8-26-2019 10:30 AM

A Study in Three Practical Management Science Problems

John S.F. Lyons, *The University of Western Ontario*

Supervisor: Peter C. Bell, *The University of Western Ontario*
: Mehmet A. Begen, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Business

Abstract

This study of practical problems in Management Science (MS) describes novel mathematical models for three different decision settings. It addresses questions of: (a) what optimal route should be taken through a time-windows and topographically complex network; (b) what optimal sequencing of scheduled surgeries best coordinates flow of patients through central recovery; and (c) what prices should be charged and what stock amounts should be produced for two markets or channels to maximize profit explicitly, given various capacity and uncertainty conditions.

The first problem is in a sport analytics context, using a novel Integer Programming and big data from Whistler-Blackcomb ski resort. The second is to coordinate dozens of surgeries at London Health Sciences Centre, using a novel Constraint Programming model mapped to and parameterized with hospital data, including a tool for visualizing process and patient flow. The third problem is relevant to almost any business with a secondary market or sales channel, as it helps them identify profit optimal prices based on simple demand estimates and cost information they can easily provide for their own setting.

The studies use fundamentally different operational research techniques, in each case uniquely extended to the problem setting. The first two are combinatorial problems, neither one extremely beyond human cognitive ability, and both involving lots of uncertainty, and thus the sort of problem managers tend to dismiss as not efficient or practical to solve analytically. We show in the first study that vastly more skiers could achieve the challenge by following our route recommendation, unintuitive as are some of its elements, initially. In the second study, our scheduling model consistently outperforms currently unstructured-independent approach at the hospital. The final study is mathematical but demonstrates that by considering distinct market costs in pricing a firm can invariably earn more profit.

Keywords

route optimization, sports analytics, operating room scheduling, integer programming, dual channel pricing, constraint programming

Dedication

This thesis dissertation is dedicated first and foremost to my wife (Laura), for her unbounded love and support. Also, our children (Connor, Emily and Victoria), whose pride and encouragement, and own personal developments have been an incalculable inspiration. Finally, to my parents, both natural (Chuck, Kay) and in-law (Bud, Carol), for instilling in their families and hence mine, a most sincere appreciation for life-long and higher learning.

Summary for Lay Audience

Three chapters of this dissertation cover a variety of important methods in management science and related disciplines e.g. statistics, economics. The problems and results are intriguing without necessarily understanding their proofs.

Chapter 2 describes what most people, even regular skiers at Whistler-Blackcomb, would not imagine as the enormity of possible routes for the problem, or the ambiguity of whether one of them is 'best'. Many will be interested to see information that can be derived from simple time-stamp data collected from electronic tickets at lift stations, and the new technology-enabled opportunities and efforts being made to 'gamify' the sport.

Few cannot relate to the problem of waiting for surgery, nor take interest in rapid and standardized operating rooms being piloted to address the problem. Chapter 3 identifies how post-surgical recovery, a step in the process rarely considered by patients, can be a limiting factor to enabling faster, more voluminous patient flow. The chapter describes the coordination challenge involved, especially its high variability and uncertainty, but also scientific approaches that can better anticipate and manage the situation despite these factors.

The first two papers demonstrate conceptually similar but fundamentally different mathematical programming approaches. One uses binary decision variables (should a specific lift-to-lift transition be included in a route), and other uses a different type of data which are interval and sequence variables (where to position intervals of patient procedures and recoveries such that they fit together 'best' in time and space.)

Some basic understanding of microeconomics is helpful to appreciate, in Chapter 4, essentially how scientific pricing works, some reason to find the same product priced differently in two places, and especially how should it be priced differently. Several propositions can serve as a guide for pricing in one's own situation, including when capacity is limited, and/or where (different) market demand uncertainties warrant consideration.

An integral motivation for my Ph.D. journey has been to acquire understanding that will allow me to impart a greater awareness of, appreciation for, and interest in scientific management techniques among non-practitioners, and I hope that is reflected in this thesis.

Co-Authorship Statement

I declare that this thesis incorporates material that is a result of joint research. Chapter 2 “Solving the Whistler-Blackcomb Mega Day Challenge”, Chapter 3 “Elective Surgery Scheduling to Improve Perioperative Patient Flow” and Chapter 4 “The Effect of Revenue Versus Profit Maximization on Firm Profits” are co-authored with Dr. Peter C. Bell and Dr. Mehmet A. Begen. As the first author, I was in charge of all aspects of these projects including formulating research questions, literature review, model formulation, programming, and preparing the first and the following complete drafts of the manuscript. With the above exceptions, I certify that this dissertation and the research to which it refers, is fully a product of my own work. Overall, this dissertation includes 3 original papers.

