the beginning of the encounter a lower service level.

For heterogeneous customers we find that the sequence and duration allocation rules that are optimal for homogeneous customers are now not always optimal. For this scenario
we numerically test the robustness of the service design that is optimal for homogeneous customers. We find that the sequence that is optimal for the mean value of the rates of acclimation and memory decay for the entire population of customers, serves as a good heuristic in the heterogeneous case.

Finally, in chapter seven we conclude by providing an overview of the results and provide directions for future research. The Appendix provides detailed proofs for analytical results.
The dissertation of Aparupa Das Gupta is approved.

Rakesh Sarin
Felipe Caro
Manel Baucells

Uday S. Karmarkar, Committee Co-chair
Guillaume Roels, Committee Co-chair

University of California, Los Angeles

2014
To my family,

for their unconditional love and support.
# Table of Contents

1 Introduction ................................................................. 1

2 Literature Review ........................................................... 6

2.1 Behavioral Aspects ......................................................... 6

2.1.1 Memory Decay ....................................................... 6

2.1.2 Acclimation .......................................................... 8

2.1.3 Decision Theory ...................................................... 8

2.1.4 Service Operations ................................................... 9

3 Memory Decay and Acclimation ............................................. 11

3.1 Introduction ............................................................. 11

3.2 Customer Satisfaction Utility ......................................... 11

3.2.1 Customer Memory Process ......................................... 12

3.2.2 Customer Acclimation Process ..................................... 13

3.2.3 Customer Satisfaction Model ....................................... 16

4 Service Provider’s Design Problem: Sequence and Duration ............ 19

4.1 Introduction ............................................................. 19

4.2 Analysis ................................................................. 22

4.2.1 Variable Sequence and Fixed Duration ............................ 23

4.2.2 Fixed Sequence and Variable Duration ............................ 26

4.2.3 Variable Sequence and Variable Duration ........................ 29

5 Heuristics and Bounds ....................................................... 31

5.1 Introduction ............................................................. 31
List of Figures

3.1 Acclimation affects instantaneous utility and memory decay determines its relative weight in the overall remembered utility. ........................................ 12

3.2 Colonoscopy experiment showing dissatisfaction at the end of experience with \( w = 0.3 \) per minute ................................................................. 13

3.3 Acclimation process leads to a decay in utility over time for a fixed service level and yields negative utility after a drop in service level. .................... 15

3.4 Online measures of enjoyment of the looped song in Nelson and Meyvis [2008] as a function of time with \( \alpha = 0.1, u_0 = 60 \) where \( x(t) \) is the service level, \( b(t) \) is the acclimation level, and \( u(t) \) is the instantaneous utility level. ........ 16

3.5 The weight assigned to change in service levels between activity \( k - 1 \) and activity \( k \) is pseudoconcave in the remaining time until the end of the encounter. 17

4.1 Logarithmic mean \( m(\alpha, w) \). .................................................... 21

4.2 The optimal sequence of activities is either increasing or U-shaped, depending on the magnitude of \( T - \frac{2}{m(\alpha, w)} \). The last two activities are always in increasing order. ................................................................. 24

4.3 Optimal duration allocation for increasing subsequences. Shaded activities are allocated duration above their lower bound. ......................... 27

4.4 Optimal duration allocation for decreasing subsequences. Shaded activities are allocated duration above their lower bound. ......................... 27

4.5 Optimal sequence and duration allocation as \( m(\alpha, w) \) increases. Shaded activities are allocated duration above their lower bound. .............. 30

6.1 Function \( g(t) \) vs. time left to end the encounter \( (t) \) with \( w = 1 \) and \( \alpha = 0.5 \). 43
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Three stylized models of a service encounter</td>
<td>20</td>
</tr>
<tr>
<td>5.1</td>
<td>Optimality Gap of VSFD Heuristic</td>
<td>33</td>
</tr>
<tr>
<td>5.2</td>
<td>Heuristic Solution vs LINDOGlobal Solution for FSVD</td>
<td>35</td>
</tr>
<tr>
<td>5.3</td>
<td>Optimality Gaps for the VSVD heuristic with $T &gt; \frac{1}{m(\alpha,w)}$</td>
<td>39</td>
</tr>
<tr>
<td>5.4</td>
<td>Optimality Gaps for the VSVD heuristic with $T &gt; \frac{2}{m(\alpha,w)}$</td>
<td>40</td>
</tr>
<tr>
<td>6.1</td>
<td>Range of parameters</td>
<td>47</td>
</tr>
<tr>
<td>6.2</td>
<td>Duration $T$ for $t_i \sim \text{Gamma}(k, \theta)$</td>
<td>47</td>
</tr>
<tr>
<td>6.3</td>
<td>Variance($m(\alpha,w)$) corresponding to $\alpha \sim \text{Truncated Normal}(\sigma(\alpha), (0, \infty))$ and $w \sim \text{Truncated Normal}(\sigma(w), (0, \infty))$</td>
<td>48</td>
</tr>
<tr>
<td>6.4</td>
<td>Average suboptimality gap in % for crescendo</td>
<td>48</td>
</tr>
<tr>
<td>6.5</td>
<td>Average suboptimality gap in % for mean $(\bar{\alpha}, \bar{w})$</td>
<td>49</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

This dissertation would not have been possible without the guidance from my committee members. I am extremely grateful to my advisors, Professor Uday S. Karmarkar and Professor Guillaume Roels, who have been exceptional mentors and have facilitated my development as a scholar. I learnt from them how to formulate an idea into a model, and how to write technical results effectively. I feel privileged to have them as my mentors. I sincerely thank them for their encouragement and support during this period. I am indebted to Professor Rakesh Sarin for offering a brilliant course in Decision Theory that inspired me to conduct my research at the interface of Decision Theory and Operations Management. I am very thankful to Professor Felipe Caro, whose mentorship taught me how to think about broad insights from complicated mathematical analysis. I also had the opportunity to be a Teaching Assistant for the courses he offered in Operations Management. I thoroughly enjoyed attending his classes and I learnt a lot regarding the practical side of Operations Management. I am very grateful to Professor Manel Baucells for offering his invaluable insights and for giving me the opportunity to learn how to review research papers. His insightful suggestions enabled me to significantly improve my research. I am sincerely thankful to all my committee members for taking time out from their busy schedules to attend my practice presentations and for offering me invaluable advise on how to improve my research and my presentation skills.

I would also like to thank many other faculty members who helped me in the process: Professor Reza Ahmadi, Professor Kumar Rajaram, and Professor Christiane Barz. They all imparted knowledge in various disciplines of Operations Management and equipped me with the analytical skills required to address research questions. I am also sincerely grateful to Professor Charles Corbett for the advice, guidance, and encouragement that he offered me when I was a Teaching Assistant for the course he offered in Operations Management.

In addition, I am very grateful to the UCLA Anderson Graduate School of Management and the UCLA Graduate Division for the generous funding over the past five years. I would also like to gratefully acknowledge the Easton Technology Leadership Program at the An-
derson School for the financial support that has provided me with much needed resources to conduct research and to travel to conferences to present my work. I also thank Lydia Heyman for her administrative support during the PhD program.

I thank all my colleagues in the PhD program for helping provide a collegial and fun work environment during the past five years. We had some wonderful memorable times together both within and outside the office.

Finally, I wish to acknowledge the unconditional love and support that I have received from my family throughout the years of my graduate studies. In particular, I would like to thank my husband, Abhimanyu, for his unwavering support and patience which helped me to move forward throughout my PhD program.
VITA

2006
B.S. and M.S., Math and Computing.
Indian Institute of Technology, Kharagpur.

2008
M.S., Industrial Engineering.
Purdue University, West Lafayette.

2014 (expected)
M.S., Statistics.
University of California Los Angeles.

2009–2014
Graduate Student Researcher.
UCLA Anderson School of Management, Los Angeles.

PRESENTATIONS


6. Best Student Paper Award-Service Science Session, INFORMS 2013, The Design of Experiential Services: Optimal Sequence and Duration of Service Activities.
CHAPTER 1

Introduction

Managing customer experience during service encounters is a powerful way of improving customer satisfaction and differentiating a firm from its competitors [Pine and Gilmore, 1998, Voss and Zomerdijk, 2007, DeVine and Gilson, 2010]. For instance, the Walt Disney Company’s entertainment parks and Benihana’s restaurants are two well known examples of companies offering unique service experiences to their customers [Heskett et al., 1997]. Maximizing customer satisfaction often entails understanding customer behavior. DeVine et al. [2012] indeed report that companies focusing on the human side of service delivery have been able to improve customer satisfaction scores by up to 30% while reducing costs.

Since an experience is a dynamic process, the utility a customer derives from a service experience evolves dynamically, in response to various stimuli experienced during the encounter. For instance, in a restaurant, such stimuli could include the level of sweetness of the food or the level of noise in the room. In principle, these stimuli are multidimensional (e.g., physiological and cognitive) and the way they affect a customer’s utility can be complex. For instance, the utility a customer may derive from a good bottle of wine may depend not only on its taste (physiological), but also on the knowledge of its price (cognitive); see [Plassmann et al., 2008]. Because these stimuli may change over time, a customer’s utility may evolve dynamically [Bitran et al., 2008, Baucells and Sarin, 2012]. The total utility, or satisfaction, generated through an experience, and evaluated at the end of the encounter, will therefore somehow integrate all these instantaneous utilities [Edgeworth, 1881].

The utility formation and integration processes are subject to well-documented behavioral biases. For instance, the total utility from an experience tends to be highly correlated to the peak and the final utilities [Varey and Kahneman, 1992]. In addition, greater weight may be
placed on the initial experience (aka the primacy effect) due to attention decrement [Crano, 1977] or on the final experience (aka the recency effect) due to memory decay [Ebbinghaus, 1913]. Biases also affect how much (instantaneous) utility is derived from an outside stimulus. For instance, acclimation makes people adapt to states but react to changes [Hsee and Tsai, 2008] and satiation reduces the marginal utility obtained from accumulated consumption [Baucells and Sarin, 2013]. In addition, phenomena, such as savoring and dread, may also affect the utility experienced before an experience [Loewenstein, 1987].

As a result, maximizing customer satisfaction requires a deep understanding of customer behavior [Chase and Dasu, 2008], both during and between encounters [Bitran et al., 2008]. In fact, Cook et al. [2002] argue that service encounters can be designed with the same depth and rigor as manufacturing processes by drawing concepts from psychology and sociology.

While research on service operations has traditionally focused on capacity sizing and pricing decisions [Haviv, 2013, Talluri and Van Ryzin, 2004], designing a service experience involves many other decisions. Consider for instance a one-week executive education course. Once the location and the program content have been selected, one must still decide the schedule of the different modules, the duration of each module, the company visits, and the timing of the main reception. Potentially, all these decisions impact customer satisfaction.

Despite the ubiquity of service experiences in today’s economy [Pine and Gilmore, 1998] and the impact of operational decisions on customer satisfaction, there exist to date few guidelines for designing service experiences. In this research, we take a first step in that direction by considering a particular service encounter with the following characteristics:

- The encounter consists of a predetermined set of homogenous activities and has a fixed total duration.
- Each activity is associated with a given unidimensional service intensity, independent of its placement in the sequence and its duration.
- Customers are captive, i.e., customers cannot arrive after the beginning of the encounter or leave before its end.
- Customers are subject to memory decay and acclimation.
The provider has control over the sequencing of activities and/or the allocation of duration to them.

Although stylized, our model could apply, as a first-order approximation, to a short executive education program on a specific topic (e.g., current trends in supply chain management). The different activities would correspond there to the different sessions offered (e.g., sustainability, offshoring), and their respective service intensities could be measured by their historical teaching ratings. Based on that information, certain schedules of sessions may be more desirable than others along the dimension of teaching quality. Other examples of service encounters where sequence and duration of events affect the overall experience are concerts, magic shows, museum tours across multiple galleries, city walking or bus tours, or medical procedures like colonoscopy, or dental visits.

In such settings, we show how to sequence and allocate duration to activities in a service encounter so as to maximize customer satisfaction, as evaluated at the end of the encounter. The implicit assumption behind this objective is that greater customer satisfaction will drive future sales, either through repeated purchases or referrals [Heskett et al., 1997, Bitran et al., 2008]. Because our choice of design variables pertains to the encounter, what happens before or after (anticipation and recall) is out of scope of our analysis.

We refer to the combination of sequence and duration of activities as a service design, and assume that it is under the control of the service provider, if it can be changed. Depending on the nature of the service offered, the service provider may be able to alter the sequence and duration of service activities. In particular, in some service encounters (e.g., a music concert with multiple performances) it is feasible to change the order of activities but not their durations. Other service encounters (e.g., treatments in a spa) offer a fixed sequence of activities, but their durations can vary. Finally, some encounters (e.g., fitness classes) offer the flexibility to alter both the sequence and duration of activities during the service encounter. Based on the nature of services, we consider three stylized models of a service encounter, 1) variable sequence and fixed duration (VSFD), 2) fixed sequence and variable duration (FSVD), and 3) variable sequence and variable duration (VSVD).

Based on their experimental findings, psychologists have proposed catchy sequencing
guidelines, such as “finish strong”, “get the bad experience out of the way early”, “segment
the pleasure, combine the pain” [Chase and Dasu, 2001]. Following those guidelines, many
service firms (e.g., Royal Caribbean cruise line, office furniture manufacturer Herman Miller)
have ensured that their customers start or end their service experience on a high note [Voss
and Zomerdijk, 2007]. Although these best practices are often justified on the grounds of the
primacy and recency effects, these effects tend to describe more the symptoms than the true
underlying behavioral biases (which could be, for instance, attention decrement and memory
decay). Moreover, these best-practice sequencing rules can be justified on the grounds of
other biases, such as satiation [Baucells and Sarin, 2013], in a more subtle way. Finally,
these guidelines are often imprecise about what should happen between the beginning and
the end of the sequence. For instance, whereas a U-shape sequence is generally recognized
as an optimal design, either to maximize a weighted sum of peak utility, end utility, and
the spread between the peak and the end utilities [Dixon and Verma, 2013] or to maximize
discounted utility in the presence of satiation [Baucells and Sarin, 2013], there is no guide-
line about the steepness of the gradients, nor about when a plain crescendo or diminuendo
dominates a U-shape sequence.

In this research, we derive formal sequencing and duration allocation guidelines in the
presence of two specific behavioral biases: memory decay and acclimation.

Memory decay refers to the fact that memory of human beings fade with the passage of
time. Hence, individuals tend to pay more importance to events closer to the present than
to the ones farther away from it. As Kahneman [2010] puts it, we dont choose between
experiences, we choose between memories of experiences.

Acclimation refers to the fact that human beings adapt, or acclimate, to a given envi-
rionment over a period of time, but react to changes. As Ariely [1998] suggests, because
physiological adaptation causes diminishing nerve activation for longer experiences, the in-
tensity of experience becomes less sensitive to a constant service level.

Although other behavioral biases (e.g., satiation) may affect customer satisfaction, we
focus on these two biases for the following reasons. First, they have a long tradition in
psychology and have therefore received strong empirical support as reviewed in Section 2.
Second, these seemingly simple biases are sufficiently rich to explain classical findings in psychology, such as that more pain can be preferred to less [Kahneman et al., 1993]. Third, they appear mathematically symmetrically in our models, and lead to the same designs individually, but differently in combination.

While memory decay favors a high service level near the end, acclimation favors a steep rise in service levels. Despite the conceptually different mechanisms, we show that a service design that is optimal for memory decay alone is also optimal for acclimation alone. In particular without precedence constraints, it is optimal to sequence activities in increasing order of service levels so as to finish at a peak and, with variable durations, to lengthen the duration of those activities with the highest service level.

Taken together, acclimation and memory decay favor a maximum gradient of service level closer to the end of the encounter. Specifically, when both phenomena are present, it may be optimal to sequence activities in a U-shaped fashion, beginning and finishing with high service levels, and to lengthen the duration of activities with the lowest service level.

Hence, even though memory decay and acclimation act in the same way when considered individually, they act in opposite directions when considered jointly. Finally, since the resulting design problems are hard to solve optimally, we propose heuristics based on the analytical characterization of the optimal design.
CHAPTER 2

Literature Review

This research is related to three bodies of research: the study of human behavior as in psychology and marketing, models of experienced utility theory in decision theory, and the operational aspects of service design in operations management. We next review work from these three streams of literature that is relevant for this research.

2.1 Behavioral Aspects

Research in psychology and marketing has identified many factors that influence the retrospective evaluation of experiences, including the trend of an experience [Loewenstein and Prelec, 1993], its rate of change [Hsee and Abelson, 1991], and its maximum and final intensities [Frederickson and Kahneman, 1993]. Underlying these factors are psychological phenomena such as adaptation, loss aversion, or memory decay. We contribute to this literature by adopting a design perspective on two such phenomena, namely, memory decay and acclimation.

2.1.1 Memory Decay.

In their seminal work, Varey and Kahneman [1992], Frederickson and Kahneman [1993], and Kahneman et al. [1993] find that the maximum and final intensities of an experience play pivotal roles in influencing the remembered utility from the overall experience. Studying satisfaction from listening to a piece of music, Rozin et al. [2004] also find that satisfaction is highly correlated with the peak and the final intensities of the piece of music. Although Kahneman’s work suggests that the overall evaluation of an experience can be summarized by
a weighted average of the peak and final intensities (i.e., the peak-end rule), regardless of the
duration of the experience, subsequent research has also shown the importance of duration
that people are sensitive to duration if subject to changing levels of painful stimuli. Ariely
[1998] moreover finds that retrospective evaluation of painful experiences is influenced not
only by the final pain intensity, but also by the intensity trend during the latter half of the
experience. In the aforementioned music study, Rozin et al. [2004] also find that the difference
in intensity between two consecutive moments affects global satisfaction. Studying memory
of ads, Baumgartner et al. [1997] report that, even if people prefer advertisements with peaks
and high ends, retrospective evaluation of advertisements is not always independent of their
durations, especially if duration affects the intensity of the peak experience and the final
moment.

In this research, we adopt an agnostic approach on the peak-end rule, i.e., we do not
cast the remembered utility as a function of the peak and the end experience. Rather, we
represent memory decay with a simple model of reverse discounted utility.

There are several models of memory decay. Studying how much memory of monosyllabic
words was retained after several days, Ebbinghaus (1913) constructed an exponential memory
decay curve. Later, Wickelgren [1974] found that a power function was a slightly better
predictor of memory retention; yet, both models account for 99% of the variance of the
original data collected by Ebbinghaus [Wixted and Carpenter, 2007]. Naik [1999] uses an
exponential decay model to estimate the half-life of advertising campaigns. In this research,
we choose to model memory decay with an exponential curve for mathematical tractability
and consistency with the traditional forward-discounting utility models.