Acknowledgments

The candidate wishes to gratefully acknowledge the support and assistance of his supervisors, Dr. Peter C. Bell and Dr. Mehmet A. Begen, for their candid and constructive criticism and challenging questions in response to which the papers were greatly improved. The candidate would also like to thank collaborators from the Whistler-Blackcomb ski resort and London Health Sciences Centre Perioperative Unit for their support of research on which the first and second articles in this thesis are based.

Table of Contents

List of Tables

List of Figures

Chapter 1

1 Introduction

1.1 Motivation

This three-paper thesis has been developed from my perspective of having spent two-and-a-half decades in management positions within both the public and private sector. It is purposefully broad and practically-oriented in keeping with similar features of the Management Science (MS) discipline. The subject areas chosen for this thesis cover three important classes of problems: two of them concerning manufacturing/service and logistic operations, routing and scheduling; and another concerning a critical marketing function, pricing. Within my practical experience, I have witnessed decision-making in all these areas, often following intuition that is expressed with the greatest conviction, but based on little of the scientific methods that are the domain of MS. This thesis, in its whole, serves as a brief but rich survey of several important MS methods, presented in the context of three relatable problems.

1.2 Overview

Chapter 2, the first article of this thesis entitled "*Solving the Whistler-Blackcomb Mega Day Challenge*"[1], addresses a routing problem of which tens of thousands additional skiers become aware every year. It is part of the ski resort's online ski-gaming community initiative called WB+. Those who consider undertaking the challenge, to ride all 24 lift systems spanning North America's largest ski area in a single day, quickly realize that a viable path is not easy to identify from a trail map. It is a combinatorial problem that is seemingly solvable by detailed inspection, but without great confidence, and subject to little margin for error to be successful, especially for less expert skiers (for whom it is likely most meaningful.) Working with WB+ management and large daily datasets of time-stamp skier ride information I developed and parameterized a mathematical (mixed linear-integer) program to determine an optimal route for a skier of the lowest percentile ability that could reasonably accomplish the feat. The approach is

[1] Lyons, J. S. F., P. C. Bell and M. A. Begen (2018). "Solving the Whistler-Blackcomb Mega Day Challenge." Interfaces 48(4): 323-339.

mathematically unique for application to time-windows constrained routing problems involving fixed and variable segment completion times/costs and an asymmetric graph of origin-destination connections.

Chapter 3, the second article of this thesis is entitled "*Elective Surgery Scheduling to Improve Perioperative Patient Flow*". It addresses a high priority problem for public health care systems that is to improve surgical throughput and reduce patient wait-times. This paper/chapter focuses on the daily operational aspect of OR scheduling and recovery. referred to collectively as perioperative care. I worked with a hospital which performs several dozen elective surgeries across 16-18 operating rooms (ORs) daily, in almost all cases requiring time for patient recovery in a central, limited capacity post-anesthesia care unit (PACU.) I obtained historical data to develop a model for predicting surgery and recovery times, and to parameterize a constraint programming (CP) model I also developed that coordinates sequences and timing of procedures across ORs to prevent surgery delays arising from PACU patient overload. Finally, I developed a tool for visualizing OR and PACU bed occupancies and patient statuses over the course of a day, both forecasted by the model and according to data collected during schedule execution, to assist management in identifying schedule-based sources of congestion, and thereby strategies to avert surgical delays.

Chapter 4, the third article of this thesis entitled "*The Effect of Revenue Versus Profit Maximization on Firm Profits*" explores a common problem of a firm simultaneously choosing prices and quantities of its product in distinct markets or channels with different demand characteristics. Although the field of revenue management literature is large, there has been a lack of attention paid to the nature and combination of costs between different sales and delivery channels, and potential impact of accounting for these costs or not in pricing decisions, with consequent profit outcomes. The paper begins with a simple deterministic unconstrained problem and set of solution decisions and outcomes. It then progressively develops easily comparable expressions regarding optimal decisions and outcomes for other forms of the problem (constrained and stochastic). These are presented to provide strong intuition for the differences between problem forms and between channel-specific decisions for profit maximization (PM) versus revenue maximization (RM). The paper includes a series of propositions and an algorithm for dealing with the most difficult stochastic, constrained problem form, the combination of which can serve as a useful aide-memoire for managers dealing with these types of problems.