The extent of memory decay is of course context-dependent: While data from Ebbinghaus
suggests that the half life of the memory of monosyllabic words is less than a couple of
hours, Naik [1999] reports that the half life of an advertising campaign can be as long as
three months. On the other hand, Kahnemenan’s physiological experiments mentioned above
identifying the peak-end rule lasted less than a few minutes.
2.1.2 Acclimation.

There is considerable evidence available regarding adaptation, or acclimation as it is referred for short experiences, in life experiences [Helson, 1964, Brickman and Campbell, 1971, Brickman et al., 1978]. In physiology, acclimation goes as far back as Newton’s law of cooling [Carlson, 1963, Incropera et al., 2011]. Ariely [1998] observes that adaptation can explain the impact of duration on retrospective evaluations of patterned stimuli and its lack thereof for constant stimuli. Nelson and Meyvis [2008] experimentally verify that, contrary to conventional wisdom, interruptions may make retrospective evaluation of pleasant experiences more enjoyable and that of unpleasant experiences more irritating, and they attribute this result to the adaptation effect. Further, Nelson et al. [2009] show that if there is less adaptation during an experience then interruptions need not make a positive experience more enjoyable.

In this research we model acclimation as a dynamic reference effect consistent with decision theory [Baucells and Sarin, 2013]: At each point in time, a customer’s instantaneous utility is a function of the difference between the current service level and the current reference level, and that reference level dynamically adapts to the service level. Similar to memory decay, the speed of acclimation is highly context dependent. For physiological processes (e.g., gustatory or olfactory adaptation), complete adaptation is often in the order of minutes [Gent and McBurney, 1978, Dalton and Wysocki, 1996] or even seconds [Dawes and Watanabe, 1987]. On the other hand, [Nelson and Meyvis’ 2008] experiments on commercial interruptions span several minutes.

2.1.3 Decision Theory

Kahneman et al. [1997] propose that there exist two kinds of experienced utility, the one reported in real time (instantaneous) and the other based on retrospective evaluation (remembered). Using this idea, our model represents the remembered utility as a function of the instantaneous utility.

Baucells and Sarin [2013] propose six laws that govern satisfaction including habit for-
ation, satiation, and social comparison. Consistent with the adaptation or habit formation models in Constantinides [1990], Wathieu [1997], and Baucells and Sarin [2013], our acclimation model adapts the reference level to the current service level.

We contribute to this literature by adopting a different perspective on experience design, focusing on maximizing ex-post satisfaction as opposed to ex-ante discounted utility. Proposing an analytical utility model for adaptation before, during, and after an event, Baucells and Belleza [2012] show that the utility function during anticipation is unimodal (decreasing, increasing, or U-shaped) because the magnitude of effective consumption, i.e., the difference between the conceptual consumption of experience during anticipation and the current reference level, decreases over time, whereas the discount function due to adaptation, which is inversely proportional to the effective consumption, increases over time. We focus here on the event itself, considering an initial reference level, which captures the customer anticipation at the start of the event, and a moment of recall at the end of the encounter, when the remembered utility is evaluated. Moreover, the discount factor due to acclimation remains constant, i.e., is independent of the magnitude of the instantaneous utility. With this utility model, we find that the optimal sequence of activities under certain conditions is U-shaped in service levels. Consistent with our result, Baucells et al. [2013] empirically find that the sequences of songs in music concerts and individual playlists exhibit a U-shape, based on their popularity or individual ratings respectively.

2.1.4 Service Operations

The management of service operations typically consists of managing capacity (e.g., staffing) [Pinedo, 2005], scheduling resources [Pinedo, 2005, Sampson and Weiss, 1995], and pricing [Talluri and Van Ryzin, 2004]. In the context of service encounter design, Bellos and Kavadias [2011] study how to allocate tasks between the service provider and the customer and how to price the service. In contrast to their model, which considers customer experience as random, we adopt here a behavioral model of customer experience. Moreover, our decision variables are not pricing and task allocation, but task sequencing and timing. Hence, our approach can be viewed as complementary to theirs.
Lately, some papers have studied how operational decisions are affected by customer behavior. The literature in this area is diverse in terms of modeling paradigms and behavioral regularities. In particular, Nasiry and Popescu [2011] study pricing and Dixon and Verma [2013], Dixon and Thompson [2013b,a] study event scheduling decisions in the light of the peak-end rule. Aflaki and Popescu [2013], Popescu and Wu [2007], and Nasiry and Popescu [2011] study how loss aversion affects pricing and service level policies. Caro and Martínez-de-Albéniz [2012] show how satiation affects assortment decisions. Adelman and Mersereau [2013] study how customer memory effects impact a supplier’s profits when the supplier dynamically allocates limited capacity among a portfolio of customers. Focusing on queueing experiences, Plambeck and Wang [2012, 2013] consider the effect of hyperbolic discounting on unpleasant services, Carmon et al. [1995] show how customer dissatisfaction can be reduced by spreading the service across the waiting period in queues, and Huang et al. [2013] show that estimating waiting times is impeded by customer bounded rationality.

This work is most closely related to Dixon and Thompson [2013b,a], who study an event scheduling problem by creating bundles of events in the presence of peak-end, trend, and other outcomes of psychological phenomena. Unlike these two papers, we directly model the psychological phenomena, and not their outcomes. Moreover we adopt an analytical approach, whereas they focus on the development of metaheuristics.
CHAPTER 3

Memory Decay and Acclimation

3.1 Introduction

We consider a service provider who seeks to maximize ex-post customer satisfaction from a service encounter. The service encounter consists of a sequence of \( n \) activities. Each activity \( i \in \{1,..,n\} \) is designed to offer a fixed service level \( x_i \), for a duration \( t_i \) with \( \sum_{i=1}^{n} t_i = T \). We assume that the service level is unidimensional and is independent of the activity’s placement in the sequence or its duration. For instance, each day of a week-long executive education course may consist of different sessions (e.g., operations, finance), each with a specific duration (e.g., one or two hours) and a specific service level (e.g., based on instructor’s ratings). We consider the following two design variables: 1) the sequence of activities \((i_1,..,i_n)\) where \( i_k \in \{1,..,n\} \) is the \( k \)th activity in the sequence and 2) the duration for each activity \( t_i, \forall i \). For any sequence \((i_1,..,i_n)\), let \( T_k = \sum_{j=1}^{k} t_i \) be the time passed by the end of the \( k \)th activity and \( T_k = \sum_{j=k}^{n} t_i \) be the time remaining from the beginning of the \( k \)th activity to the end of the encounter. We next build the customer satisfaction utility model and then consider the service provider’s design problem.

3.2 Customer Satisfaction Utility

We consider a customer who is subject to memory decay and acclimation. We assume the customer’s instantaneous experienced utility at activity \( i_k \) is captured by a unidimensional utility function \( u_{ik}(t) \) where \( T_{k-1} \leq t \leq T_k \). Acclimation affects how much instantaneous utility \( u_{ik}(t) \) is obtained from the service level \( x_k \) of activity \( k \), whereas memory decay
determines the relative weight of the contribution of \( u_{ik}(t) \) in the overall remembered utility \( S((i_1, \ldots, i_n), t) \) from the experience as illustrated in Figure 3.1.

![Diagram](image)

Figure 3.1: Acclimation affects instantaneous utility and memory decay determines its relative weight in the overall remembered utility.

3.2.1 Customer Memory Process.

We model the memory decay process based on the exponential nature of forgetting proposed by Ebbinghaus [1913]. According to this model, memory loss occurs rapidly initially, but more gradually subsequently. The memory decay process thus weighs experiences at the end of the encounter more heavily than the ones at the beginning of the encounter. Accordingly, memory decay operates like a backward discounting process. When the customer discounts past experiences exponentially with rate \( w \in [0, \infty) \), her remembered utility or satisfaction from the service experience is given by

\[
S((i_1, \ldots, i_n), t) = \sum_{k=1}^{n} \int_{T_{k-1}}^{T_k} u_{ik}(t)e^{-w(T-t)}dt. \tag{3.1}
\]

With no acclimation, if \( u_{ik}(t) = x_{ik}, \forall k \), then satisfaction from memory decay is given by a discounted sum of the difference of service levels across consecutive activities. Formally,

\[
S((i_1, \ldots, i_n), t) = \sum_{k=1}^{n} (x_{ik} - x_{ik-1}) \frac{(1 - e^{-wT_k})}{w}.
\]

As an illustration, we apply Equation 3.1 to the data obtained from the colonoscopy experiment in Redelmeier and Kahneman [1996]. In Figure 3.2 we plot the instantaneous utility
Figure 3.2: Colonoscopy experiment showing dissatisfaction at the end of experience with $w = 0.3$ per minute

reported by patients experiencing two different colonoscopy procedures, one short (left) and one long (right). The memory decay rate is assumed to be $w = 0.3$ per minute, which corresponds to a half life of about 2.3 minutes. With these parameters, we find that the patient going through the short operation experiences a total discomfort of $S = -12.1$, whereas the patient going through the long operation experiences a total discomfort of $S = -6.9$. Hence in our model, the patient subject to the longest operation overall reports less discomfort than the patient subject to the shortest operation, providing theoretical support to the authors’ claim that more pain can be preferred to less. Although there exist of course other interpretations of that result (including acclimation), this illustrates that memory decay alone is rich enough to explain this surprising outcome.

### 3.2.2 Customer Acclimation Process.

The acclimation process affects the instantaneous experienced utility. Let $b(t)$ be the customer’s reference level at time $t$. Similar to Baucells and Sarin [2013], we assume that the instantaneous utility experienced at time $t$ is a function of the difference between the service level and the reference point:

$$u_{ik}(t) = U(x_{ik} - b(t))$$

As defined by Kahneman and Tversky [2000], the instantaneous experienced utility, $u_{ik}(t)$, captures both the valence (good or bad) and the intensity (mild to extreme) of the instantaneous experience.
The rate of change of the reference point is proportional to the difference between the service level and the reference point, akin to the adaptation models used in life sciences (e.g., Overbosch 1986), or more generally, to Newton’s law of cooling:

\[
\frac{db(t)}{dt} = \alpha(x(t) - b(t)),
\]

where \( \alpha > 0 \) is the rate of acclimation. The value of \( \alpha \) can be estimated empirically (e.g., see Dawes and Watanabe 1987). A high value of \( \alpha \) implies that the customer is less influenced by past service levels. This model is similar to the classical habituation model of Wathieu [1997], although we do not restrict the acclimation parameter \( \alpha \) to be less than one. Solving this differential equation for activity \( i_k \) yields,

\[
b(t) = x_{i_k} - (x_{i_k} - b(T_{k-1}))e^{-\alpha(t-T_{k-1}}), T_{k-1} \leq t \leq T_k, \tag{3.2}
\]

where the initial reference level \( b(0) \) captures the history of past experiences.

We next make the following three assumptions:

1. The acclimation parameter \( \alpha \) is common across all activities of the encounter.
2. The reference point at the end of an activity carries over to the beginning of the following activity, i.e., \( \lim_{t \uparrow T_k} b(t) = \lim_{t \downarrow T_k} b(t) \).
3. The utility function is linear, \( U(x - b) = u_0 + (x - b) \), where \( u_0 \) is the intrinsic utility from the experience, normalized to zero in the sequel since it does not affect the design decisions.

The first two assumptions are direct consequences of our assumptions that activities are homogenous, such as different sessions in a one-day executive program, different massages in a spa treatment, or different tricks in a magic show. If activities were heterogeneous (e.g., waiting for a table and eating dinner), then potentially the level of acclimation could differ between activities and the reference point could be discontinuously reset once a new activity is started.

The last assumption is perhaps the most limiting. In practice, utility could be nonlinear, potentially exhibiting different characters for gains and losses [Kahneman and Tversky, 2000]. Moreover, it could be non-monotone. For instance, if \( x \) represents the outside temperature, it is likely that \( u(x) \) would peak when \( x \) is close to 70 degrees.
Under these assumptions, acclimation may yield negative utility even if the service level is positive, in contrast to satiation [Baucells and Sarin, 2013].

Figure 3.3: Acclimation process leads to a decay in utility over time for a fixed service level and yields negative utility after a drop in service level.

Figure 3.3 shows the acclimation process experienced by a customer during a service encounter consisting of three activities $x_1, x_2, x_3$ such that $x_2 > x_3 > x_1$. The acclimation level $b(t)$ always trails the service level. Therefore, for a constant service level $x(t)$, the instantaneous utility decreases over time and eventually becomes zero [Ariely, 1998]. In contrast, an upward or downward jump in service level across consecutive service levels leads to a rapid increase or decrease in the instantaneous utility level respectively.

Expanding (3.2), we obtain

$$b(t) = x_{i_k} - \left( (x_{i_1} - b(0)) + \sum_{j=2}^{k} (x_{i_j} - x_{i_{j-1}}) e^{\alpha T_{j-1}} \right) e^{-\alpha t}, T_{k-1} \leq t \leq T_k.$$  

Therefore, the instantaneous utility is given by

$$u_{i_k}(t) = (x_{i_1} - b(0)) e^{-\alpha t} + \sum_{j=2}^{k} (x_{i_j} - x_{i_{j-1}}) e^{-\alpha(t-T_{j-1})}, T_{k-1} \leq t \leq T_k. \quad (3.3)$$

The instantaneous experienced utility, $u_{i_k}(t)$, can thus be expressed as a discounted sum of the past changes in service levels. In this case, discounting operates forward in time, in contrast to memory decay, which applies backward discounting.

As an illustration, we apply the acclimation model to the data from the break insertion experiment performed by Nelson and Meyvis [2008, p 660]. The inclusion of break in the
Figure 3.4: Online measures of enjoyment of the looped song in Nelson and Meyvis [2008] as a function of time with $\alpha = 0.1$, $u_0 = 60$ where $x(t)$ is the service level, $b(t)$ is the acclimation level, and $u(t)$ is the instantaneous utility level.

experience disrupts the adaptation process. As shown in Figure 3, the acclimation model indicates that if there is no break (left figure), the instantaneous experienced utility gradually decreases over time, whereas if there is an unpleasant break (right figure), the instantaneous utility initially drops and then goes back up at then end of the break. Moreover, the utility experienced after the unpleasant break dominates the utility of the customers who didn’t experience the break, consistent with Figure 3 in Nelson and Meyvis [2008]. This example illustrates that our acclimation model, although simple, can explain why inserting unpleasant breaks may result in higher utility following the break. Moreover, we can show that, with acclimation and memory decay, the overall satisfaction from the experience may be higher with an unpleasant break than without, as was observed by Nelson and Meyvis [2008].

3.2.3 Customer Satisfaction Model.

Combining (3.1) and (3.3) leads to the following model of cumulative remembered utility in the presence of acclimation and memory decay:

$$S((i_1, \ldots, i_n), t) = \sum_{k=1}^{n} \int_{T_{k-1}}^{T_k} u_{i_k}(t)e^{-w(T-t)} dt = \sum_{k=1}^{n} (x_{i_k} - x_{i_{k-1}}) \frac{(e^{-\alpha T_k} - e^{-w T_k})}{w - \alpha}.$$  

(3.4)
Equation (3.4) reveals that, under our modeling assumptions, memory decay and acclimation mathematically play a symmetric role on ex-post satisfaction. Specifically, the customer weights a change in service intensity from the \((k-1)\)th to the \(k\)th activity by

\[
\Phi(\alpha, w, T_k) = e^{-\alpha T_k} - e^{-w T_k} = \int_0^{T_k} e^{-\alpha t} e^{-w (T_k - t)} dt. \tag{3.5}
\]

Whereas memory decay “discounts” the service level increment \((x_i^k - x_i^{k-1})\) backwards from the end of the encounter to the activity time, acclimation “discounts” it forward, from the beginning of the encounter to the activity time. Although the mathematical symmetry between memory decay and acclimation follows from our modeling assumptions (e.g., exponential decay, linear utility function), we nevertheless find it intriguing given the different natures of these two phenomenon and the way they affect total satisfaction, see Figure 3.1.

Rewriting (3.4) as follows:

\[
S = \sum_{k=1}^{n} (x_i^k - x_i^{k-1}) \Phi(\alpha, w, T_k) = \sum_{k=1}^{n} x_i^k (\Phi(\alpha, w, T_k) - \Phi(\alpha, w, T_{k+1})) \tag{3.6}
\]

the mathematical symmetry between acclimation and memory decay suggests that memory decay which is fundamentally a weighting of the activity service levels \(x_i^k\), can also be
interpreted as a weighting of the activity increments \((x_{ik} - x_{ik-1})\), and, conversely, that
acclimation, which is fundamentally a weighing of the service level increments, can also
be interpreted as a weighting of the service levels themselves. In addition, \(\Phi(\alpha, w, T_k)\) is
concave-convex, and the inflection point occurs at \(T_k = \frac{2}{m(\alpha, w)}\); see Lemma A3. Based on
(3.6), this shows that the weights on service levels are decreasing up to time \(\frac{2}{m(\alpha, w)}\) and then
increasing.
CHAPTER 4

Service Provider’s Design Problem: Sequence and Duration

4.1 Introduction

We consider a service provider who seeks to optimize the design of a service encounter to maximize ex-post customer satisfaction. We assume that the population of customers is homogeneous, i.e., they have the same values of the parameters \( \alpha \) and \( w \) or that the service design can be adapted to each customer, based on their specific degrees of memory decay and acclimation (provided these have been elicited).

Every service encounter is composed of activities that offer pleasant experiences (e.g., a performance during a concert) and unpleasant experiences (e.g., a painful medical procedure). Due to resource limitations, activity durations are typically constrained. In particular, we assume that activity durations must lie within certain bounds, i.e., \( \tau_i \leq t_i \leq \tau_i \) such that \( \tau_i > 0, \tau_i > 0, \forall i \). When \( \tau_i = 0 \), we are effectively allowing for activity \( i \) not to be included in the encounter. Moreover, we consider a situation in which the total duration of the service encounter is fixed to \( T \), i.e., \( \sum_i t_i = T \), as with say a two-hour long concert. Relaxing this constraint can be easily accommodated, and a cost of time could also be allocated to the objective. However, in practice, different prices may be associated with different encounter durations (e.g., spa treatments). Hence, changing the total duration may have additional impact on customer satisfaction, in contrast to duration of specific activities, within a particular encounter of fixed duration (e.g., massage of different parts of the body), which are typically not specified contractually.

In addition to having some control over the duration of activities, a service provider may
Table 4.1: Three stylized models of a service encounter

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Fixed</th>
<th>Variable</th>
<th>Fixed</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Duration</td>
<td>VSFD</td>
<td>VSVD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable Duration</td>
<td>FSVD</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

also have control over the sequence of activities. The admissible sequence of activities can be restricted by precedence constraints; for instance, taking X-rays must be performed before a medical procedure. Let $\mathcal{P}$ be the set of all possible feasible sequences of activities.

Then, using (3.4), the service provider’s design problem (SPDP) is given by

$$\max_{(i_1,..,i_n),t} S((i_1,..,i_n),t) = \sum_{k=1}^{n} (x_{i_k} - x_{i_{k-1}})\Phi(\alpha,w,T_k)$$

subject to

$$t_i \leq t_i \leq T, \forall i$$

(4.2)

$$\sum_{i=1}^{n} t_i = T$$

(4.3)

$$(i_1,..,i_n) \in \mathcal{P}.$$ (4.4)

In the sequel, we will denote the optimal sequence by $(i_1^*,..,i_n^*)$ and the optimal duration by $t^*$. In Appendix B, we propose a mathematical programming formulation of SPDP.

The function $\Phi(\alpha,w,T_k)$ is pseudo-concave and peaks when $T_k = \frac{1}{m(\alpha,w)}$; see Lemma A1 in appendix. Accordingly if the service encounter consists of only one activity, i.e., $n = 1$, such that $S = (x_1 - b(0))\Phi(\alpha,w,t_1)$, and $x_1 > b(0)$, satisfaction peaks when the duration is $\frac{1}{m(\alpha,w)}$. Intuitively, the duration must be long enough to have the opportunity to enjoy that high service level, but short enough to avoid acclimating to it.

Because the weighting function (3.5) is symmetric in $\alpha$ and $w$, i.e., $\Phi(\alpha,w,T_k) = \Phi(w,\alpha,T_k)$ a design that is optimal with pure acclimation $\alpha$ and no memory decay is also
optimal with no acclimation and memory decay \( w = \alpha \).

More generally, one can take advantage of that mathematical symmetry to reduce the problem dimensionality. Instead of characterizing a problem instance by a pair of acclimation and memory decay parameters \((\alpha, w)\), one can characterize it by a single number, namely the logarithmic mean between the acclimation rate and the memory decay rate [Carlson, 1972]:

\[
m(\alpha, w) = \frac{w - \alpha}{\ln w - \ln \alpha}.
\]

Accordingly, all our design results will be expressed in terms of \( m(\alpha, w) \).

![Figure 4.1: Logarithmic mean \( m(\alpha, w) \).](image)

The mean \( m(\alpha, w) \) is naturally increasing in both \( \alpha \) and \( w \) such that, for any \( \alpha > 0 \), \( \lim_{w \to \infty} m(\alpha, w) = \infty \) and \( \lim_{w \to 0} m(\alpha, w) = 0 \) and for any \( w > 0 \), \( \lim_{\alpha \to \infty} m(\alpha, w) = \infty \) and \( \lim_{\alpha \to 0} m(\alpha, w) = 0 \). Moreover, \( m(\alpha, w) \leq \min\{\alpha, w\} \), see Figure (4.1). When there is no ambiguity, we will use \( \Phi(t) \) instead of \( \Phi(\alpha, w, t) \) henceforth.

Using these two properties, we next examine the optimal design strategy to maximize ex-post satisfaction. Different service industries may face different degrees of flexibility regarding the sequencing and the allocation of duration to activities. In certain industries (e.g., music performances), the sequence of activities is variable, but the duration of activities is fixed (VSFD). In SPDP, this can be modeled by setting \( \tau_i = \tau_i \forall i \). In other industries (e.g., spa treatment), the sequence of activities is fixed, but the duration is variable (FSVD).
In SPDP, this can be modeled by having only one element in \( P \). Finally, other industries (e.g., fitness classes) are characterized by both variable sequences and variable durations (VSVD). Overall, this leads to three different design problems, see Table 4.1. To obtain a general characterization, we will assume that, whenever the sequence is variable, there are no precedence constraints, i.e., \( P \) is the set of all permutations.

### 4.2 Analysis

In this section, we characterize the optimal sequence and duration allocation solutions, sequentially considering variable sequence with fixed duration (VSFD), variable duration with fixed sequence (FSVD), and variable sequence and variable duration (VSVD) settings. Since the resulting problems are in general hard to solve, despite the analytical characterization of the structure of the optimal solution, we propose design heuristics in chapter five. Underlying the optimal design in each scenario are the following two properties of memory decay and acclimation:

- Memory decay favors a high service level at the end of the encounter.
- Acclimation favors large positive gradients for successive service stages.

Together, they maximize the gradient of the service level near the end of the service encounter. This will lead to the following optimal designs:

- With only acclimation or only memory decay (or more generally, when \( m(a,w) \) is small),
  - sequence activities in a crescendo (increasing service level sequence) and
  - allocate maximum duration to the activities with the highest service intensity

- With both high degrees of acclimation and memory decay (i.e., when \( m(a,w) \) is large),
  - sequence activities in a U-shape and
  - allocate maximum duration to the activities with the lowest service intensity.
4.2.1 Variable Sequence and Fixed Duration

VSFD comprises all service encounters that allow the service provider to change the sequence of activities keeping their respective durations fixed. For instance, consider a session at an INFORMS conference with multiple presentations each by a different speaker. In this scenario, the service level could be the rating of the presentation skills of each speaker.

In terms of complexity, VSFD can be formulated as a single machine job scheduling problem with the objective of maximizing the total weighted profit \( \sum_{k} x_{ik} f(T_{ik}) \), where \( f(T_{ik}) = \Phi(T - T_{ik} + t_{ik}) - \Phi(T - T_{ik}) \). Because \( f(T_{ik}) \) is pseudoconvex in \( T_{ik} \) (see Lemma A-2 in Appendix A), the computational complexity of VSFD, even without precedence constraints, is an open problem [Höhn and Jacobs, 2012].

We find that, individually, both memory decay and acclimation favor designs with increasing service levels. However, when they act together, the optimal sequence is in general U-shaped, as shown in Proposition 1.

**Proposition 1** When durations are fixed, there exists an optimal sequence \( (i_1^*, ..., i_n^*) \) such that for any two consecutive activities \( x_{i_k^*} \) and \( x_{i_{k+1}^*} \) the following applies:

(i) If both activities start and finish within \( [0, T - \frac{2}{m(\alpha,w)}] \), then \( x_{i_k^*} > x_{i_{k+1}^*} \).

(ii) If both activities start and finish within \( [T - \frac{2}{m(\alpha,w)}, T] \) or if \( k = n - 1 \), then \( x_{i_k^*} < x_{i_{k+1}^*} \).

Proposition 1 is illustrated in Figure 4.2 for different values of \( m(\alpha,w) \). For small \( m(\alpha,w) \) the optimal sequence is increasing over time. Otherwise, the optimal sequence is U-shaped with a steep rise at the end. Unlike the optimal allocation of a good’s consumption under either habit formation [Wathieu, 1997] or satiation [Baucells and Sarin, 2010] a decreasing sequence of service levels is however never optimal. This is because memory decay always favors a positive ending of utilities; and to ensure positive instantaneous utility, the last two service levels should be in increasing order. By contrast, consumption planning models typically discount the future and therefore tend to favor immediate consumption.

When either \( \alpha \) or \( w \) is small (i.e., when \( m(\alpha,w) \leq \frac{2}{T} \)), it is optimal to have activities sequenced in increasing order of service levels, i.e., a crescendo pattern is optimal. When \( w \) is small, the customer gives equal weight to past and recent utilities. Because of acclimation,
any drop in service level may create a negative instantaneous utility. Hence with small \( w \), it is optimal to have a monotonically increasing sequence of service levels. By contrast when \( \alpha \) is small, the customer acclimates to the service level of an activity very slowly. Therefore, the higher the service level, the larger the magnitude of instantaneous utility. Because of memory decay, these high utilities are more likely to be remembered if they happen at the end of the encounter. Hence, with small \( \alpha \), it is also optimal to have a monotonically increasing sequence of service levels. To summarize, when either \( \alpha \) or \( w \) is small, the service provider should sequence activities in increasing order of service levels, as shown in Corollary 1.

**Corollary 1** In VSFD, for \( \alpha = 0 \) or \( w = 0 \), the optimal sequence of activities \((i_1^*, ..., i_n^*)\) is such that \( x_{i_1^*} < ... < x_{i_n^*} \), i.e., is in increasing order of service levels.

By contrast when \( \frac{2}{T} < m(\alpha, w) \), Proposition 1 shows that the optimal sequence of activities is U-shaped: The service level should fall gradually in the beginning and rise steeply at the end. This U-shaped encounter maximizes the gradient of service level towards the end of the encounter and is desirable when \( m(\alpha, w) \) is large, i.e., when both \( \alpha \) and \( w \) are high. Intuitively, high memory decay \( w \) makes the customer heavily discount past experiences. On the other hand, high acclimation \( \alpha \) makes the current utility depend less on past service levels. Therefore, the service provider can gradually decrease the service level in the early
part of the sequence, which will be forgotten due to high memory decay, so as to ensure a steep positive gradient towards the end.

U-shaped sequences of activities appear in many experiences, such as concerts [Baucells et al., 2013], concertos [Rozin et al., 2004], opera season scheduling [Dixon and Thompson, 2013b], movies and theme parks [Lawrence, 2006]. U-shaped sequences turn out to be optimal in consumption planning in the presence of either acclimation [Wathieu, 1997] or satiation [Baucells and Sarin, 2010]. We thus reach similar conclusions (albeit slightly different since decreasing sequence is never optimal) from a different perspective, by making evaluation retrospective (and subject to memory decay) instead of prospective (and subject to the usual time discounting). In retrospective evaluation, U-shaped sequences of activities could emerge for other reasons. For instance, a combination of recency (due to memory decay) and primacy (due to attention decrement) could lead to similar sequence. We thus offer a complementary, yet more subtle, justification of the optimality of U-shaped sequence. Moreover, what happens within the U-shape may differ. In particular, maximizing the weighted sum of peak intensities, final intensities, spread between peak and final intensities, and slope of the trend leads to a steep fall and gradual rise, whereas the combination of acclimation and memory decay suggests the opposite. Incidentally, it may not always be optimal to place the activity with the highest service level at the end, contrary to what may imply the peak-end rule [Kahneman et al., 1993]. When the activity with the highest service level has a long duration, the customer becomes acclimated to that high level of service and her utility diminishes over time. In that case, it may then be optimal to move that activity at the beginning of the encounter so as to ensure a steep gradient near the end. For instance, when \((x_1, x_2, x_3, x_4) = (2, 5, 7, 10)\) and \((t_1, t_2, t_3, t_4) = (5, 4, 3, 8)\), the optimal sequence when \(w = 1\) and \(\alpha = 0.7\) is \((i_1, i_2, i_3, i_4) = (4, 2, 1, 3)\), i.e., the activity with the highest service level is placed at the beginning of the encounter.

Hence, we find that, although it is always optimal to finish on a high note, it may not be optimal to finish on the highest note; one should therefore take the “finish strong” recommendation in context. In addition, although it is often recommended to “get the bad experience out of the way early”, this may not be optimal in the presence of both acclimation
and memory decay. Finally, while it is recommended to “segment the pleasure, combine the pain”, we observe that it may be optimal to combine the pleasure if either memory decay or acclimation is insignificant.

Although proposition 1 characterizes the optimal sequence, it offers little indication on how to construct the optimal U-shape. This problem turns out to be exponentially complex and commercial solvers may fail to return a solution when the number of activities becomes large. In Appendix C.1, we propose a heuristic to construct a U-shape sequence, so as to ensure a steep gradient near the end. Numerical experiments suggest that the heuristic’s suboptimality gap remains moderate even for a large number of activities.

4.2.2 Fixed Sequence and Variable Duration

FSVD comprises all service encounters that have a fixed sequence of activities but offer the flexibility to alter their duration. For instance, a medical procedure performed by the dentist will have a more or less fixed sequence of activities but the time spent on each activity can be varied at the discretion of the dentist. In terms of complexity in general, FSVD is only a well-behaved concave optimization problem when the sequence of activities is increasing in service levels, i.e., \( x_{i_1} < \cdots < x_{i_n} \), (see Lemma A-4 in Appendix A); otherwise it may have multiple local optima. Yet, we show in Appendix C.2 that a simple coordinate ascent algorithm performs reasonably well. The following proposition characterizes the optimal duration allocation across a subsequence of activities.

**Proposition 2** For a fixed subsequence of activities \((x_l, \ldots, x_r)\) where \(l \geq 1\), the optimal duration allocation is given as follows:

1. Suppose that \(m(\alpha, w) < \frac{1}{T} \), and that there exists an activity \(q, l < q < r\), such that \(t^*_q > \tau_q\). Then if \(x_l < \cdots < x_r\), then \(t^*_j = \tau_j, j = q + 1, \ldots, r\); and if \(x_l > \cdots > x_r\), \(t^*_j = \tau_j, j = l, \ldots, q - 1\).

2. Suppose that \(\frac{1}{T} < m(\alpha, w) < \frac{1}{\sum_{i=r}^{l} \tau_i}\). Then if \(x_l < \cdots < x_r\) such that \(t^*_i = \tau_i, t^*_r = \tau_r\), then \(t^*_i = \tau_i, \forall i, l < i < r\); and if \(x_l > \cdots > x_r\) such that \(t^*_l > \tau_l, t^*_r > \tau_r\), then \(t^*_i > \tau_i, \forall i, l < i < r\).

3. Suppose that \(\frac{1}{\sum_{i=r}^{l} \tau_i} < m(\alpha, w) \) and that there exists an activity \(q, l < q < r\), such that...
Then if \( x_l < \cdots < x_r \), then \( t^*_j = \tau_j, j = l, \ldots, q - 1 \); and if \( x_l > \cdots > x_r \), \( t^*_j = \tau_j, j = q + 1, \ldots, r \).

Figure 4.3: Optimal duration allocation for increasing subsequences. Shaded activities are allocated duration above their lower bound.

Figure 4.4: Optimal duration allocation for decreasing subsequences. Shaded activities are allocated duration above their lower bound.

We illustrate Proposition 2 by considering two service encounters: (i) with all increasing service levels (see Figure 4.3) and (ii) with all decreasing service levels (see Figure 4.4). In the figures, the shaded activities are allocated duration above their lower bound, i.e., \( t^*_i > \tau_i \).

Proposition 2 shows that when either \( \alpha \) or \( w \) is low (i.e., \( m(\alpha, w) < \frac{1}{T} \)), then duration should be allocated to the activities in decreasing order of service levels (see (i) in Figures 4.3 and 4.4). When the customer gives equal weight to all instantaneous utilities (\( w \approx 0 \)), this design strategy naturally ensures that the maximum time is spent at high service levels.
When there is no acclimation (i.e., $\alpha \approx 0$), this design strategy brings closer to the end of the service encounter the activities associated with a high service level, so that the customer does not heavily discount them.

By contrast when both $\alpha$ and $w$ are very high (i.e., when $m(\alpha, w) > \frac{1}{T_n}$), then duration should be allocated in increasing order of service levels so as to push the steep rise and gradual fall in service level towards the end of the encounter (see (iii) in Figures 4.3 and 4.4). Large memory decay implies that the customer forgets the initial experience of the encounter. Therefore with an increasing sequence, it is optimal to allocate duration to the activities in the beginning so as to bring closer to the end the steep rise in service levels. Conversely in a decreasing sequence, it is optimal to allocate duration at the end so as to push the steep drop in service levels as far as possible from the end of the sequence and make the gradient of service level near the end tend to zero.

In the intermediate cases when $\frac{1}{T} < m(\alpha, w) < \frac{1}{T_n}$, duration can be allocated in any of the three possible ways shown in Figures 4.3 and 4.4, respectively for an increasing and a decreasing sequence. However, duration should never be allocated only to the activities in the middle of an increasing sequence or in the extremities of a decreasing sequence.

Figure 4.3 (iii) indicates that, in the presence of both acclimation and memory decay, the optimal design should induce a steep gradient at the end of the encounter, even if that involves spending less time on high-service level activities. A grand finale is desirable, but it needs to be short and contrasting with the preceding activities. This thus corroborates the optimality of the U-shaped sequence in VSFD, which induced some respite before the final ascent in service levels.

Similarly, Figure 4.4 (iii) indicates that, in the presence of both acclimation and memory decay, it is preferable to spend more time at low service levels than to provide great service initially and then create disappointment near the end of the encounter. In fact, if $\tau_i = 0, \forall i$, it may be optimal to allocate zero duration to the initial activities, associated with the highest service level, effectively removing them from the encounter.

Overall, Figures 4.3 and 4.4 suggest that the optimal design should strive for stability at low levels of service and sharp transition to and from (if it must happen) high levels of


4.2.3 Variable Sequence and Variable Duration

Some service encounters, like personal fitness classes or museum tours across multiple galleries, allow the provider to change both the sequence and the duration of activities. In terms of complexity, VSVD is at least as difficult as FSVD or VSFD.

Combining Propositions 1 and 2 leads to the following characterization of the optimal sequence and duration allocation.

**Corollary 2** In VSVD,

1) If \( m(\alpha, w) < \frac{2}{T} \), a sequence with increasing service levels is optimal. Moreover,
   
   1.1) If \( m(\alpha, w) < \frac{1}{T} \), the optimal duration allocation is characterized by an index \( k \) such that \( t_{ij} = \tau_{ij}, 1 \leq j \leq k - 1, \ t_{ik} = \tau, \tau_{ik} \leq \tau \leq \tau_{ik}, \ \text{and} \ \ t_{ij} = \tau_{ij}, k < j \leq n \).

   1.2) If \( \frac{1}{T} < m(\alpha, w) \leq \frac{1}{\Sigma_{in}} \), the optimal duration allocation is such that, if \( t_{il}^* = \tau_{il} \) and \( t_{ir}^* = \tau_{ir} \), then \( t_{ij}^* = \tau_{ij}, \ \forall j, l < j < r \).

   1.3) If \( \frac{1}{\Sigma_{in}} < m(\alpha, w) \), then the optimal duration allocation is characterized by an index \( k \) such that \( t_{ij}^* = \tau_{ij}, j = 1, \cdots, k - 1, \ t_{ik}^* = \tau, \tau_{ik} \leq \tau \leq \tau_{ik}, \ \text{and} \ t_{ij}^* = \tau_{ij}, j = k + 1, \cdots, n \).

2) If \( \frac{2}{T} < m(\alpha, w) \), a U-shaped sequence bottoming out at activity \( k \) is optimal such that:
   
   for \( x_{ik} < \cdots < x_{in} \), if \( \exists l, r \) such that \( t_{il}^* = \tau_{il} \) and \( t_{ir}^* = \tau_{ir} \), then \( t_{ij}^* = \tau_{ij}, \ \forall j, k \leq l < j < r \leq n \)

   for \( x_{i1} > \cdots > x_{ik} \), if \( \exists l, r \) such that \( t_{il}^* > \tau_{il} \) and \( t_{ir}^* > \tau_{ir} \), then \( t_{ij}^* > \tau_{ij}, \ \forall j, 1 \leq l < j < r \leq k \).