The first two papers are practical in a literal sense, relating to real-world problems and data of the ski resort and hospital for whom we worked on the research. The third paper is practical in the different sense of having broad utility, providing a fresh view of an old subject, and raises several interesting considerations for managers facing the problem of dual channel pricing and inventory planning.

Chapter 5 concludes the thesis.

Chapter 2

2 Solving the Whistler-Blackcomb Mega Day Challenge

2.1 Abstract

The Whistler-Blackcomb (WB) Mega Day Challenge requires a skier to ride all 24 lifts at the resort in a single day. Among over two million skiers annually at WB, only 313 completed the challenge in fourteen months following the introduction of a system that tracks lift use by skier. Apart from the physical challenge, the difficulty is to find a route that matches one's skill level while accounting for variable lift opening and closing times.

We use data from WB's radio-frequency identification (RFID) ticketing system to estimate ski times between lifts for skiers of various skill levels. We then formulate and solve the problem by a combined, iterative integer programming and heuristic approach, up to the highest feasible skier skill level. The problem's distinctive features preclude use of known solution methods for similar problems, so we use a practical, staged solution approach.

Our results include a recommended route that enables the greatest number of skiers, roughly the fastest quartile, to achieve the challenge. We also provide a benchmark, that skiers who can ski a particular common run in 12 minutes or less, should be able to complete the challenge. In three months following communication of our recommended solution, the rate at which Mega Days were successfully completed increased by two-thirds from the previous seven skiing months.

2.2 Introduction

Whistler-Blackcomb (WB) ski area located in British Columbia, Canada, hosted the 2010 Winter Olympics alpine events and is one of North America's largest ski resorts. WB spans more than 12 square miles across two mountains (Whistler and Blackcomb) each with more than 4500 vertical feet of lift ski access.

In 2015 WB implemented an RFID ticketing system that enabled the resort to track movements of skiers, frequently numbering more than 15,000 per day, as they board 24 lift systems via 27 distinct access points scattered throughout the resort. This new RFID system offers management benefits such as increased revenues through reduced ticket fraud and

increased speed of service along with some reduction in staff costs. It also provides WB skiers with access to a web-based portal called WB+, where they can view their personal statistics (e.g. number of rides and vertical metres/feet accumulated).

Whistler-Blackcomb was acquired by Vail Ski Resorts in 2017. WB+ follows another similar system in North America operated across several properties of Vail Ski Resorts, called EpicMix©, that has been described as the 'gamification' of skiing (Khan 2010, Sean 2013). Detailed data on individual skiers enables Whistler-Blackcomb and Vail Ski Resorts to implement a motivational program where skiers can earn a variety of 'badges' (called 'pins' in EpicMix) based on their performance. For example, the "Mount Everest" badge is earned by riding a number of lifts in a single day such that the sum of their vertical rises exceeds 29,029 feet. As in this example, some badges are based on total elevation and do not require riding specific lifts, whereas others such as the "Whistler Complete" or "Blackcomb Complete" badges are earned by riding all RFID-enabled lifts on one mountain or the other in the course of a single day.

Among the most challenging is the "Mega Day" badge that is earned by a skier who, on a single day, rides all 24 lift systems on both mountains (12 on Whistler, 11 on Blackcomb, plus the 'Peak-2-Peak' lift that spans the adjoining valley).

Data made available to us showed that only a very small number (~0.1%) of skiers earn the Mega Day badge on any given day and so we set out to help WB management to improve the marketability of the WB award program by highlighting the Mega Day Challenge as the pinnacle of this program. An important part of this effort was to demonstrate that this accomplishment is not just for expert skiers, but can be earned by skiers of modest ability if they follow a route that suits their skill level.

Our aim was to find minimum-time routes for skiers of varying ability, with emphasis on finding a route that would enable the greatest number of skiers to earn their WB Mega Day badge, that is, the minimum-time route for the least capable skier among those realistically able to complete the challenge. There are generally multiple trails that a skier can take from lift to lift, and the matter of which trail is best in terms of speed and navigability may differ between skiers. More advanced skiers may follow steep trails that are shorter and faster for them, but that are difficult and ultimately slower (effectively longer) for less capable skiers.

The former group have more route flexibility, as a less advanced skier's feasible route is always feasible to the more advanced skier (although it may not be the latter's shortest time route for their ability and pace).