Overall, Corollary 2 shows that the optimal solution for VSVD is obtained by constructing the steepest positive gradient near the end. Figure 4.5 represents the change in service design as \( m(\alpha, w) \) increases, i.e., as either acclimation or memory decay or both become more intense. Initially, when \( m(\alpha, w) \) is low, crescendo is the optimal sequence, with duration initially allocated to the activities at the end of the sequence, and then, as \( m(\alpha, w) \) increases, also at the beginning of the sequence, so as to induce a steeper gradient near the end of the encounter. As \( m(\alpha, w) \) keeps increasing, the optimal sequence may switch to a U-shape
Figure 4.5: Optimal sequence and duration allocation as $m(\alpha, w)$ increases. Shaded activities are allocated duration above their lower bound.

sequence, as in (iv) and (v), and more duration is then be allocated to the activities in the middle of the encounter, so as to induce a more gradual fall at low levels of service and a steeper rise at higher service levels.
5.1 Introduction

In chapter four we discussed the computational complexity of three types of service encounters, namely, variable sequence and fixed duration (VSFD), fixed sequence and variable duration (FSVD), and variable sequence and variable duration (VSVD). All the three problem formulations are computationally complex and in general cannot be easily solved to optimality. In this chapter we propose heuristics that can be implemented to provide near optimal designs of a service encounter in terms of the sequence of activities and the duration allocated to them.

The heuristics for VSFD and FSVD are compared against the true optimal solution for problem instances that can be solved to optimality. For VSFD, we use enumeration to find the optimal solution, whereas for FSVD we use problem instances that can be solved to near optimality by commercial solvers. For benchmarking the performance of the heuristic corresponding to VSVD, we formulate an upper bound that can be computed efficiently.

5.2 Variable Sequence Fixed Duration

Because commercial solvers failed to return a feasible solution within the set time limit for \( n \geq 9 \),\(^1\) we propose a heuristic based on Proposition 1 that states that the optimal sequence should be U-shaped with activities appearing within the duration \([0, T - \frac{1}{m(\alpha,w)}]\) sequenced in decreasing order of service levels whereas activities appearing within the duration \([T - \frac{1}{m(\alpha,w)}
\]

\(^1\)We solved VSFD with BARON and LINDOGlobal available on NEOS with a maximum run time of 500,000 seconds for \( n = 6, \ldots, 10 \).
$2^{-1/m(α,w)}, T]$ are sequenced in increasing order of service levels.

The heuristic assigns activities sequentially, going backwards in time. It selects activities in decreasing order of service levels, starting from the activity with the highest service level, until time $T - 2^{-1/m(α,w)}$ is reached. At that point, it sequences the remaining activities in increasing order of service level, starting from the activity with the lowest service level. Doing so results in a U-shaped sequence, or in an only increasing sequence if $T - 2^{-1/m(α,w)} < 0$. The formal steps appear in Algorithm 1. VSFD can be formulated as a mixed integer non-linear programming problem (see Appendix B).

We tested the performance of the heuristic on small problem instances that can be solved optimally by exhaustive enumeration. We generated the problem instances with $α$ and $w$ uniformly drawn between 0 and 1 such that the condition $T > 2^{-1/m(α,w)}$ is satisfied since an increasing sequence is known to be optimal otherwise. The values of $t_i$’s were uniformly drawn from the interval $[1, 100]$, and the service levels ($x_i$’s) were drawn from a Uniform [1,100] or from a Gamma distribution with shape ($k$) and scale ($θ$) parameters, with either $k = 1$ and $θ = 2$ or $k = 9$ and $θ = 0.5$, respectively generating an exponential distribution and a bell-shaped distribution. Table 5.1 reports the mean, median, standard deviation, minimum, and maximum suboptimality gaps of our heuristic relative to the optimal solution found through complete enumeration, over 100 randomly generated instances for each value of $n$.

---

**Algorithm 1 Heuristic for VSFD**

1: Sort activities such that $x_1 > .. > x_n$.
2: Comment: opt is the optimal sequence.
3: Initialize: time= 0, $i = n, j = 1$, $k = n$.
4: while time $\leq 2^{-1/m(α,w)}$ do
5: opt[$i$]=x[$j$]
6: $i=i-1$
7: time=time+$t[j]$
8: $j=j+1$
9: end while
10: while $i > 0$ do
11: opt[$i$]=x[$k$]
12: $i=i-1; k=k-1$
13: end while
Table 5.1: Optimality Gap of VSFD Heuristic

<table>
<thead>
<tr>
<th></th>
<th>$x_i \sim Uniform[1, 100]$</th>
<th></th>
<th>$x_i \sim Gamma(1, 2)$</th>
<th></th>
<th>$x_i \sim Gamma(9, 0.5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Std Dev</td>
<td>Minimum</td>
<td>Maximum</td>
</tr>
<tr>
<td>6</td>
<td>0.77%</td>
<td>0.44%</td>
<td>0.95%</td>
<td>0.00%</td>
<td>4.66%</td>
</tr>
<tr>
<td>7</td>
<td>1.17%</td>
<td>0.86%</td>
<td>1.03%</td>
<td>0.00%</td>
<td>4.29%</td>
</tr>
<tr>
<td>8</td>
<td>1.43%</td>
<td>0.93%</td>
<td>1.96%</td>
<td>0.00%</td>
<td>5.68%</td>
</tr>
<tr>
<td>9</td>
<td>1.88%</td>
<td>1.44%</td>
<td>1.69%</td>
<td>0.00%</td>
<td>7.59%</td>
</tr>
<tr>
<td>10</td>
<td>1.92%</td>
<td>1.39%</td>
<td>1.75%</td>
<td>0.00%</td>
<td>7.68%</td>
</tr>
</tbody>
</table>

Overall, Table 5.1 shows that the average suboptimality gap for this heuristic is within 2%, and the maximum gap is within 8%. The suboptimality gaps naturally increase with $n$, but the deterioration in performance remains moderate relative to the exponential increase in run time of the enumeration method. Moreover, the gaps appear to be robust to the distribution of service levels.

5.3 Fixed Sequence Variable Duration

The heuristic for FSVD is based on the fact that given two activities $i$ and $j$, if activity $i$ offers higher satisfaction margin than activity $j$, i.e., $\frac{\partial S(t)}{\partial t_i} > \frac{\partial S(t)}{\partial t_j}, \forall t_i \in [\tau_i, \tau_i + \Delta t], \forall t_j \in [\tau_j, \tau_j + \Delta t]$ then allocating duration $\Delta t$ to activity $i$ will increase the level of satisfaction more than allocating it to activity $j$.

We solved FSVD with LINDOGlobal. The commercial solver returned the optimal or a near optimal solution (with a 1% optimality gap) for small problem instances, but it failed to solve larger problem instances (e.g., $n \geq 24$) with decreasing subsequences of service levels within one week. Since the objective function of FSVD with a decreasing sequence is pseudoconvex (Lemma A-4), the maximization problem may have multiple local optima when $\tau_{i_{n}} < \frac{1}{m(\alpha, w)} < T$, and there is no guarantee that a solution can be obtained in a reasonable amount of time.

We therefore propose a modified coordinate ascent heuristic to solve FSVD that takes less computational time on average than LINDOGlobal. The heuristic starts with $t^0 = t$ and
iteratively allocates time increments $\Delta t$ to the activity with the highest marginal return, i.e.,

$$\arg \max_i \frac{\partial S(t)}{\partial t_i}$$

until the total duration is equal to $T$. The steps are detailed in Algorithm 2.

**Algorithm 2** Heuristic for FSVD

1: Comment: $\Delta t$ is the step size. $I$ is the set of activities such that $t_i = \tau_i$, $\forall i \in I$.
2: Initialize: $t^0 = \tau$; $m = 0$; $I = \emptyset$
3: while $T - \sum_{i=1}^{n} t^m_i > 0$ do
4: $i = \arg \max_k k \in I \frac{\partial S(t^m)}{\partial t_k}$
5: $\Delta t_i = \min\{\Delta t, T - \sum_{j=1}^{n} t^m_j, \tau_i - t^m_i\}$
6: if $\Delta t_i = 0$ then
7: $I = I \cup \{i\}$
8: end if
9: $t^{m+1}_i = t^m_i + \Delta t_i$
10: $m = m + 1$
11: end while

To compare the heuristic solution against the solution obtained from the solver, we considered four patterns of sequences of service levels, namely, increasing, decreasing, U-shaped, and Λ-shaped. For each pattern of sequence, we considered sequences of different lengths, i.e., $n = 8, 16, 24, 32$. We randomly generated 100 instances for each $n$. All instances were generated with $\alpha$ and $w$ uniformly drawn from $(0, 1)$ and with $x_i, \tau_i, \tau_i, \forall i$ uniformly drawn from $[1, 100]$. The total duration of the service encounter was fixed to $T = \sum_i \tau_i + \sum_i (\tau_i - \tau_i)/n$. For each $n$, the U-shape sequence was created by forming a decreasing subsequence with $n/2$ randomly selected activities and an increasing subsequence with the remaining $n/2$ activities. The Λ-shape was constructed similarly. The time increment for the heuristic was fixed as $\Delta t = 0.01$.

Table 5.2 benchmarks the solution as well as the running times from the heuristic to the best solution found by LINDOGlobal, within a 1% optimality gap. All codes were implemented in MATLAB 7.12.0. and GAMS. The LINDOGlobal solver was executed using the NEOS optimization server. The computational tests on MATLAB were run on a workstation with a 2.26 GHz Intel Core 2 Duo processor, 8 GB of RAM, and Mac OS X 10.6.8 as the operating system.

Apart from the sequences of activities with increasing service levels, which yield a concave
Table 5.2: Heuristic Solution vs LINDOGlobal Solution for FSVD

<table>
<thead>
<tr>
<th></th>
<th>Mean Run Time</th>
<th>No. of Aborted Instances by Solver</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Heuristic</td>
<td>Solver</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>8</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>16</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>24</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>32</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>16</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>24</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>32</td>
<td>0.11%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

maximization problem (Lemma A-4), all other sequences had some cases aborted by NEOS when the solver failed to find an optimal or a near-optimal solution with a 1% optimality gap in 500,000 seconds (reported in the last column in Table 5.2). Only the cases for which the solver returned an optimal or near-optimal solution were included when computing the performance gaps.

Overall, the FSVD heuristic gives solutions that closely match the solver solutions (within 1.5%) in a significantly smaller run time (i.e., in less than 1 second). Moreover, the computation time of the coordinate ascent heuristic appears less sensitive to the size or specific shape of the sequence.
5.4 Variable Sequence Variable Duration

VSVD is at least as complex as VSFD since an instance of VSFD can be constructed from VSVD by setting $\tau_i = \tau_i$, $\forall i$. When $T < 2\frac{1}{m(a, w)}$, the sequence of service levels is increasing, and therefore the duration allocation problem is pseudoconcave (see Lemma A-4 in Appendix A). When $T > 2\frac{1}{m(a, w)}$, Corollary 2 shows that a U-shaped sequence is optimal and the duration allocation problem may therefore have multiple local optima. In general, commercial solvers like BARON and LINDOGlobal on NEOS failed to return a feasible solution in one week for many instances with $n \geq 10$. We therefore propose a heuristic for SPDP when $T > 2\frac{1}{m(a, w)}$ and we present an upper bound on SPDP against which we benchmark the performance of our heuristic.

5.4.1 Upper Bound

We introduce an alternate formulation of SPDP that allows the sequence to be either U-shaped or monotonically increasing in service levels. Without loss of generality, we assume that $x_1 \geq \cdots \geq x_n$. A U-shaped sequence consists of a decreasing subsequence followed by an increasing one. Let $y_i^s \in \{0, 1\}, \forall i, \forall s$ be a variable indicating if activity $i$ belongs to subsequence $s$, where $s = l$ for the decreasing subsequence on the left and $s = r$ for the increasing subsequence on the right of the U-shape. The variables $t_i^s, \forall i, \forall s$ denote the duration allocated to activity $i$ when it is in subsequence $s$. With these decision variables, SPDP-U is formulated as:

$$
\max_{y^l, y^r, t^l, t^r} S(y^l, y^r, t^l, t^r) = \sum_{i=1}^{n} (x_i - x_{i-1})\Phi(T - \sum_{j=1}^{i-1} t_j^iy_j^i) + \sum_{i=1}^{n-1} (x_i - x_{i+1})\Phi(\sum_{j=1}^{i} t_j^iy_j^i)
$$

subject to

$$
y_i^l + y_i^r = 1, \forall i \tag{5.1}
$$

$$
1 \leq \sum_{i=1}^{n-1} y_i^r \tag{5.2}
$$

$$
y_i^s \tau_i \leq t_i^s \leq y_i^s \tau_i, s = l, r, \forall i \tag{5.3}
$$

$$
\sum_{i=1}^{n} (y_i^l t_i^l + y_i^r t_i^r) = T \tag{5.4}
$$
where \( x_0 = b(0) \). The objective function consists of a decreasing and an increasing subsequences, such that each subsequence is composed of all the activities. When \( t_i^l = 0 \), the coefficient of \( x_i \) in the decreasing subsequence is equal to zero, i.e., no weight is associated with activity \( i \) in the decreasing subsequence. Similarly if \( t_i^r = 0 \), no weight is associated with activity \( i \) in the increasing subsequence. Constraints (5.1) allow an activity to be either a part of the decreasing or a part of the increasing subsequence. Constraint (5.2) eliminates the possibility of having only a decreasing sequence of activities. Therefore, constraints (5.1)-(5.2) ensure a U-shaped or a monotonically increasing sequence. The remaining constraints (5.3)-(5.4) correspond to the duration constraints. In particular, constraints (5.3) ensure that if \( y_i^s = 0 \), then the corresponding duration \( t_i^s = 0 \). This rules out the possibility of allocating duration to an activity when it does not belong to the corresponding subsequence. In Appendix C, Lemma C-1, we show that SPDP-U is equivalent to SPDP when \( T > \frac{2}{m(a,w)} \).

The upper bound is obtained by dropping constraints (5.1)-(5.2), relaxing constraint (5.3) by replacing the lower bounds \( \tau_i \) by zero \( \forall i \), and relaxing (5.4) by converting the equality sign to an inequality sign, i.e., \( \sum_{i=1}^{n} (y_i^l t_i^l + y_i^r t_i^r) \leq T \). We furthermore relax constraint (5.5) by allowing \( y_i \) to be a continuous variable between 0 and 1, and then introduce a new set of variables, namely, \( \tilde{t}_i^s = t_i^s y_i^s, s = l, r \). After dualizing the relaxed constraint (5.4), we obtain the upper bound \( UB = \min_{\lambda \geq 0} UL(\lambda) + UR(\lambda) + \lambda T \), in which

\[
UL(\lambda) = \max_{\mathbf{t}} UL(\lambda, \mathbf{t}) = \sum_{i=1}^{n} (x_i - x_{i-1}) \Phi(T - \sum_{j=1}^{i-1} t_j) - \lambda \sum_{i=1}^{n} t_i
\]

subject to \( 0 \leq t_i \leq \tau_i, \forall i \) (5.6)

\[
UR(\lambda) = \max_{\mathbf{t}} UR(\lambda, \mathbf{t}) = \sum_{i=1}^{n-1} (x_i - x_{i+1}) \Phi(\sum_{j=1}^{i} t_j) - \lambda \sum_{i=1}^{n} t_i
\]

subject to \( 0 \leq t_i \leq \tau_i, \forall i \) (5.7)

in which we renamed the duration variables \( \tilde{t}_i^l = t_i, \forall i \) in \( UL(\lambda) \) and \( \tilde{t}_i^r = t_i, \forall i \) in \( UR(\lambda) \), for simplicity of exposition. The Lagrangian relaxation \( UB = \min_{\lambda \geq 0} UL(\lambda) + UR(\lambda) + \lambda T \) is then solved by the subgradient method [Fisher, 2004].
Next we show that \( UR(\lambda) \) and \( UL(\lambda) \) can be solved in polynomial time. Without loss of
generality, we assume that the service levels of all activities are unique, i.e., \( x_1 > \cdots > x_n \).
If we have \( x_i = x_{i-1} \), for some \( i \), we can delete activity \( i-1 \) and increase the upper bound
on \( t_i \) to \( \tau_i + \tau_{i-1} \) since

\[
UL(\lambda, t) = \sum_{k=1}^{i-1} (x_k - x_{k-1}) \Phi(T - \sum_{j=1}^{k-1} t_j) + (x_i - x_{i-1}) \Phi(T - \sum_{j=1}^{i-1} t_j) + \sum_{k=i+1}^{n} (x_k - x_{k-1}) \Phi(T - \sum_{j=1}^{k-1} t_j) - \lambda \sum_{i=1}^{n} t_i
\]

and similarly, for \( UR(\lambda, t) \).

Using the above result we show in Lemma C-2 and Lemma C-3 that it takes \( O(n^2) \) and
\( O(n) \) points to be considered for finding the solution to \( UL(\lambda) \) and \( UR(\lambda) \) respectively. The
proofs are available in Appendix C.

### 5.4.2 Heuristic and Computational Study

The heuristic first finds a U-shaped sequence and then allocates duration for the corre-
sponding FSVD problem so as to maximize the gradient of service levels at the end of the
encounter. We next detail the two steps of the heuristic.

Step 1: **Sequencing.** The heuristic constructs a U-shaped sequence. Without loss of gener-
ality, we assume that \( x_1 \geq \cdots \geq x_n \). The following optimization problem freely allocates
duration among all activities, so as to maximize customer satisfaction by generating a steep
gradient in service level.

\[
\max_{t} S(t) = \sum_{i=1}^{n-1} (x_i - x_{i+1}) \Phi(\sum_{j=1}^{i} t_j)
\]

subject to \( 0 \leq t_i \leq \tau_i, \forall i \)

\[ \sum_{i=1}^{n} t_i \leq T. \]
Table 5.3: Optimality Gaps for the VSVD heuristic with $T > \frac{1}{m(\alpha,w)}$

<table>
<thead>
<tr>
<th>n</th>
<th>Solver</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>5</td>
<td>0.72%</td>
<td>0.75%</td>
</tr>
<tr>
<td>6</td>
<td>1.09%</td>
<td>0.71%</td>
</tr>
<tr>
<td>7</td>
<td>1.42%</td>
<td>1.30%</td>
</tr>
<tr>
<td>8</td>
<td>2.19%</td>
<td>2.01%</td>
</tr>
<tr>
<td>9</td>
<td>2.67%</td>
<td>2.32%</td>
</tr>
</tbody>
</table>

Compared to SPDP, the lower bounds on activity duration (4.2) are set to zero and the constraint on the total duration of the service encounter (4.3) is relaxed to an inequality. Moreover, the sequence is fixed. If an activity is allocated zero duration, it suggests that it does not contribute to creating a steep gradient and should therefore be moved at the beginning of the encounter. Accordingly, let $L$ be the set of activities that are allocated zero duration, i.e., if $t^*$ denotes the optimal solution then $L = \{j | t^*_j = 0\}$. Therefore, for all $j \in L$, we order activities in decreasing order of $x_j$’s and for all $j \notin L$, we order activities in increasing order of $x_j$’s. The combination of the decreasing subsequence followed by the increasing subsequence forms a U-shaped sequence.

Step 2: Duration Allocation. For the U-shaped sequence obtained in Step 1, we apply the coordinate ascent method described in Algorithm 2 to allocate duration.

To test the performance of the heuristic against the optimal solution, when computable, and an upper bound otherwise, we randomly generated problem instances with $\alpha$ and $w$ uniformly drawn from $[0, 1]$, and $x_i, \tau_i, \bar{\tau}_i, \forall i$ uniformly drawn from $[1, 100]$. The total duration of the service encounter was set as $T = \sum_i \tau_i + \sum_i (\tau_i - \bar{\tau}_i)/n$ and the condition $T > \frac{2}{m(\alpha,w)}$ was verified.

Table 5.3 compares the solution obtained from the heuristic with the optimal solution and the upper bound when $n \leq 9$. For every instance of size $n$, the gaps were evaluated over 100 instances, and the corresponding sample was used for evaluating the mean, median, standard deviation, minimum, and maximum optimality gap. We observe that the gaps against the optimal solution are within 3% on an average, whereas against the upper bound the average gaps are within 6%.
Table 5.4: Optimality Gaps for the VSVD heuristic with $T > \frac{2}{m(\alpha, w)}$

<table>
<thead>
<tr>
<th>n</th>
<th>Mean</th>
<th>Median</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5.99%</td>
<td>4.91%</td>
<td>3.18%</td>
<td>0.75%</td>
<td>10.53%</td>
</tr>
<tr>
<td>15</td>
<td>6.91%</td>
<td>5.95%</td>
<td>2.97%</td>
<td>0.83%</td>
<td>11.79%</td>
</tr>
<tr>
<td>20</td>
<td>7.09%</td>
<td>6.81%</td>
<td>3.01%</td>
<td>1.03%</td>
<td>13.86%</td>
</tr>
<tr>
<td>25</td>
<td>8.01%</td>
<td>7.77%</td>
<td>3.86%</td>
<td>1.99%</td>
<td>14.01%</td>
</tr>
<tr>
<td>30</td>
<td>9.81%</td>
<td>8.97%</td>
<td>3.55%</td>
<td>2.08%</td>
<td>14.19%</td>
</tr>
</tbody>
</table>

Table 5.4 compares the performance of the heuristic against the upper bound for larger instances, over a sample of 100 instances for each $n$. The gap of the heuristic is on average less than 10% of the upper bound, although it increases with $n$. Given that the upper bound appears to be less tight for larger values of $n$ from Table 5.3, the performance of the heuristic could be within a few percent of the optimal solution, even for large problem instances.
CHAPTER 6

Model Extensions

6.1 Introduction

In this chapter we look at extensions of the service provider’s design problem (SPDP). These extensions are primarily obtained by relaxing the assumptions that were made in formulating the problem.

The first model extension relaxes the assumption that service levels of the activities in the encounter are exogenously defined. In practice, service providers are often responsible for choosing the service level they would like to offer to their customers [Aflaki and Popescu, 2013]. For instance, during a city bus tour, the tour guide decides what quality of food will be offered to the customer and at which location the bus would stop to allow tourists to visit the attraction as opposed to seeing it from inside the bus. Similarly, in a spa, the service provider decides what should be the quality of the products and the equipments that will be offered to the customer, or an instructor of a fitness class decides how much personal attention she will provide to the customers during various sessions. Hence, we will analyze the SPDP problem with fixed sequence and duration but variable service level.

We also extend the SPDP model to consider the scenario where the service provider offers service to customers who are very diverse in their rates of acclimation and memory decay. So far we have assumed that the customers are homogeneous and have the same rates of acclimation and memory decay. In such scenarios the service provider can offer a common service design that will be optimal for all the customers. In practice, this assumption may fail and the service provider may be serving a diverse population of customers who are heterogeneous in terms of the rates at which they acclimate to an experience or forget past
experiences. We therefore consider the problem of sequencing and duration allocation in the presence of a heterogeneous population of customers.

### 6.2 Endogenous Service Levels

So far while analyzing the SPDP problem we assumed that the service levels $x_i$ for any activity $i \in \{1, \ldots, n\}$ are exogenously defined and the service provider can only make decisions regarding the sequence of events and the duration allocated to them. We also impose budget constraint on the cost of service levels that the service provider offers. For instance, for a five day cruise, the itinerary may be fixed but the cruise line decides the service level comprising on-board entertainment programs for each day. Moreover, there might be a limit on the service level offered each day, and the total budget available to be allocated for service across all days.

The corresponding problem can be formulated in terms of $T_k = \sum_{i=k}^{n} t_i$ as follows.

$$
\max \ S(x) = \sum_{i=1}^{n-1} x_i [\Phi(\alpha, w, T_{i+1}) - \Phi(\alpha, w, T_i)] + x_n \Phi(\alpha, w, T_n) - x_0 \Phi(\alpha, w, T) \tag{6.1}
$$

such that

$$
x_i \leq x_i \leq \bar{x}_i, \forall i \tag{6.2}
$$

$$
\sum_{i=1}^{n} c_i(x_i) \leq B. \tag{6.3}
$$

where constraint (6.2) is the limit on the service levels of individual activities, and inequality (6.3) is the budget constraint on the total cost of service to be offered across all activities. The cost function $c_i(x_i)$ gives the cost incurred in offering service level $x_i$ for activity $i$. The total budget available to the service provider is assumed to be $B$.

We observe that the coefficient of the service level $x_i$ for any arbitrary activity $i$ can be represented as

$$
g(t_i) = \Phi(\alpha, w, \Delta t + t_i) - \Phi(\alpha, w, \Delta t)$$

where $\Delta t = \sum_{k=i+1}^{n} t_k$. In figure 6.1 we see that this function is positive and decreasing when $x_i$ is closer to the end of the encounter, whereas it is negative and increasing when $x_i$
is farther away from the end of the encounter. This indicates that the activities near the end contribute more to increase the satisfaction of the customer and the activities at the beginning of the encounter contribute less towards the overall satisfaction. To simplify the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{g_t.pdf}
\caption{Function $g(t)$ vs. time left to end the encounter ($t$) with $w = 1$ and $\alpha = 0.5$.}
\end{figure}

analysis we further assume that the cost function is linearly increasing in service level, i.e., $c_i(x_i) = a_i x_i, \forall i$ where $a_i \in \mathbb{R}^+$. In this case the following proposition holds true.

**Proposition 3** When service levels are endogenous the following apply:

1. If $\bar{T}_{ik+1} < \frac{1}{m(\alpha,w)} < \bar{T}_{ik}$ then $x^*_i \equiv \underline{x}_i, j = 1, ..., k$ and $x^*_i > \underline{x}_i, j = k + 1, ..., n$.

2. If $\frac{1}{m(\alpha,w)} < \bar{T}_{ik}$, then $x^*_i > \underline{x}_i$ and $x^*_i = \underline{x}_i, k = 1, ..., n - 1$.

The proof of this result is provided in Appendix A. Proposition 3 indicates that the last activity is always allocated a service level greater than the lower limit set at $\underline{x}_i$.

An alternate proof for this result is obtained by re-writing the problem as follows.

\[
\max_y S(y) = \sum_{i=1}^{n-1} (y_i + \underline{x}_i) [\Phi(\alpha, w, \bar{T}_{i+1}) - \Phi(\alpha, w, \bar{T}_i)] + (y_n + \underline{x}_n) \Phi(\alpha, w, \bar{T}_n) - x_0 \Phi(\alpha, w, T)
\]  

(6.4)
such that
\[ 0 \leq y_i \leq x_i - x_i, \forall i \]  
\[ \sum_{i=1}^{n} y_i a \leq B - \sum_{i=1}^{n} x_i a \]  
(6.5)  
(6.6)

where \( x_i = y_i + x_i, \forall i \). Note that all activities with negative values of the coefficient \( g(t_i) < 0 \) should be allocated service level \( x_i^* = x_i \). The remaining problem reduces to a continuous knapsack problem with \( g(t_i) > 0 \) and can be solved by a greedy rule that states that we can sort activities in decreasing order of the ratio \( r_i = \frac{\Phi(\alpha, w, T_{i+1}) - \Phi(\alpha, w, T_i)}{a_i}, i = 1, \ldots, n - 1, \)
\[ r_n = \frac{\Phi(\alpha, w, T_n)}{a_n} \]  
and then allocate the maximum possible service level to each activity in decreasing order of \( r_i \) till the limit on budget \( B \) is reached.

The result indicates that all activities that are at a temporal distance from the end of the service encounter below a certain threshold will be allocated maximum possible service level and those activities that are at a distance above the threshold will be allocated lowest value of their service level. The threshold value \( \Delta t = \frac{1}{m(\alpha, w)} \) is determined by the rate of acclimation and memory decay. For a given set of activities with fixed durations \( t_i, \forall i \), when \( m(\alpha, w) \) is large then fewer activities are allocated high service level, whereas when \( m(\alpha, w) \) is small then more activities are allocated high service level. Intuitively the result implies that when \( \alpha \) and \( w \) are large then few activities that are nearest to the service encounter influence the overall satisfaction of the customer due to high rates of acclimation and memory decay. Whereas when \( w \) or \( \alpha \) is small then more activities influence the overall satisfaction since the customer either acclimates slowly or has a good memory.

6.3 Heterogeneous Customers

We assume that there are \( N \) classes of customers with corresponding rates of acclimation and memory decay given by \( (\alpha_c, w_c), c = 1, \ldots, N \). The probability of \( (\alpha, w) \) belonging to the \( c \)th class is given by \( p_c \) such that \( \sum_{c=1}^{N} p_c = 1 \).

The objective is to maximize the average satisfaction from the remembered utility of the experience. The problem formulation for service provider’s design problem for heterogeneous
customers (SPDP-H) is given by:

$$E[S((i_1, ..., i_n), t)] = \sum_{k=1}^{n} (x_{i_k} - x_{i_{k-1}}) E[\Phi(\alpha, w, T_k)].$$

(6.7)

such that

$$(i_1, ..., i_n) \in \mathcal{P}$$

(6.8)

$$\tau_i \leq t_i \leq \bar{\tau}_i, \forall i$$

(6.9)

$$\sum_i t_i = T.$$  

(6.10)

where $x_{i_0} = b(0)$ and $\mathcal{P}$ is the set of all permutations of activities. We define,

$$E[\Phi(\alpha, w, T_k)] = \sum_{c=1}^{N} p_c \Phi(\alpha_c, w_c, T_k).$$

In this case all the constraints regarding sequencing and duration allocation remain unchanged. Next we analyze the optimal service design for a heterogeneous population of customers.

### 6.3.1 Variable Sequence Fixed Duration

We have shown that for a given $(\alpha, w)$, the optimal sequence is either crescendo or U-shaped. Next we consider VSFD in the scenario where the pseudoconcave function $\Phi(\alpha, w, t)$ is replaced by $E[\Phi(\alpha, w, t)]$.

**Proposition 4** When durations are fixed, there exists an optimal sequence $(i_1^*, ..., i_n^*)$ such that for any two consecutive activities $x_{i_k}^*$ and $x_{i_{k+1}}^*$ the following applies:

(i) If both activities start and finish within $[0, T - \frac{2}{\min_c m(\alpha_c, w_c)}]$, then $x_{i_k}^* > x_{i_{k+1}}^*$

(ii) If both activities start and finish within $[T - \frac{2}{\max_c m(\alpha_c, w_c)} , T]$ or if $k = n - 1$, then $x_{i_k}^* < x_{i_{k+1}}^*$.

Proposition 4 shows that the U-shaped sequence need not be optimal in the presence of a heterogeneous population of customers. This can be verified by a simple counter example where $n = 6, c = 6$. The rates of acclimation $\alpha = [0.0200, 0.0427, 0.0169, 0.0752, 0.0368, 0.0942]$ and
\[ w = [0.0172, 0.8291, 0.6266, 0.5387, 0.6505, 0.7266] \text{ for } c = 1, \ldots, 6 \text{ respectively.} \] The probabilities for belonging to a class is \( p = [0.0396, 0.3677, 0.0060, 0.1233, 0.0754, 0.3881] \). The service level of the activities are \( x = [120, 90, 65, 42, 38, 15] \) and the corresponding durations are \( t = [12, 7, 6, 1, 12, 8] \). By performing enumeration on all possible sequences, the optimal sequence is given by \( x^* = [42, 65, 90, 38, 15, 120] \).

Also if \( m(\alpha_c, w_c), \forall c \) coincide then there is a unique optimal solution as stated in Corollary 3.

**Corollary 3** If \( m(\alpha_1, w_1) = \cdots = m(\alpha_N, w_N) \) then if \( T - \frac{2}{m(\alpha_1, w_1)} > 0 \) the optimal sequence is U-shaped, otherwise it is crescendo.

Though the optimal solution need not be U-shaped but the last two activities are always in increasing order of service levels even in the presence of a heterogeneous population of customers. Moreover, from Corollary 5 we see that if \( T < \frac{2}{\min_c m(\alpha_c, w_c)} \) then the optimal sequence for every individual is crescendo and thereby that becomes the unique optimal service design for the entire population of customers.

**Corollary 4** When \( T < \frac{2}{\min_c m(\alpha_c, w_c)} \), a sequence with increasing service levels is optimal.

### 6.3.1.1 Robustness

In practice, a service provider may not be able to adjust the service design to different customers, each characterized by individual rates of acclimation (\( \alpha \)) and memory decay (\( w \)). Alternatively, the provider may not be able to precisely assess their rates. For any given \( \alpha \) and \( w \), the optimal sequence will either be a crescendo or U-shaped. If a service provider does not offer personalized services then she designs a service encounter with a common sequence of activities for all customers. In this section we compare two design strategies, (i) sequence of increasing service levels (crescendo), and (ii) optimal sequence corresponding to mean values \( \bar{\alpha} \) and \( \bar{w} \) (mean sequence) for the population of customers. We compare the satisfaction obtained from crescendo and mean sequences with the optimal satisfaction for a randomly selected sample of customers and evaluate which sequence performs better.
Table 6.1: Range of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acclimation rate</td>
<td>$\alpha \sim \text{Truncated Normal}$</td>
</tr>
<tr>
<td>Memory decay rate</td>
<td>$w \sim \text{Truncated Normal}$</td>
</tr>
<tr>
<td>Duration of activities</td>
<td>$t \sim \text{Gamma}(k, \theta)$</td>
</tr>
<tr>
<td>Service level</td>
<td>$x \sim \text{Gamma}(2,2)$</td>
</tr>
</tbody>
</table>

Table 6.2: Duration $T$ for $t_i \sim \text{Gamma}(k, \theta)$.

<table>
<thead>
<tr>
<th>$(k, \theta)$</th>
<th>(0.5,0.6)</th>
<th>(1,5)</th>
<th>(1,10)</th>
<th>(2,8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>3</td>
<td>30</td>
<td>75</td>
<td>125</td>
</tr>
</tbody>
</table>

across all customers on an average. Therefore, any customer $m$ with rates of acclimation and memory decay given by $(\alpha_m, w_m)$ will have an optimal design corresponding to $m(\alpha_m, w_m)$ giving satisfaction $S^*_m$. If this customer is subjected to a sequence that is crescendo then her corresponding satisfaction is denoted by $S_m$ whereas satisfaction from mean sequence will be denoted as $\overline{S}_m$. We evaluate the suboptimality gaps corresponding to crescendo and mean sequence by $\frac{100(S^*_m - S_m)}{S^*_m}$ and $\frac{100(S^*_m - \overline{S}_m)}{S^*_m}$ respectively.

The problem instances are generated as follows. We selected a service encounter with seven activities, i.e., $n = 7$. We drew a sample for service level $x \sim \text{Gamma}(2,2)$ and fixed it across all problem instances that were tested. Table 6.1 shows the distribution of the problem parameters. We drew one sample for the activity durations $t_i \sim \text{Gamma}(k, \theta)$, for four different pairs of values of $(k, \theta)$ as shown in Table 6.2. The corresponding total duration $T$ was evaluated for each of the four samples as shown in Table 6.2. We assume that the rates of acclimation and memory decay of all customers belong to Truncated normal distribution with variance of respective parameters as stated in Table 6.3. Hence, for a given sample of customers of size 1500 we obtain the corresponding variance in their values of $m(\alpha, w)$. Therefore, there are four instances of duration allocated to activities and four population of customers where each population is identified by the variance in $m(\alpha, w)$.
Table 6.3: Variance\((m(\alpha, w))\) corresponding to \(\alpha \sim \text{Truncated Normal}(\sigma(\alpha), (0, \infty))\) and \(w \sim \text{Truncated Normal}(\sigma(w), (0, \infty))\).

\[
\begin{array}{|c|cccc|}
\hline
(\sigma(\alpha), \sigma(w)) & (0.01, 0.05) & (0.1, 0.5) & (1.0, 0.8) & (1.2) \\
\hline
\text{Variance}(m(\alpha, w)) & 0.1 & 0.3 & 0.5 & 1.0 \\
\hline
\end{array}
\]

Table 6.4: Average suboptimality gap in % for crescendo

<table>
<thead>
<tr>
<th>Variance ((m(\alpha, w)))</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.10</td>
<td>0.12</td>
<td>24.45</td>
</tr>
<tr>
<td>30</td>
<td>15.82</td>
<td>42.84</td>
<td>53.84</td>
<td>70.25</td>
</tr>
<tr>
<td>75</td>
<td>32.45</td>
<td>60.32</td>
<td>70.96</td>
<td>85.55</td>
</tr>
<tr>
<td>125</td>
<td>54.69</td>
<td>81.05</td>
<td>87.71</td>
<td>96.71</td>
</tr>
</tbody>
</table>

across the population. In all, there are sixteen problem instances for which the average gap from crescendo and mean sequence are evaluated.

The optimal sequence for any \((\alpha, w)\) is obtained by enumeration. We then evaluated the average gap by taking the mean of gaps of 1500 randomly selected customers in Tables 6.4 and 6.5.