Our approach was to solve the problem starting with an advanced skier's ability, which we characterize as a 1st percentile skier, and then solve for increasing skier percentiles or decreasing ability. At some percentile, which turned out to be the 28th on our scale, a skier can expect to complete the challenge only by following a specific route, as there is little or no spare time to do otherwise.

Our results show that the WB Mega Day Badge challenge is achievable by the top quartile of typical skiers at Whistler-Blackcomb, although a much smaller fraction have done so to-date. In three months following electronic newsletter communication of our recommended solution to skiers at WB, the rate at which Mega Day Challenges were successfully completed increased 67%. Our findings also suggest that if a skier can ski a common run from Roundhouse lodge to the Whistler Village base in 12 minutes or less, they are capable of completing the Mega Day Challenge. Finally, we found that if WB were to keep the Fitzsimmons lift open an extra hour (even if only on weekends), it could increase the accessibility of the Mega Day badge to a wider range of skiing abilities.

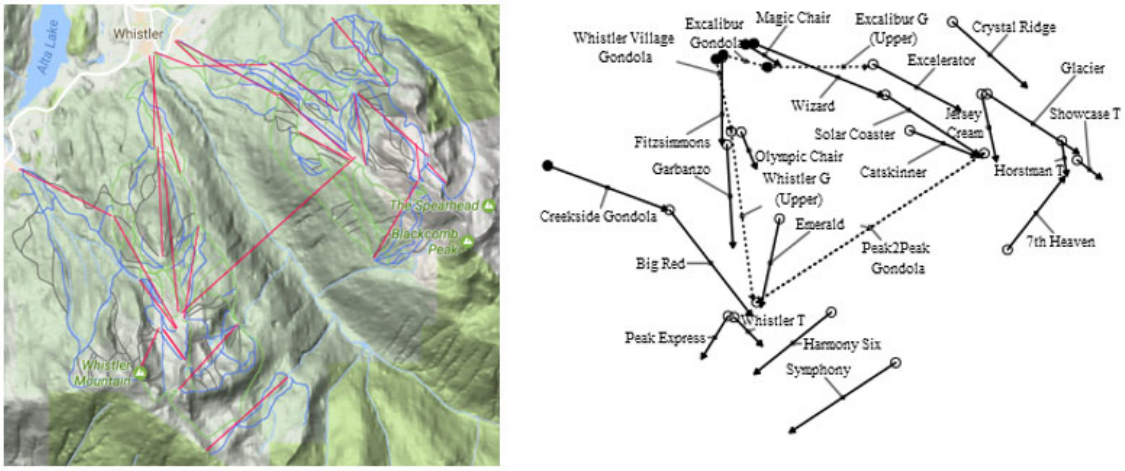
In the next section, we define the problem and point out the relationships between this Mega Day routing problem and other similar problems in the literature. We then provide a mathematical formulation of our model, relating its components to our staged solution approach. We discuss the RFID scan data made available to us, and how it has been used to parameterize the model. Finally, we present our computational results and discuss how they have been used by WB management.

2.3 The "Mega Day" Challenge

A skier earns the Mega Day badge by having his or her lift ticket scanned at each of the 24 RFID-enabled lift systems that make up the WB ski area in a single day. Typically, scanners are located at the bottom entrance of the lift system, so a route that visits the bottom of all 24 lifts would suffice. While this seems like a straightforward routing problem, several features make this variant unique.

First, the resort includes three “lifts” that are in fact “lift systems” which include more than one location where a rider can get on or off. One of these lifts is the ‘Peak-2-Peak’ gondola, spanning the valley between Whistler and Blackcomb mountains that can be ridden in either direction. The other two are the Excalibur and Whistler Village gondolas that are unidirectional but have mid-stations that can serve as alternate entry-exit points.

Figure 1 depicts a North-facing map of the Whistler-Blackcomb area on the left and a schematic of its 24 lift systems on the right. The dotted lines represent gondolas with multiple and/or bi-directional segments. A skier must begin at one of seven potential starting points along the two mountain bases, indicated by solid black circles (upper left portion of the image). These lifts are the Creekside Gondola, Whistler Village Gondola, Fitzsimmons Chair, Excalibur Gondola base, Magic Chair, Wizard Chair, and Excalibur Gondola mid-station (accessible from an upper level parking area).



Source: www.googlemaps.com

Figure 2-1 Map of Whistler-Blackcomb & Schematic of Lift System

In Figure 2, lift entry points are represented as circles, with seven possible starting points indicated by black fill, and lifts with multi-segment shown as dotted lines

We consider a *route* to be a sequence of lift segments ridden that meets the requirements of the Mega Day Challenge. A route does not necessarily involve riding every segment of every lift and may include riding some lift segments more than once.

The seven possible starting lifts are spread over five kilometres, and open at slightly staggered times. Upper level and back valley lifts open progressively later in the morning, only after a

sufficiegnt number of skiers are in a position to reach these lifts. These remote lifts also close earlier in the afternoon, in order to move skiers down toward the base of the resort and allow ski patrols to complete sweeps before dusk.

A skier must check into his or her final lift before it closes for the day. Neither the ride time on that final lift ride nor the following ski-out have any bearing on completion of the challenge.

In combination, these features make the problem unique among routing problems, and difficult to optimize. A key source of this difficulty is that we cannot impose hard time window constraints on all lift segments, because we don't know in advance if they will be part of the chosen route. Also, the natural optimization objective of minimizing time gives preference to routes whose final lifts are most remote because the subsequent, long ski-out time is not counted in the objective calculation. On the other hand, early closing times of remote lifts make them poor choices as final lifts for slower (higher percentile) skiers, the very ones for whom solving the problem is most important. For these reasons we have taken a practical approach that solves the problem in three stages.

The objective of our work was to create interest in the Mega Day Challenge among skiers by providing recommendations and guidelines on the routes they should choose according to their ability. In particular, we sought to identify the route that is feasible for the greatest number of skiers, from the most 'expert' down to some skill level below which a skier cannot reasonably expect to be able to complete the challenge.

2.4 RFID Ticket System Information

We obtained a full day's worth of data for Thursday, 25 February 2016 from the WB IT Business Support Team to develop and initially validate our model. We later obtained three additional days of RFID scan data to refine our model parameter estimates. A summary of the scan data sets is shown in Table 1.

Date (2016)	19 Feb	25 Feb	29 Mar	7 Apr
Unique RFID passes	19,591	15,417	17,152	11,665
Regular skiers	17,481	13,210	15,027	9,999
Total # Scans	158,759	144,185	144,661	94,400
Regular skier scans	145,823	128,689	130,609	83,641
Mega Day (MD) badges	1	14	16	11
MD routes w/ minimum lift	0	1	4	1

Table 2-1 Four Day RFID Lift Scan Data Summary

Unique RFID passes represent the number of individuals riding at least one lift on each date. Regular skiers *excludes* RFID passes used for only one ride (mostly employees ascending to work at the upper mountain lodges) as well as a small number of contractors and volunteers such as law enforcement, who may have shared the RFID pass among different skiers. Total # Scans represent all lift rides taken on each date, whereas Regular skier scans includes only rides taken by Regular skiers. The number of skiers who earned a Mega Day badge on each date is recorded. The final row in the table gives the number of distinct Mega Day routes taken on the date, which involved only the fewest possible number (24) of rides.