Gaps in Table 6.4 show that as variability in \(m(\alpha, w)\) increases, the suboptimality gap for crescendo increases. Also, for a given population of customers with fixed variance in \(m(\alpha, w)\), the longer the total duration \(T\) the greater is the suboptimality gap for crescendo sequence. This is because as \(T\) or \(m(\alpha, w)\) increases, the quantity \(T - \frac{2}{m(\alpha, w)}\) also increases, and the optimal sequence of any individual customer tends to be U-shaped and in the extreme case, the optimal sequence has only two activities in increasing order of service level. Hence, the
Table 6.5: Average suboptimality gap in % for mean ($\bar{\alpha}, \bar{w}$)

<table>
<thead>
<tr>
<th>Variance ($m(\alpha, w)$)</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.10</td>
<td>0.22</td>
<td>4.62</td>
</tr>
<tr>
<td>30</td>
<td>12.22</td>
<td>20.32</td>
<td>24.16</td>
<td>9.01</td>
</tr>
<tr>
<td>75</td>
<td>7.09</td>
<td>27.37</td>
<td>8.66</td>
<td>3.01</td>
</tr>
<tr>
<td>125</td>
<td>31.75</td>
<td>23.32</td>
<td>9.30</td>
<td>2.14</td>
</tr>
</tbody>
</table>

benefit of designing the sequence for the average population parameters $\bar{\alpha}$ and $\bar{w}$ is greater for service encounters that are longer and have higher variance in $m(\alpha, w)$ of the population.

By comparing the gaps from crescendo and mean sequences as stated in Tables 6.4 and 6.5 we can clearly see that the mean sequence is performing better than crescendo when the variability in the rates of acclimation and memory decay is higher. This shows that the U-shaped sequence is more robust towards variation across population than the crescendo sequence of activities.

### 6.3.2 Fixed Sequence Variable Duration

Next we look at service encounters with fixed sequence where the service provider has to decide how much duration should be allocated across all activities. Therefore, we drop the constraint (6.8) from SPDP-H. We find that the presence of heterogeneous customers changes the characterization of the optimal solution.

**Proposition 5** For a fixed subsequence of activities $(x_1, \ldots, x_r)$, the optimal duration allocation is such that:
1. Suppose that \( T < \frac{1}{\max_c m(\alpha_c, w_c)} \), and there exists an activity \( m, l < m < r \), such that \( t^*_m > \tau_m \). Then if \( x_l < \cdots < x_r \), then \( t^*_j = \tau_j, j = m+1, \ldots, r \); and if \( x_l > \cdots > x_r \), \( t^*_j = \tau_j, j = l, \ldots, m-1 \).

2. Suppose that \( \frac{1}{\min_c m(\alpha_c, w_c)} < \sum_{i=r}^n \tau_i \) and that there exists an activity \( m, l < m < r \), such that \( t^*_m > \tau_m \). Then if \( x_l < \cdots < x_r \), then \( t^*_j = \tau_j, j = l, \ldots, m-1 \); and if \( x_l > \cdots > x_r \), \( t^*_j = \tau_j, j = m+1, \ldots, r \).

From Proposition 5 we identify the two scenarios where the optimal duration allocation is similar to the result we have for homogeneous customers. When \( \max_c m(\alpha_c, w_c) \) is very small then duration should be allocated in increasing order of service levels for both increasing and decreasing subsequences. This corresponds to the scenario when all the customers in the population have low rates of memory decay or acclimation. Whereas when \( \min_c m(\alpha_c, w_c) \) is very large then it is optimal to allocate duration in decreasing order of service level for any subsequence that is either all increasing or all decreasing in service levels. This corresponds to the scenario where all the customers in the population have high rates of acclimation and memory decay.

**Corollary 5** When \( m(\alpha_1, w_1) = \cdots = m(\alpha_N, w_N) \) then the duration allocation is the same as that for the homogeneous customer population with \( m(\alpha, w) = m(\alpha_c, w_c) \) for any \( c \).

Corollary 5 states that if all the customers have the same value of \( m(\alpha, w) \), then a common duration allocation rule will be optimal for all, and the optimal solution can be obtained by treating them as a homogeneous population of customers.

### 6.3.3 Variable Sequence Variable Duration

We next consider a service encounter where both sequence and duration are variable and the service provider serves a population of customers with diverse rates of acclimation and memory decay.

**Corollary 6** In VSVD,
1) If \( T < \frac{2}{\min_c m(\alpha_c, w_c)} \), a sequence with increasing service levels is optimal. Moreover,

1.1) If \( T < \frac{1}{\max_c m(\alpha_c, w_c)} \) the optimal duration allocation is characterized by an index \( k \) such that \( t_{ij} = \tau_{ij}, 1 \leq j \leq k - 1, t_{ik} = \tau, \tau_{ik} \leq \tau \leq \tau_{ik}, \) and \( t_{ij} = \tau_{ij}, k < j \leq n. \)

1.2) If \( \frac{1}{\min_c m(\alpha_c, w_c)} < \frac{1}{\tau_{in}} \), then the optimal duration allocation is characterized by an index \( k \) such that \( t_{i}^{*} = \tau_{i}, j = 1, \ldots, k - 1, t_{ik}^{*} = \tau, \tau_{ik} \leq \tau \leq \tau_{ik}, t_{ij}^{*} = \tau_{ij}, j = k + 1, \ldots, n. \)

2) If \( \frac{2}{\min_c m(\alpha_c, w_c)} < T \), a noisy U-shaped sequence with a peak at the start and the end is optimal such that:

for any subsequence \( (x_l, \ldots, x_r) \), if \( \frac{1}{\min_c m(\alpha_c, w_c)} < \sum_{i=r}^{l} \tau_{i} \) and that there exists an activity \( m, l < m < r, \) such that \( t_{m}^{*} > \tau_{m} \). Then if \( x_l < \cdots < x_r \), then \( t_{j}^{*} = \tau_{j}, j = l, \ldots, m - 1; \) and if \( x_l > \cdots > x_r \), \( t_{j}^{*} = \tau_{j}, j = m + 1, \ldots, r. \)

Corollary 6 combines the result from Propositions 4 and 5. We find that when all customers have very small rates of memory decay or acclimation \( i.e., T < \frac{1}{\max_c m(\alpha_c, w_c)} \) then the optimal sequence for the entire population is a crescendo, i.e., increasing order of service levels and the optimal duration allocation is in increasing order of service levels. Whereas, when the rates of memory decay and acclimation are very large then it is optimal to have a noisy U-shaped (starting and ending with peaks in addition to having one or more peaks in the middle) sequence and duration should be allocated across any subsequence of activities in decreasing order of service levels.
CHAPTER 7

Conclusions and Future Research

In this research, we analytically determine how to sequence and allocate duration to activities in a service encounter so as to maximize ex-post customer satisfaction in the presence of memory decay and acclimation. Memory decay favors high service levels at the end of the encounter. Whereas, acclimation favors a high gradient of service level from one stage to the next, thereby inducing a steep rise or a gradual fall.

We show that, individually, the two behavioral biases give the same monotonic service design. However, when considered jointly, they can act as opposing forces. In particular, if either memory decay or acclimation is low, it is optimal to sequence activities in increasing order of service levels and to lengthen the duration of activities with the highest service levels. In contrast, when both memory decay and acclimation are high, it is optimal to sequence activities in a U-shape fashion and to lengthen the duration of activities with the lowest service levels, located in the middle of the sequence.

Overall, our results suggest that “finishing strong” is a valid recommendation in the presence of either memory decay or acclimation. But when both effects are present, the design recommendation is more subtle; specifically, as one should aim at maximizing the gradient near the end. While an anti-climax is never desirable, a never-ending grand finale should also be avoided. Moreover, it may be a good idea to introduce some respite just before the grand finale so as to accentuate its strong character. Similar to the (controversial) idea of adding light pain at the end to induce customers to forget a more painful event [Kahneman et al., 1993], it may be desirable to voluntarily degrade the service level of middle activities to accentuate the strong character of the ending activities as empirically observed by Nelson and Meyvis [2008]. Pursuing this logic one step further, it may be desirable to even get
rid of some very high-intensity activities in order to avoid a decline in service level if they are positioned at the beginning or accentuate the final gradient in service levels if they are sequenced towards the end. In fact in the presence of memory decay and acclimation, decreasing the service level on every activity may generate greater satisfaction, provided that this leads to a steep gradient at the end. Hence, properly sequencing and allocating duration to a service encounter can more than compensate for inferior service levels.

Another key insight from this research is the fact that in the presence of customers who are heterogeneous in terms of their rates of acclimation and memory decay, the service provider should sequence the activities in the U-shaped order that is optimal for the average rate of acclimation and memory decay of the population. This U-shaped sequence offers a higher average satisfaction across the population of customers than when activities are sequenced as a crescendo, i.e., in increasing order of service levels.

This thesis takes a first step toward designing service encounters using an analytical framework, by capturing behavioral phenomena that influence the evaluation of past experiences by customers. We have employed a parsimonious models to generate insights from this framework. Natural research extensions could relax the assumptions outlined in the introduction. For example,

- What is the optimal total duration of a service encounter?

- What is the optimal service design in the presence of other behavioral biases, such as satiation or loss aversion? The optimal future consumption plan in the presence of satiation is in general U-shaped [Baucells and Sarin, 2013], is that also true with retrospective evaluation?

- How should a service be designed when customers can join after the beginning of the encounter or leave before the end?

- What is the optimal service design when service experience is multidimensional?

- Does the optimal service design change if we assume the power law of forgetting in place of exponential decay of memory [Wixted and Carpenter, 2007]?
We hope that this research will induce further interest in engineering service experiences to maximize customer satisfaction.
CHAPTER 8

APPENDICES

8.1 Appendix A: Proofs

Lemma A-1 The function $\Phi(t)$ is strictly pseudoconcave in $t$ with a stationary point at $\frac{1}{m(\alpha,w)}$.

Proof Suppose $\alpha < w$. We have:

$$\Phi'(t) = \frac{e^{-\alpha t} - e^{-wt}}{w - \alpha} > 0 \Leftrightarrow w > \alpha e^{t(w-\alpha)}, \forall w > \alpha \Leftrightarrow \frac{\ln w - \ln \alpha}{w - \alpha} > t. $$

Since $\Phi'(t) > 0$ when $t < \frac{1}{m(\alpha,w)}$, $\Phi'(t) = 0$ when $t = \frac{1}{m(\alpha,w)}$, and $\Phi'(t) < 0$ when $t > \frac{1}{m(\alpha,w)}$, $\Phi(t)$ is strictly pseudoconcave. Note that the result is symmetric for $w < \alpha$.

Lemma A-2 The function $f(T_{ik}) = \Phi(T - T_{ik} + t_{ik}) - \Phi(T - T_{ik})$ is strictly pseudoconvex in $T_{ik}$.

Proof Suppose $\alpha < w$. We have:

$$f'(T_{ik}) = \frac{\alpha e^{-\alpha(T-T_{ik}+t_{ik})} - w e^{-w(T-T_{ik}+t_{ik})}}{w - \alpha} - \frac{\alpha e^{-\alpha(T-T_{ik})} - w e^{-w(T-T_{ik})}}{w - \alpha}$$

Consider $t_0 = T - \frac{1}{(w-\alpha)} \ln \frac{w(1-e^{-wT_{ik}})}{\alpha(1-e^{-\alpha t_{ik}})}$. Since $f'(T_{ik}) < 0$ when $T_{ik} < t_0$, $f'(T_{ik}) = 0$ when $T_{ik} = t_0$, and $f'(T_{ik}) > 0$ when $T_{ik} > t_0$, $f(T_{ik})$ is strictly pseudoconvex. Note that the result is symmetric for $w < \alpha$.

Lemma A-3 $\Phi(t)$ is concave $\forall t \in [0, \frac{1}{m(\alpha,w)}]$ and convex $\forall t \in [\frac{1}{m(\alpha,w)}, T]$.
Proof Since, $\Phi''(t) = \frac{\alpha^2 e^{-\alpha t} - w^2 e^{-\alpha t}}{w - \alpha}$, therefore, $\Phi''(t) \leq 0, \forall t \in [0, \frac{1}{m(\alpha, w)}]$ and $\Phi''(t) \geq 0, \forall t \in [\frac{1}{m(\alpha, w)}, T]$. Hence, $\Phi(t)$ is concave $\forall t \in [0, \frac{1}{m(\alpha, w)}]$ and convex $\forall t \in [\frac{1}{m(\alpha, w)}, T]$. \hfill

Proof of Proposition 1. The proof uses Lemma A-3. Let $S^*$ be the satisfaction obtained from the optimal sequence $(i_1^*, \ldots, i_n^*)$. For any $j$, let $S_j^*$ be the satisfaction obtained by interchanging $i_{j-1}^*$ and $i_j^*$ in the optimal sequence. Therefore,

$$S^* - S_j^* = (x_{i_{j-1}^*} - x_{i_j^*}) \left( (-\Phi(T_j) + \Phi(T_{j+1})) + (\Phi(T_{j-1}) - \Phi(t_{i_{j-1}^*} + T_{j+1}))) \right).$$

We next consider three scenarios: (i) when the two activities start and end within $[0, T - \frac{1}{m(\alpha, w)}]$, (ii) when the two activities start and end within $[T - \frac{1}{m(\alpha, w)}, T]$, and (iii) when $j = n$.

(i) If $i_{j-1}^*$ and $i_j^*$ start and finish within $[0, T - \frac{1}{m(\alpha, w)}]$, then by Lemma A-3, the corresponding $\Phi'(t)$ is increasing $\forall t \in [\frac{1}{m(\alpha, w)}, T]$. Hence,

$$\Phi(T_j) - \Phi(T_{j+1}) < \Phi(t_{i_{j-1}^*} + T_j) - \Phi(t_{i_{j-1}^*} + T_{j+1}).$$

Therefore, by optimality of $(i_1^*, \ldots, i_n^*)$, we have $x_{i_j^*} < x_{i_{j-1}^*}$.

(ii) If $i_{j-1}^*$ and $i_j^*$ start and finish within $[T - \frac{1}{m(\alpha, w)}, T]$, then by Lemma A-3, the corresponding $\Phi'(t)$ is decreasing $\forall t \in [0, \frac{1}{m(\alpha, w)}]$. Hence,

$$\Phi(T_j) - \Phi(T_{j+1}) > \Phi(t_{i_{j-1}^*} + T_j) - \Phi(t_{i_{j-1}^*} + T_{j+1}).$$

Therefore, by optimality of $(i_1^*, \ldots, i_n^*)$, we have $x_{i_j^*} > x_{i_{j-1}^*}$.

(iii) By interchanging the last two activities $i_{n-1}^*, i_n^*$, we obtain a suboptimal sequence with corresponding satisfaction $S_n^*$. Therefore,

$$S^* - S_n^* = (x_{i_n^*} - x_{i_{n-1}^*}) \left( \Phi(t_{i_n^*}) + \Phi(t_{i_{n-1}^*}) - \Phi(t_{i_{n-1}^*} + t_{i_n^*}) \right)$$

$$= \frac{x_{i_n^*} - x_{i_{n-1}^*}}{w - \alpha} \left\{ (1 - e^{-\alpha t_{i_n^*}})(1 - e^{-\alpha t_{i_{n-1}^*}}) - (1 - e^{-\alpha t_{i_{n-1}^*}})(1 - e^{-\alpha t_{i_n^*}}) \right\}.$$ 

Hence, by optimality of $(i_1^*, \ldots, i_n^*)$, we have $x_{i_n^*} > x_{i_{n-1}^*}$. \hfill
Lemma A-4 Suppose that $x_i$ is the service level of the activity at the $i$th position in the sequence. When $x_1 < \cdots < x_n$, FSVD is strictly pseudoconcave, whereas, when $x_1 > \cdots > x_n$, FSVD is strictly pseudoconvex.

Proof We show that the corresponding Hessian ($H$) for $S(t)$ is negative definite for pseudoconcave and positive definite for pseudoconvex, at a stationary point of $S(t)$. By equation (4.1), $S(t) = \sum_{i=1}^{n}(x_i - x_{i-1})\Phi(T_i)$. Taking the partial derivatives of $S(t)$ with respect to $t_i, \forall i$ we obtain

$$\frac{\partial S(t)}{\partial t_i} = \sum_{j=1}^{i} (x_j - x_{j-1})\Phi'(T_j), \forall i.$$ 

So the stationary point has to satisfy the following set of equations:

$$\sum_{j=i}^{n} t_j = \frac{1}{m(\alpha, w)}, \forall i$$

(8.1)

Therefore, $t^* = (0, \ldots, 0, \frac{1}{m(\alpha, w)})$ is the unique stationary point satisfying equations (A-1).

We evaluate the Hessian at $t^*$ to obtain for any $y \in \mathbb{R}^n$

$$y^T H y = \alpha \left( \frac{\alpha}{w} \right)^{\omega-\alpha} \left( -(x_2 - x_1)(\sum_{i=2}^{n} y_i)^2 - (x_3 - x_2)(\sum_{i=3}^{n} y_i)^2 \cdots - (x_n - x_{n-1})y_n^2 \right).$$

Clearly, for $x_1 < \cdots < x_n$, $y^T H y < 0$, and for $x_1 > \cdots > x_n$, $y^T H y > 0 \ \forall y \in \mathbb{R}^n$.

Lemma A-5 In FSVD the necessary conditions for optimality are given by

$$\frac{\partial S(t)}{\partial t_i} = \lambda_i + \mu, \forall i \in I_U$$

$$\frac{\partial S(t)}{\partial t_i} = \mu, \forall i \in I_M$$

$$\frac{\partial S(t)}{\partial t_i} = \mu - \lambda_i, \forall i \in I_L,$$

(8.2)

and constraints (4.2) and (4.3) where $\mu \in \mathbb{R}, \lambda_i \geq 0$ and $\lambda_i \geq 0$. $I_U, I_M, I_L$ are three disjoint sets such that $I = I_U \cup I_M \cup I_L$ and $t_i = \tau_i, \forall i \in I_U$, $t_i = \tau, \forall i \in I_M$, $t_i = \tau_i, \forall i \in I_L$.

Proof The necessary conditions for optimality of $t^*$ are given by the following Karush-Kuhn-Tucker conditions [Boyd and Vandenberghe, 2004]:

57
1. The stationarity condition gives $\frac{\partial S(t^*)}{\partial t_i} = \mu + \Delta_i - \overline{\lambda}_i = 0, \forall i$ at $t^*$.

2. Complementary slackness gives $\mu(T - \sum_{i=1}^n t^*_i) = 0$, $\overline{\lambda}_i(t^*_i - t^*_i) = 0, \forall i$ and $\Delta_i(t^*_i - \tau_i) = 0, \forall i$.