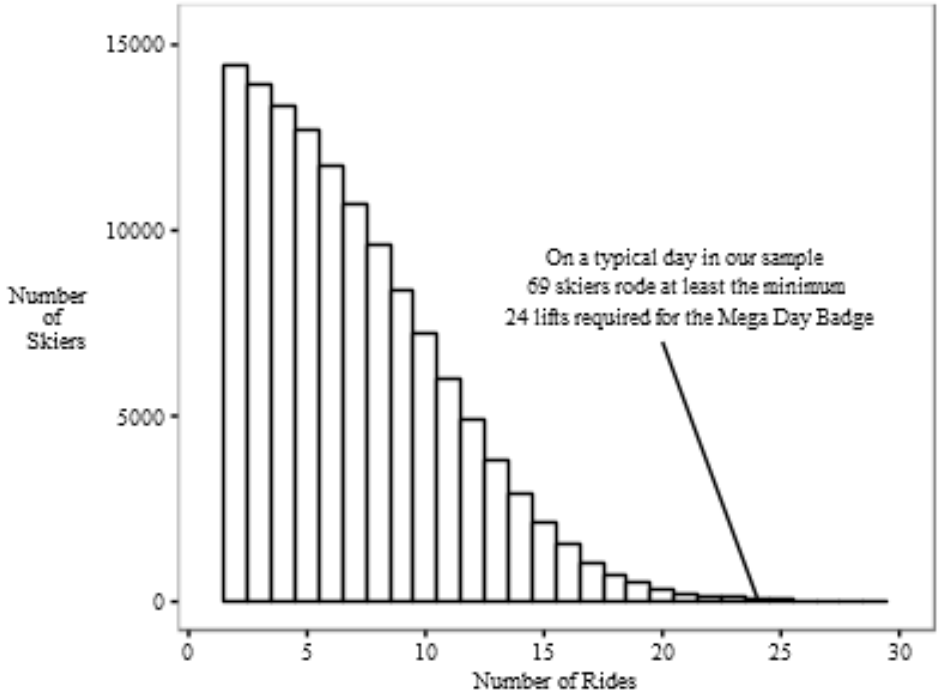


Figure 2-2 Distribution of Rides Per Skier on 25 Feb 2016

From our initial dataset, we determined that among 13,210 regular skiers, only 69 (~0.5%) of them rode 24 or more lifts. Most of these 69 skiers did not earn a Mega Day badge, as they rode fewer than 24 *distinct* lifts, rather multiple rides on some lifts and zero on others.

We determined the existence of *at least* one Mega Day route involving only the minimum 24 rides, by identifying three skiers who completed the Mega Day challenge together in this manner on 25 February 2016. The group started at 9:42 am and checked into their final lift at 15:01 pm. This duration of 319 minutes is roughly three-quarters of the overall time available from first lift opening (8:15 am) to the resort's general closing time (15:45 pm), after which only one small base lift (Magic Chair) remains open, typically for an additional 60-90 minutes.

This evidence suggests that the Mega Day is quite challenging on one hand, but should be achievable by reasonably advanced skiers, and not exclusively expert skiers, given that this group on 25 Feb achieved it with more than two hours of open resort time to spare. A key question we sought to answer is how advanced must be one's skiing ability in order to have a reasonable chance of earning the Mega Day badge. The closer a particular skier's ability is to that threshold, the more important will be that skier's choice of which route to navigate through the mountain lift network. A feasible route for this marginal Mega Day skier will be feasible for any more advanced skier.

2.5 Background Literature and Related Research

Routing problems have a long history in operations research extending back to 1736 when Leonhard Euler laid the foundations of graph theory, a commonly-used technique to model such problems. (Biggs *et al*. 1976). From the earliest days of electronic computing, the well-known Traveling Salesman Problem (TSP) and its variants have served as benchmarks of combinatorial complexity and standards by which solution computational approaches, heuristics and algorithms, are often compared (Cook 2012). In the early 1950s, a team from The Rand Corporation (Dantzig *et al*. 1954) developed an approach to large-scale instances of these problems that would be described five decades later as "the Big Bang" that "all successful TSP solvers echo" (Jünger *et al*. 2010). In their case, the specific problem was to find and prove a shortest travel distance route passing through Washington, DC and 48 capitals of the lower-mainland U.S. states.

Over the years, many variations of the problem have been proposed and many more approaches have been employed to solve them. Entirely new classes have developed, Vehicle Routing Problems (Eksioglu *et al*. 2009) being among the broadest of them. We briefly identify a few related problem types that share critical characteristics with the Mega Day Challenge.

The Orienteering Problem, also known as the Selective Traveling Salesman Problem (Vansteenwegen *et al*. 2011), has the objective of finding a route through a network of checkpoints, each of which has a certain score, where not all checkpoints must or can be visited within a given time frame. The Mega Day Challenge is also selective, as it allows for only a subset of lift segments to be visited. But its objective is binary (successful completion or not) and the rewards from visiting different lift segments have indistinguishable bearing on the decision of which lift segments to visit.

The lifts network at Whistler-Blackcomb includes only 285 feasible transitions (excluding same lift returns), of which 198 comprise 99 bidirectional connections. Only one pair of lift segments has *both* feasible connections *and* equal transition times in both directions. This makes our problem one of a broad class of Asymmetric Traveling Salesman Problems (Öncan *et al*. 2009).

Many routing problem variants involve time windows, and/or service times at the destinations (Kantor and Rosenwein 1992, Focacci *et al*. 2002, Campbell *et al*. 2011, Tas *et al*. 2016). Various approaches have been developed for the former, including the use of time buckets (Dash *et al*. 2012), adding variables for each destination's arrival time and constraining them to fall within that destination's time window (Desrosiers *et al*. 1995), and by constraint logic programming (Pesant *et al*. 1998). However, these approaches were developed for problems in which all destinations must be visited whereas with respect to the second of these approaches for example, an unvisited lift segment has no arrival time in the Mega Day problem and so cannot be constrained to lie within a specified time interval. Ride versus ski times in the Mega Day problem could be seen as analogous to service versus travel times, since we consider entire lift segments as nodes, from entrance to exits with fixed ride times in between. However, bundling ride and ski times in our model allows us to reduce by half the number nodes to define the problem.