3. Primal feasibility implies $t^*$ satisfies the constraints (4.2) and (4.3).

4. Dual feasibility implies $\mu \in \mathbb{R}$ and $\Delta_i, \overline{\lambda}_i \geq 0, \forall i$.

From the complementary slackness condition, $\Delta_i = 0, \forall i \in I_U$ and $\overline{\lambda}_i = 0, \forall i \in I_L$ and both $\Delta_i = 0, \overline{\lambda}_i = 0, \forall i \in I_M$. □

**Proof of Proposition 2.** We use Lemmas A-1 and A-5 for this proof. We have

$$\frac{\partial S(t^*)}{\partial t_i} = \sum_{j=1}^i (x_j - x_{j-1})\Phi'(T_j), \forall i,$$

where $x_0 = b(0)$.

1. Suppose that $x_l < \cdots < x_r$. Since $T < \frac{1}{m(\alpha, w)}$, $\Phi'(t) > 0, \forall t \in [0, T]$ by Lemma A-1, and therefore $\frac{\partial S(t)}{\partial t_1} < \cdots < \frac{\partial S(t)}{\partial t_r}, \forall t \geq (\tau_1, \ldots, \tau_n)$. Hence, from condition (A-2), if $t_m > \tau_m$, $\frac{\partial S(t)}{\partial t_m} = \mu$ or $\frac{\partial S(t)}{\partial t_m} = \mu + \overline{\lambda}_m$ at $t^*$. Therefore, $\frac{\partial S(t^*)}{\partial t_j} = \mu + \overline{\lambda}_j, j = m + 1, \ldots, r$ at $t^*$. From Lemma A-5, $t^*_j = \tau_j, j = m + 1, \ldots, r$. The proof for a decreasing subsequence is similar.

2. We show the result by contradiction. Suppose that $x_l < \cdots < x_r$ and that $t_l = \tau_l$ and $t_r = \tau_r$. Therefore, condition (A-2) implies that $\frac{\partial S(t^*)}{\partial t_i} = \mu - \Delta_i$ and $\frac{\partial S(t^*)}{\partial t_r} = \mu - \overline{\lambda}_r$. If for any $i, l < i < r$, we have $t^*_i > \tau_i$, then $\frac{\partial S(t^*)}{\partial t_i} = \mu$ or $\frac{\partial S(t^*)}{\partial t_r} = \mu + \overline{\lambda}_i$, by condition (A-2). If $\sum_{i=1}^n t^*_h \leq \frac{1}{m(\alpha, w)}$, then $\Phi'(t) > 0$ for $t \leq \sum_{i=1}^n t^*_h < \frac{1}{m(\alpha, w)}$ from Lemma A-1 and therefore $\frac{\partial S(t^*)}{\partial t_i} > \frac{\partial S(t^*)}{\partial t_r}$, thereby a contradiction. If $\sum_{i=1}^n t^*_h > \frac{1}{m(\alpha, w)}$, then $\Phi'(t) < 0$ for $t \geq \sum_{i=1}^n t^*_h$ from Lemma A-1 and therefore $\frac{\partial S(t^*)}{\partial t_i} > \frac{\partial S(t^*)}{\partial t_r}$, thereby a contradiction. Therefore, from Lemma A-5, $t^*_l = \tau_l, \forall l \leq i < r$. The proof for a decreasing subsequence is similar.

3. Suppose that $x_l < \cdots < x_r$. When $\sum_{i=r}^n \tau_i > \frac{1}{m(\alpha, w)}$, we have $\Phi'(t) < 0, \forall t \in (\sum_{i=r}^n \tau_i, T]$ by Lemma A-1, and therefore $\frac{\partial S(t^*)}{\partial t_l} > \cdots > \frac{\partial S(t^*)}{\partial t_r}, \forall t \geq (\tau_1, \ldots, \tau_n)$. From condition (A-2), since $t_m > \tau_m$, $\frac{\partial S(t^*)}{\partial t_m} = \mu$ or $\frac{\partial S(t^*)}{\partial t_m} = \mu + \overline{\lambda}_m$ at $t^*$. Therefore, $\frac{\partial S(t^*)}{\partial t_j} = \mu + \overline{\lambda}_j, j = l, \ldots, m - 1$ at $t^*$. Therefore, from Lemma A-5, $t^*_j = \tau_j, j = l, \ldots, m - 1$. The proof for a decreasing subsequence is similar. □
Lemma A-6 For endogenous service level the necessary conditions for optimality are given by

\[
\frac{\partial S(t)}{\partial x_i} = \lambda_i + a_i \mu, \forall i \in I_U \\
\frac{\partial S(t)}{\partial t_i} = a_i \mu, \forall i \in I_M \\
\frac{\partial S(t)}{\partial t_i} = a_i \mu - \lambda_i, \forall i \in I_L,
\]

and constraints (6.2) and (6.3) where \( \mu \geq 0, \lambda_i \geq 0 \) and \( \lambda_i \geq 0 \). \( I_U, I_M, I_L \) are three disjoint sets such that \( I = I_U \cup I_M \cup I_L \) and \( x_i = \pi_i, \forall i \in I_U \), \( x_i = x, x_i < x < \tau_i, \forall i \in I_M \) and \( x_i = \bar{x}_i, \forall i \in I_L \).

**Proof** The necessary conditions for optimality of \( x^* \) are given by the following Karush-Kuhn-Tucker conditions [Boyd and Vandenberghe, 2004]:

1. The stationarity condition gives \( \frac{\partial S(x)}{\partial x_i} - a_i \mu + \lambda_i - \lambda_i = 0, \forall i \) at \( x^* \).
2. Complementary slackness gives \( \mu(B - \sum_{i=1}^n a_i x_i^*) = 0, \lambda_i(\pi_i - x_i^*) = 0, \forall i \) and \( \lambda_i(x_i^* - \bar{x}_i) = 0, \forall i \).
3. Primal feasibility implies \( x^* \) satisfies the constraints (6.2) and (6.3).
4. Dual feasibility implies \( \mu, \lambda_i, \lambda_i \geq 0, \forall i \).

From the complementary slackness condition, \( \lambda_i = 0, \forall i \in I_U \) and \( \lambda_i = 0, \forall i \in I_L \) and both \( \lambda_i = 0, \lambda_i = 0, \forall i \in I_M \).

**Proof of Proposition 3** For fixed sequence and duration the service provider’s problem reduces to a linear program in \( x_{ik} \) with upper bound, lower bound, and budget constraints. The coefficient of \( x_{ik} \) is \( g(t_k) = \Phi(\alpha, w, T_k) - \Phi(\alpha, w, T_{k+1}), \forall k = 1, \ldots, n-1 \). From Lemma A-6 we obtain \( g(t_k) = \lambda_{ik} - \lambda_{ik} + a_{ik} \mu, \forall k \) where \( \lambda_{ik} \geq 0 \) and \( \lambda_{ik} \geq 0 \) are the Lagrangian multipliers for the upper and lower bound constraints (6.2) respectively. Whereas \( \mu \geq 0 \) is the multiplier corresponding to (6.3). The complementary slackness condition gives \( \lambda_{ik}(\pi_{ik} - x_{ik}) = 0, \lambda_{ik}(x_{ik} - \bar{x}_{ik}) = 0, \forall k \) and \( \mu(B - \sum_{k=1}^n a_{ik} x_{ik}) = 0 \).

1. If \( T_{ik+1} < \frac{1}{m(\alpha, w)} < T_{ik} \), then \( g(t_{ij}) < 0, \forall j = 1, \ldots, k \) and \( g(t_{ij}) > 0, \forall j = k + 1, \ldots, n \). If the optimal solution is such that \( x_{ij}^* = \pi_{ij} \) we get \( g(t_{ij}) = \lambda_{ij} + a_{ij} \mu, \) if \( \pi_{ij} < x_{ij}^* < \bar{x}_{ij} \) we get \( g(t_{ij}) = a_{ij} \mu, \) and if \( x_{ij}^* = \bar{x}_{ij} \) we get \( g(t_{ij}) = -\lambda_{ij} + a_{ij} \mu \). Therefore, the optimal solution is
given by $x^*_i > x^*_j$, $j = k + 1, \ldots, n$ and $x^*_i = x^*_j$, $j = 1, \ldots, k$.

2. If $\frac{1}{m(\alpha, w)} < T_{i_n}$ then $\Delta g(t_{i_n}) < 0, \forall j = 1, \ldots, n - 1$ and $g(t_{i_n}) = \Phi(\alpha, w, T_{i_n}) > 0$. From the KKT condition we get that the optimal solution is $x^*_i < x^*_j$ and $x^*_i = x^*_j, \forall j = 1, \ldots, n - 1$.

**Proof of Proposition 4** Let $E[S^*]$ be the expected satisfaction corresponding to the optimal sequence $(i^*_1, \ldots, i^*_n)$ and $E[S_k]$ be the expected satisfaction corresponding to the sequence obtained by interchanging activities $i^*_k$ and $i^*_{k+1}$. Then

$$E[S^*] - E[S_k] = (x^*_{i_{k+1}} - x^*_k)((E[\Phi(\alpha, w, T_k)] - E[\Phi(\alpha, w, T_{k+2})]) - (E[\Phi(\alpha, w, T_k)] - E[\Phi(\alpha, w, t^*_k + T_{k+2})])).$$

We consider three scenarios, (i) when both the activities start and end within $[0, T - \frac{2}{\min m(\alpha_c, w_c)}]$, (ii) when both the activities start and end within $[T - \frac{2}{\max m(\alpha_c, w_c)}, T]$, and (iii) when $k = n - 1$.

(i) From pseudoconcavity of $\Phi(\alpha_c, w_c, t)$, $\forall c$, we have for any given $(\alpha_c, w_c)$, if both the activities start and finish within $[0, T - \frac{2}{m(\alpha_c, w_c)}]$ then

$$\Phi(\alpha_c, w_c, T_k) - \Phi(\alpha_c, w_c, T_{k+2}) < (\Phi(\alpha_c, w_c, T_k) - \Phi(\alpha_c, w_c, t^*_k + T_{k+2})).$$

Therefore, $E[(\Phi(\alpha_c, w_c, T_k) - \Phi(\alpha_c, w_c, T_{k+2})) - (\Phi(\alpha_c, w_c, T_k) - \Phi(\alpha_c, w_c, t^*_k + T_{k+2}))] < 0, \forall c$. Hence, by optimality of $(i^*_1, \ldots, i^*_n)$ we have $x^*_k > x^*_{k+1}$.

(ii) Let $i^*_k$ and $i^*_{k+1}$ start and finish within $[T - \frac{2}{\max m(\alpha_c, w_c)}, T]$. We know $\Phi''(\alpha, w, t) = \frac{\alpha^2e^{-\alpha t} - w^2e^{-wt}}{we^{-\alpha t}} < 0$ when $t < \frac{2}{m(\alpha, w)}$ for any $(\alpha, w)$. Therefore,

$$\Phi(\alpha_c, w_c, T_k) - \Phi(\alpha_c, w_c, T_{k+2}) > \Phi(\alpha_c, w_c, T_k) - \Phi(\alpha_c, w_c, t^*_k + T_{k+2}), \forall c.$$
(iii) $E[S^*] - E[S_{n-1}] = (x_{i_n^*} - x_{i_{n-1}^*})(E[\Phi(\alpha, w, t_{i_n^*})] + E[\Phi(\alpha, w, t_{i_{n-1}^*})] - E[\Phi(\alpha, w, t_{i_{n-1}^*} + t_{i_n^*})]).$

Since, $\Phi(\alpha, w, t_{i_n^*}) + \Phi(\alpha, w, t_{i_{n-1}^*}) - \Phi(\alpha, w, t_{i_{n-1}^*} + t_{i_n^*}) > 0$ for any $(\alpha, w)$, we have $x_{i_n^*} - x_{i_{n-1}^*} > 0$ for optimality. 

**Lemma A-7** In FSVD for heterogeneous customers the necessary conditions for optimality are given by

$$\frac{\partial E[S(t)]}{\partial t_i} = \lambda_i + \mu, \forall i \in I_U$$

$$\frac{\partial E[S(t)]}{\partial t_i} = \mu, \forall i \in I_M$$

$$\frac{\partial E[S(t)]}{\partial t_i} = \mu - \lambda_i, \forall i \in I_L,$$

and constraints (6.9) and (6.10) where $\mu \in \mathbb{R}$, $\lambda_i \geq 0$ and $\bar{\lambda}_i \geq 0$. $I_U, I_M, I_L$ are three disjoint sets such that $I = I_U \cup I_M \cup I_L$ and $t_i = \tau_i, \forall i \in I_U$, $t_i = \tau, \tau_i < \tau < \tau_i, \forall i \in I_M$ and $t_i = \tau_i, \forall i \in I_L$.

**Proof** The necessary conditions for optimality of $t^*$ are given by the following Karush-Kuhn-Tucker conditions [Boyd and Vandenberghe, 2004]:

1. The stationarity condition gives $\frac{\partial E[S(t)]}{\partial t_i} = \mu + \lambda_i - \bar{\lambda}_i = 0, \forall i$ at $t^*$.
2. Complementary slackness gives $\mu(T - \sum_{i=1}^n t_i^*) = 0, \bar{\lambda}_i(t_i^* - \tau_i) = 0, \forall i$ and $\lambda_i(t_i^* - \tau_i) = 0, \forall i$.
3. Primal feasibility implies $t^*$ satisfies the constraints (6.9) and (6.10).
4. Dual feasibility implies $\mu \in \mathbb{R}$ and $\lambda_i, \bar{\lambda}_i \geq 0, \forall i$.

From the complementary slackness condition, $\lambda_i = 0, \forall i \in I_U$ and $\bar{\lambda}_i = 0, \forall i \in I_L$ and both $\lambda_i = 0, \bar{\lambda}_i = 0, \forall i \in I_M$.

**Proof of Proposition 5** We know that

$$\frac{\partial E[S(t^*)]}{\partial t_i} = \sum_{k=1}^i (x_k - x_{k-1})E[\Phi'(\alpha, w, T_k)].$$

therefore, from lemma A-7, we can construct the necessary conditions for optimality of $t^*$ given by the Karush-Kuhn-Tucker conditions.

(i) If $T < \frac{1}{\max_c m(\alpha_c, w_c)}$ then $T < \frac{1}{m(\alpha_c, w_c)}, \forall c$, therefore $\Phi'(\alpha_c, w_c, t) > 0, \forall c, \forall t$. The
remaining result follows from Proposition 2 for homogeneous customers.

(ii) \( \frac{1}{\min_c m(\alpha_c, w_c)} < \sum_{i=r}^{n} \tau_i \) then \( \frac{1}{m(\alpha_c, w_c)} < \sum_{i=r}^{n} \tau_i, \forall i \), therefore \( \Phi'(\alpha_c, w_c, t) < 0, \forall c, \forall t \).
The remaining result follows from Proposition 2 for homogeneous customers.

The proof for the decreasing subsequence is similar. ■

8.2 Appendix B: Mathematical Programming Formulation of SPDP

Define \( z_{ij} \) as a binary variable indicating if activity \( i \) directly or indirectly precedes \( j \) and define \( c_i \) as the completion time of activity \( i \). We keep constraints (4.2)-(4.3) and express constraint (4.4) in SPDP in terms of \( z_{ij} \)'s and \( c_i \)'s. Constraints (8.5)-(8.6) ensure that for any pair of activities, one has to precede the other. Let \( Q \) be the set of precedence constraints between pairs of activities. Then constraint (8.7) ensures that activity \( j \) precedes activity \( k \) for all such pairs of precedence constraints in \( Q \). Constraint (8.8) ensures that the completion time of an activity must be greater than its duration and smaller than the total duration of the service encounter. Finally, constraint (8.9) ensures that for any pair of activities, the completion time of the preceding activity is less than that of the succeeding activity.

\[
\max_{z, t, c} S(z, t, c) = \sum_{i=1}^{n} x_i \left( \Phi(T - c_i + t_i) - \Phi(T - c_i) \right)
\]
subject to

\[ z_{ij} + z_{ji} = 1, \forall i, j, i \neq j \] (8.5)
\[ z_{ii} = 0, \forall i \] (8.6)
\[ z_{jk} = 1, \forall (j, k) \in Q \] (8.7)
\[ t_i \leq c_i \leq T, \forall i \] (8.8)
\[ c_i \leq c_j - t_j + (1 - z_{ij})T, \forall i, j, i \neq j \] (8.9)
\[ z_{ij} \in \{0, 1\}, \forall i, j. \]
8.3 Appendix C: Upper Bound Formulation

Lemma C-1 When \( T > \frac{1}{2m(\alpha, w)} \), SPDP is equivalent to SPDP-U.

Proof Proof. We establish the equivalence result in two steps by showing that (i) any optimal solution for SPDP is feasible in SPDP-U and has the same objective value, and (ii) any optimal solution for SPDP-U is feasible in SPDP and has the same objective value.

(i) From Proposition 1 we know that any optimal solution of SPDP for \( T > \frac{1}{2m(\alpha, w)} \) will either be monotonically increasing or U-shaped in service levels. Let \( ((i_1^*, \ldots, i_n^*), t^*) \) be an optimal solution of SPDP such that \( x_{i_1^*} \geq \cdots \geq x_{i_k^*} \geq x_{i_{k+1}} \leq x_{i_{k+2}} \leq \cdots \leq x_{i_n^*} \), for \( 0 \leq k \leq n - 2 \). We construct a corresponding feasible solution for SPDP-U by assigning \( y_{i,j}^r = 0, t_{i,j}^r = 0, y_{i,j}^l = 1, t_{i,j}^l = t_{i,j}^r, \forall j = 1, \ldots, k \) and \( y_{i,j}^l = 0, t_{i,j}^l = 0, y_{i,j}^r = 1, t_{i,j}^r = t_{i,j}^*, \forall j = k + 1, \ldots, n \).

By construction, (5.1), (5.2), and (5.5) are satisfied. Since \( \sum_{j=1}^n t_{i,j}^* = T \) and \( \tau_{i,j} \leq t_{i,j}^* \leq \tau_{i,j}, \forall j, (5.3)-(5.4) \) are also satisfied. Hence, such a solution is feasible in SPDP-U.

We next show that the objective values are equal. The objective value for SPDP at the optimal point is given by \( S^{SPDP}(i_1^*, \ldots, i_n^*, t^*) = \sum_{j=1}^n (x_{i_j^*} - x_{i_{j-1}^*}) \Phi(t_{i_j}^* + \cdots + t_{i_n}^*) \)

\[ = \sum_{j=1}^{k+1} (x_{i_j^*} - x_{i_{j-1}^*}) \Phi(T - \sum_{p=1}^{j-1} t_{i_p}^*) + \sum_{j=k+2}^{n} (x_{i_j^*} - x_{i_{j-1}^*}) \Phi(\sum_{p=j}^{n} t_{i_p}^*) \]

\[ = \sum_{j=1}^{k+1} (x_{i_j^*} - x_{i_{j-1}^*}) \Phi(T - \sum_{p=1}^{j-1} t_{i_p}^*) + \sum_{j=k+2}^{n} (x_{i_j^*} - x_{i_{j-1}^*}) \Phi(\sum_{p=1}^{j} t_{i_p}^*) \]

\[ = \sum_{j=1}^{k+1} (x_{i_j^*} - x_{i_{j-1}^*}) \Phi(T - \sum_{p=1}^{j-1} t_{i_p}^*) + \sum_{j=k+2}^{n} (x_{i_j^*} - x_{i_{j-1}^*}) \Phi(\sum_{p=1}^{i_j} t_{i_p}^*) \]

\[ = \sum_{j=1}^{n} (x_j - x_{j-1}^*) \Phi(T - \sum_{p=1}^{j-1} t_{i_p}^*) + \sum_{j=1}^{n} (x_j - x_{j-1}) \Phi(\sum_{p=1}^{j} t_{i_p}^*) \]

\[ = \sum_{j=1}^{n} (x_j - x_{j-1}) \Phi(T - \sum_{p=1}^{j-1} t_{i_p}^*) + \sum_{j=1}^{n} (x_j - x_{j-1}) \Phi(\sum_{p=1}^{j} t_{i_p}^*) \]

(8.10)

(8.11)

(8.12)

(8.13)

(8.14)
Because of constraints (5.1) and (5.3), \( x_1 \geq \cdots \geq x_n \). Since, by assumption, all activities are allocated and no activity is allocated twice, i.e., (ii) \( \sum_{j=1}^{n} t_{ij} = 1 \) and set \( t_{ij} = t_{ij}^* \). We then assign \( i_{k+1} = n \) and set \( t_{ik+1} = y_{ik}^*t_{ik}^* + y_{ik}^*t_{ik}^* \). Finally, for \( j = k+2, \ldots, n \), let \( i_j = max\{p < i_{j-1} \} \) such that \( y_{ip}^* = 1 \) and set \( t_{ij} = t_{ij}^* \).

By equation (5.1), each activity is such that either \( y_{ij}^* = 1 \) or \( y_{ij}^* = 1 \). Hence, by construction, all activities are allocated and no activity is allocated twice, i.e., \( (i_1, \ldots, i_n) \) represents a sequence of activities. Since, \( x_1 \geq \cdots \geq x_n \), we have \( x_{i_1} \geq \cdots \geq x_{i_{k+1}} \leq x_{i_{k+2}} \leq \cdots \leq x_{i_n} \). Because of constraints (5.1) and (5.3), \( \tau_{ij} \leq t_{ij} \leq \tau_{ij}, \forall j \). Moreover by (5.4), \( \sum_{j=1}^{n} t_{ij} = 1 \).

Hence, this constructed solution is feasible for SPDP.

We next show that the objective value of SPDP-U is the same as SPDP at \((y^1, y^r, t^1, t^r)\). The objective value for SPDP-U at optimality is given by

\[
S^{SPDP-U}(y^1, y^r, t^1, t^r) = S^{SPDP}(i_1, \ldots, i_n, t).
\]
Lemma C-2 Without loss of generality, suppose that $x_1 > \cdots > x_n$ and $T > \frac{2}{m(\alpha, w)}$. For any $\lambda \geq 0$ only $O(n^2)$ points need to be considered to find an optimal solution for $UL(\lambda)$.

Proof Proof. Let $t^*$ be the optimal solution for $UL(\lambda)$. Hence, $t^*$ should satisfy the first order optimality conditions:

1. The stationarity condition gives $\frac{\partial UL(\lambda, t^*)}{\partial t_i} + \mu_i - \bar{\mu}_i = 0, \forall i$.
2. Complementary slackness gives $\bar{\mu}_i(t_i - \tau_i) = 0, \forall i$ and $\mu_i(t_i) = 0, \forall i$.
3. Primal feasibility implies $t^*$ satisfies the constraints (5.6).
4. Dual feasibility implies $\mu_i, \bar{\mu}_i \geq 0, \forall i$.

These conditions reduce to the following:

$$\frac{\partial UL(\lambda, t^*)}{\partial t_i} = \bar{\mu}_i \geq 0, \forall i \text{ such that } t^*_i = \tau_i$$

$$\frac{\partial UL(\lambda, t^*)}{\partial t_i} = 0, \forall i \text{ such that } 0 < t^*_i < \tau_i$$

$$\frac{\partial UL(\lambda, t^*)}{\partial t_i} = -\mu_i \leq 0, \forall i \text{ such that } t^*_i = 0. \quad (8.15)$$

We observe that for any two consecutive activities the following holds

$$\frac{\partial UL(\lambda, t^*)}{\partial t_{i-1}} - \frac{\partial UL(\lambda, t^*)}{\partial t_i} = -(x_i - x_{i-1})\Phi'(T - \sum_{p=1}^{i-1} t_p), \forall i = 2, \ldots, n. \quad (8.16)$$

Since $\frac{\partial UL(\lambda, t^*)}{\partial t_n} = -\lambda \leq 0$, we may take $t^*_n = 0$.

Suppose that $T - \sum_{p=1}^{n} t^*_p > \frac{1}{m(\alpha, w)}$. Then we have $\frac{\partial UL(\lambda, t^*)}{\partial t_1} < \frac{\partial UL(\lambda, t^*)}{\partial t_2} < \cdots < \frac{\partial UL(\lambda, t^*)}{\partial t_{n-1}} < 0$ by Lemma A-1 and condition (8.16). By (8.15), $t^* = (0, \ldots, 0)$ is an optimal solution.

Suppose next that there exists a $k = \min\{i : i \in \{2, \ldots, n+1\} \text{ such that } T - \sum_{p=1}^{i-1} t^*_p \leq \frac{1}{m(\alpha, w)}\}$. Therefore by (8.16),

$$\frac{\partial UL(\lambda, t^*)}{\partial t_{i-1}} < \frac{\partial UL(\lambda, t^*)}{\partial t_i}, \forall i < k$$

$$\frac{\partial UL(\lambda, t^*)}{\partial t_{i-1}} \geq \frac{\partial UL(\lambda, t^*)}{\partial t_i}, \forall i \geq k, \quad (8.17)$$

since $x_1 > \cdots > x_n$ and $\Phi(t)$ is strictly pseudoconcave and maximized at $t = \frac{1}{m(\alpha, w)}$ (Lemma A-1), and the latter inequalities are strict unless there exists an $l \geq k$ such that $T - \sum_{p=1}^{i-1} t^*_p = \frac{1}{m(\alpha, w)}$ for $i = \{k, \ldots, l\}$. Therefore, in case of strict inequalities there exists at most one
\( i < k - 1 \) and at most one \( i > l \) such that \( \frac{\partial UL(\lambda, t^*)}{\partial t_i} = 0 \). By (8.15) the optimal solution is of the form \( t^* = (0, \ldots, 0, t^*_l, \tau_{l+1}, \ldots, \tau_{r-1}, t^*_r, 0, \ldots, 0) \).

Suppose next that there exists an index \( j \geq k \) such that \( T - \sum_{p=1}^{i-1} t^*_p = \frac{1}{m(\alpha, w)} \) for \( i \in \{k, \ldots, j\} \) and \( T - \sum_{p=1}^{j} t^*_p < \frac{1}{m(\alpha, w)} \). Then, \( t^*_p = 0, p = k, \ldots, j - 1 \). We consider three cases.

(i) Suppose that \( \frac{\partial UL(\lambda, t^*)}{\partial t_{k-1}} > 0 \). Therefore, \( \frac{\partial UL(\lambda, t^*)}{\partial t_i} > 0 \) for \( i = k - 1, \ldots, j \) by (8.16). By (8.15), \( t^*_p = \tau_p, p = k - 1, \ldots, j \), a contradiction since \( t^*_p = 0, p = k, \ldots, j - 1 \).

(ii) Suppose that \( \frac{\partial UL(\lambda, t^*)}{\partial t_{k-1}} = 0 \). Therefore, \( \frac{\partial UL(\lambda, t^*)}{\partial t_i} = 0 \) for \( i = k - 1, \ldots, j \) by (8.16). By assumption \( t^*_p = 0 \) for \( p = k, \ldots, j - 1 \). Therefore by (8.17),

\[
\frac{\partial UL}{\partial t_{i-1}} < 0 \text{ for } i < k - 1; \quad \frac{\partial UL}{\partial t_i} < 0 \text{ for } i > j.
\]

By (8.15), the optimal solution is therefore of the form \( t^* = (0, \ldots, 0, t^*_{k-1}, 0, \ldots, 0, t^*_j, 0, \ldots, 0) \).

(iii) Suppose that \( \frac{\partial UL(\lambda, t^*)}{\partial t_{k-1}} < 0 \). Therefore, \( \frac{\partial UL(\lambda, t^*)}{\partial t_i} < 0 \) for \( i = k - 1, \ldots, j \) by (8.16). This implies by (8.17),

\[
\frac{\partial UL}{\partial t_{i-1}} < 0 \text{ for } i \leq k - 1; \quad \frac{\partial UL}{\partial t_i} < 0 \text{ for } i > k - 1.
\]

By (8.15), the optimal solution is therefore of the form \( t^* = (0, \ldots, 0) \).

In summary, the optimal solution of \( U(\lambda, t) \) can be characterized by two indices \((l, r)\) such that

\[
t^* = (0, \ldots, 0, t^*_l, \tau_{l+1}, \ldots, \tau_{r-1}, t^*_r, 0, \ldots, 0), \quad \text{or} \quad t^* = (0, \ldots, 0, t^*_j, 0, \ldots, 0, t^*_r, 0, \ldots, 0),
\]

(8.18)

where \( t^*_l \) and \( t^*_r \), are such that \( 0 \leq t^*_l \leq \tau_l \), \( 0 \leq t^*_r \leq \tau_r \) and \( T - \sum_{p=1}^{l} t^*_p \leq m(\alpha, w) \) with \( r \geq l \). We next claim that, for any such \( l \) and \( r \) if there exists a solution such that \( 0 < t^*_l < \tau_l \) and \( 0 < t^*_r < \tau_r \), then there exists only one such solution. In particular, \( t^*_l \) and \( t^*_r \) have to satisfy (8.15), which reduces to:

\[
f_1(t_l) = (x_{l+1} - x_l)\Phi'(T - t_l) + \cdots + (x_r - x_{r-1})\Phi'(T - t_l - \tau_{l+1} - \cdots - \tau_{r-1}) = 0
\]

\[
f_2(t_l, t_r) = (x_n - x_r)\Phi'(T - t_l - \tau_{l+1} - \cdots - \tau_{r-1} - t_r) - \lambda = 0.
\]
Solving for \( f(t_i) = 0 \) gives the closed form solution

\[
\dot{t}_i = T - \frac{1}{w - \alpha} \ln \left( \frac{\left( x_{i+1} - x_i \right) + \cdots + \left( x_r - x_{r-1} \right) e^{w(\sum_{p=1}^{r-1} T_p)}}{\left( x_{i+1} - x_i \right) + \cdots + \left( x_r - x_{r-1} \right) e^{\alpha(\sum_{p=1}^{r-1} T_p)}} \right).
\]

Hence, there exists at most one \( t_i, 0 < t_i < \tau_i \) solving \( f_1(t_i) = 0 \). Moreover, because \( r \geq l \), \( T - \sum_{p=1}^{r} t_p^* \leq T - \sum_{p=1}^{l} t_p^* \leq \frac{1}{m(\alpha, w)} \), and therefore \( \Phi''(T - \sum_{p=1}^{r} t_p^*) < 0 \) by Lemma A-3. Hence, \( f_2(t_i^*, t_r) \) is a strictly increasing function of \( t_r \) given that \( x_n < x_r \) and there can be therefore at most one \( t_r, 0 < t_r < \tau_r \) solving \( f_2(t_i^*, t_r) = 0 \).

Hence, for any \( l, r \), there exists only a constant number of potential optimal solutions of type (8.18) such that \( T - \sum_{p=1}^{r} t_p^* \leq \frac{1}{m(\alpha, w)} \) and \( r \geq l \). Since there are \( O(n^2) \) different ways of selecting \( l, r \), the number of candidate solutions is therefore of the order \( O(n^2) \).

**Lemma C-3** Without loss of generality, suppose that \( x_1 > \cdots > x_n \) and \( T > 2\frac{1}{m(\alpha, w)} \). For any \( \lambda \geq 0 \), only \( O(n) \) points need to be considered to find the optimal solution for \( UR(\lambda) \).

**Proof** Proof. Let \( t^* \) be the optimal solution for \( UR(\lambda) \). Hence, \( t^* \) should satisfy the first order optimality conditions:

1. The stationarity condition gives \( \frac{\partial UR(\lambda, t)}{\partial t_i} + \mu_i - \bar{\mu}_i = 0 \), \( \forall i \).
2. Complementary slackness gives \( \bar{\mu}_i (\tau_i - t_i^*) = 0 \), \( \forall i \) and \( \mu_i (t_i^*) = 0 \), \( \forall i \).
3. Primal feasibility implies \( t_i^* \) satisfies the constraints (5.7).
4. Dual feasibility implies \( \mu_i, \bar{\mu}_i \geq 0 \), \( \forall i \). These conditions reduce to the following:

\[
\frac{\partial UR(\lambda, t^*)}{\partial t_i} = \bar{\mu}_i \geq 0, \forall i \text{ such that } t_i^* = \tau_i
\]

\[
\frac{\partial UR(\lambda, t^*)}{\partial t_i} = 0, \forall i \text{ such that } 0 < t_i^* < \tau_i
\]

\[
\frac{\partial UR(\lambda, t^*)}{\partial t_i} = -\bar{\mu}_i \leq 0, \forall i \text{ such that } t_i^* = 0.
\]

Observe that for any two consecutive activities the following holds

\[
\frac{\partial UR(\lambda, t^*)}{\partial t_{i-1}} - \frac{\partial UR(\lambda, t^*)}{\partial t_i} = (x_{i-1} - x_i) \Phi'(\sum_{p=1}^{i-1} t_p), \forall i = 2, \ldots, n.
\]

Since \( \frac{\partial UL(\lambda, t^*)}{\partial t_n} = -\lambda \leq 0 \), we may set \( t_n^* = 0 \).

Suppose next that there exists a \( k = \max\{i : i \in \{2, \ldots, n + 1\} \text{ such that } \sum_{p=1}^{i-1} t_p^* \leq \)
Therefore
\[
\frac{\partial UR(\lambda, t^*)}{\partial t_{i-1}} \geq \frac{\partial UR(\lambda, t^*)}{\partial t_i}, \quad \forall i < k
\]
and the former inequalities are strict unless there exists an \( l \leq k \) such that \( \sum_{p=1}^{i-1} t_p^* = \frac{1}{m(\alpha, w)} \) for \( i = l + 1, \ldots, k \) since \( x_1 > \cdots > x_n \) and by Lemma A-1. If no such \( k \) exists, i.e., if \( t_1 > \frac{1}{m(\alpha, w)} \), then \( \Phi'(\sum_{p=1}^r t_p^*) < 0, \forall r = 1, \ldots, n \). Therefore, \( \frac{\partial UR(\lambda, t^*)}{\partial t_1} < \cdots < \frac{\partial UR(\lambda, t^*)}{\partial t_{n-1}} < 0 \) by (8.19) and since \( \frac{\partial UR(\lambda, t^*)}{\partial t_n} \leq 0 \). Hence, the first-order conditions (8.19) imply that the optimal solution is \( t^* = (0, \ldots, 0) \).

Suppose that there exists an \( l \leq k \) such that \( \sum_{p=1}^{i-1} t_p^* = \frac{1}{m(\alpha, w)} \) for \( i = l + 1, \ldots, k \) and \( \sum_{p=1}^{l-1} t_p^* < \frac{1}{m(\alpha, w)} \). By (8.20),
\[
\frac{\partial UR(\lambda, t^*)}{\partial t_{i-1}} > \frac{\partial UR(\lambda, t^*)}{\partial t_i}, \quad \forall i = 1, \ldots, l
\]
and
\[
\frac{\partial UR(\lambda, t^*)}{\partial t_{i-1}} < \frac{\partial UR(\lambda, t^*)}{\partial t_i}, \quad \forall i = k + 1, \ldots, n.
\]
In particular, \( \frac{\partial UR(\lambda, t^*)}{\partial t_k} < \cdots < \frac{\partial UR(\lambda, t^*)}{\partial t_n} \leq 0 \) and therefore, \( \frac{\partial UR(\lambda, t^*)}{\partial t_{i-1}} = \cdots = \frac{\partial UR(\lambda, t^*)}{\partial t_k} < 0 \).
By (8.19) this implies that \( t^*_1 = \cdots = t^*_l = 0 \). Since \( \frac{\partial UR(\lambda, t^*)}{\partial t_{i-1}} > \frac{\partial UR(\lambda, t^*)}{\partial t_i}, \forall i = 1, \ldots, l \), there is at most one \( i \) such that \( \frac{\partial UR(\lambda, t^*)}{\partial t_i} = 0 \).

Therefore, by (8.19), when there exists an index \( k \) such that \( \sum_{p=1}^{k-1} t_p^* \leq \frac{1}{m(\alpha, w)} \), the optimal solution has the form \( t^* = (\tau_1, \ldots, \tau_{r-1}, t_r^*, 0, \ldots, 0) \), with at most one activity \( t_r^* \) assigned a duration \( 0 < t_r^* < \tau_r \), in which \( r \leq k \), i.e., \( \sum_{i=1}^r t_i^* \leq \frac{1}{m(\alpha, w)} \). We next claim that for any \( r \) such that \( \sum_{i=1}^r t_i^* \leq \frac{1}{m(\alpha, w)} \), there exists at most one interior solution \( t_r, 0 < t_r < \tau_r \), such that \( \frac{\partial UR(\lambda, t^*)}{\partial t_r} = 0 \). In particular, \( t_r^* \) has to satisfy the first-order optimality conditions (8.19):
\[
f(t_r) = (x_r - x_n)\Phi'(\tau_{i_1} + \cdots + \tau_{r-1} + t_r) - \lambda = 0.
\]
Because \( \sum_{p=1}^r t_p^* \leq \sum_{p=1}^{k-1} t_p^* \leq \frac{1}{m(\alpha, w)} \), \( \Phi'(\sum_{p=1}^r t_p^*) < 0 \) by Lemma A-3. Moreover, \( x_r > x_n \) by assumption. Hence, \( f(t_r) \) is a strictly decreasing function of \( t_r \). Therefore, there can be at most one solution for \( f(t_r) = 0 \).

Since there are \( O(n) \) different ways of selecting \( r \), the number of candidate solutions is therefore of the order \( O(n) \).
Bibliography