Another related sub-class of routing problem is the Steiner Traveling Salesman Problem (Letchford *et al*. 2013) and its variants, including with time windows. Their important distinguishing features relevant to our problem are: (a) that only a subset of nodes must be visited; (b) that nodes may be visited more than once; and (c) that edges between nodes may be traversed more than once. Comparison to the Mega Day problem is deceptive, however. Rather than having required and optional nodes, we have groups of lift segments requiring that one (or more) must be chosen from each group. A minimum time route could conceivably visit the same lift more than once but is much less likely to be followed by the same subsequent lift.

While there exists a large variety of closely-related problem types and solution approaches in the literature, we have not found among them any quite like the Mega Day problem. Its uniqueness is due to the *combination of* the following features:

1. The network on which it is defined is clearly incomplete and highly asymmetric.

2. Route feasibility and time minimization depend not only on which $i \rightarrow j$ lift transitions are chosen, but also the *order* in which lifts are visited.

3. The problem includes subsets of lift segments where only one segment needs to be visited (although more may be visited.)

4. There is a subset of lifts which are possible starting points.

5. The transition times from lift-to-lift are a combination of fixed lift ride times and variable ski times, the latter being a function of skier ability.

6. Like the Steiner TSP, the number of times that a lift may be visited is integer, not binary.

In addition to the related operations research literature discussed above, we identified a small body of research related to skier abilities and trail selection. Skier abilities have been characterized according to their linear velocities as measured by radar gun (Shealy *et al*. 2005) Our analysis of skier abilities is based on vertical speeds of descent. Graph theory networks have been used to model the flow of skiers, particularly to analyze the cascading impacts of trail or lift closures on skier volumes and resulting queues (De Biagi *et al*. 2013).

In that research, skiers have been segmented into three levels of ability which imply their choice between easy, intermediate and difficult trails. Our research doesn't consider skier volumes or queueing explicitly, although the latter is accommodated in our time estimates (mean queue times are embedded within mean ski times). Our application of network and graph theory is to a route optimization problem for individual skiers rather than for modeling aggregate skier flow.

2.6 Methodology

We conducted our study in two *phases*. The first included data preparation and analysis, which provided input to the second phase where instances of the problem were generated and solved for different skiers. The second phase utilized a three-*stage* solution procedure for each instance.

In the first phase, RFID scan data from the WB+ system was cleaned and shaped. We accounted for and removed exit-scans to determine what lift *rides* began at what times, per skier. We then gathered data about lifts, including lower and upper locations (latitude, longitude, altitude) and ride times. We translated pairs of successive rides into *runs*, i.e., transitions from lifts i to lifts j. We identified and removed infeasible observations (< 2%) stemming from scans that were occasionally missed between two lifts that have no direct interconnection. We disaggregated the ride and ski times, making adjustments in cases where multi-segment lifts were exited at mid-stations. Finally, we derived $i \rightarrow j$ ski times according to skier abilities, from 1st to 100th percentiles, representing fastest to slowest skiers, respectively.

Our second phase was to develop and execute an optimization model. A series of parameter-data files were derived from phase 1 output for different skier abilities. The program was executed in order of increasing skier ability. Feasible solutions to the Mega Day routing problem were obtained for skier abilities from the 1st up to and including the 28th percentile. These solutions were generated in a three-stage procedure, the first two stages being common to many integer programming approaches for solving routing problems:

Stage 1 Solve the integer program (IP) described by equations (1)-(5) below. Then determine whether the solution contains any subtours. If not, proceed to stage 3 with the full tour candidate solution.

Stage 2 Add subtour elimination constraint equations (6) as required, re-solving the IP and again determining whether the revised solution contains any new subtours. If so, repeat stage 2.

Stage 3 Determine whether the full tour candidate solution from stage 2 satisfies time windows constraint equations (7). If necessary and possible, adjust start time and/or accept delays imposed by arrivals in advance of lift opening times. If any lift closing times are violated, within the initial solution or as a result of time adjustments, return to stage 2 with a new constraint equation (6) that precludes this full tour candidate solution, and force instead a search for the next most optimal solution in terms of the objective function (1).

A final solution for a skier percentile, generated by Stage 1, is then validated as feasible by Stages 2 and 3, perhaps modified with delays in Stage 3. We nevertheless refer to this as the 'shortest time route' for the skier level, because it is based on minimization of the objective function in Stage 1, regardless of time adjustments in Stage 3, if any. (This is shown visually in the Model Results section Figure 8.)

2.7 Model Formulation

Our mathematical formulation is presented in three stages that coincide with the solution procedure.

2.7.1 Integer Program (Stage 1)

Twenty-seven discrete ski lift segments are each mapped to one of 24 lift systems. The former are required to specify which lift-to-lift connection points are chosen, and to calculate the objective function value associated with those choices, whether they comprise a full tour or multiple subtours. The mapping is required to verify whether all lift systems are represented among the origins and/or destinations of the chosen connections.

We define the following sets, parameters and variables:

N a set of discrete lift segments (generally referred to simply as 'lifts').

G a set of lift systems (groupings of one or more lift segments).

B a subset of lift segments $B \subseteq N$ that constitute feasible starting lifts.

g_m a subset of lift segments belonging to the same lift system $g_m \subset N, m \in G$.

T_i the standard *ride time* (fixed) from the base to the top of a lift segment $i \in N$.

We note that lifts may operate at various speeds in reality, depending on load and other factors. But we assume, for simplicity, fixed lift ride times as suggested by Whistler-Blackcomb personnel.

τ_{ij}^k the *ski time* from the top of lift *i* to the base of lift *j* for a skier of skill level *k*.

t_{ij}^k the total *transit time* from the base of lift *i* to the base of lift *j* for a skier of skill level *k.*

Note that $t_{ij}^k = T_i + \tau_{ij}^k$.

For contiguous lift segments i, j belonging to the same lift system g_m, we specify the ski-time $\tau_{ij}^k = 0$, reflecting the fact that a skier merely needs to remain on-board to ride the second segment.

Recognizing that transition times are, in general, specific to the skier's skill level, we drop the skill level superscript k from this point forward.

We define:

$\hat{N}$ the union of the set N with a "dummy lift" $i = 0$, that is, $\hat{N} = N \cup \{0\}$

t_{0j} the transition time assigned to a skier for travel from the initial "dummy" lift to the entrance of any possible starting lift in B. We assign $t_{0j} = 0, \forall j \in B, t_{0j} = \infty$ otherwise.

t_{i0} the transition time assigned to a skier returning from lift *i* to the "dummy", $t_{i0} = 0, \forall i$.

X_{ij} the decision variables, where $X_{ij} = 1$ if the route for a skier includes skiing (or connecting within the same lift system) from top of *i* to bottom of *j*, otherwise $X_{ij} = 0$. We also set $X_{ij} = 0$ for infeasible $i \to j$ transitions.

S_j the number of times a skier visits (rides) a lift segment j, that is, $S_j = \sum_{i \in \hat{N}} X_{ij}$

The objective is to find the set of decision variables X_{ij} that minimize the time to visit all lift systems at least once:

$$\textit{Minimize} \qquad \sum_{i \in \hat{N}} \sum_{j \in \hat{N}} X_{ij} t_{ij} \tag{1}$$

subject to

$$\sum_{j \in g_m} S_j = \sum_{j \in g_m} \sum_{i \in \hat{N}} X_{ij} \geq 1 \qquad \forall\, m \in G \tag{2}$$

$$\sum_{i \in \hat{N}} X_{ik} = \sum_{j \in \hat{N}} X_{kj} \qquad \forall\ k \in \hat{N} \tag{3}$$

$$\sum_{j \in N} X_{0j} = \sum_{i \in N} X_{i0} = 1 \tag{4}$$

X_{ij} binary, S_j integer (5)

All components of the Objective function (1) that involve the dummy lift (either $i = 0$ or $j = 0$) evaluate to zero. Consequently, neither a skier's final ride nor his or her final ski add any cost (time) to the objective function value.

(2) allow solutions to use a given lift segment more than once but specify that, for all lift systems g_m, the sum of the visits to its member lift segments $j \in g_m$ must be at least one.

(3) are for flow conservation. The number of arrivals flowing into any lift in $\hat{N}$, including the dummy, must be the same as the number of departures flowing from that lift.

(4) ensures that there is only one connection from the dummy lift 0 to (one of seven practical starting lifts in) the real lift network N, as well as only one connection from the real lift network back to the dummy lift.

(5) specify that the choice of whether a solution includes travel from i to j is binary, whereas the number of times each lift system is ridden, along any of its segments, must be an integer equal to or greater than one.

2.7.2 Subtour Elimination (Stage 2)

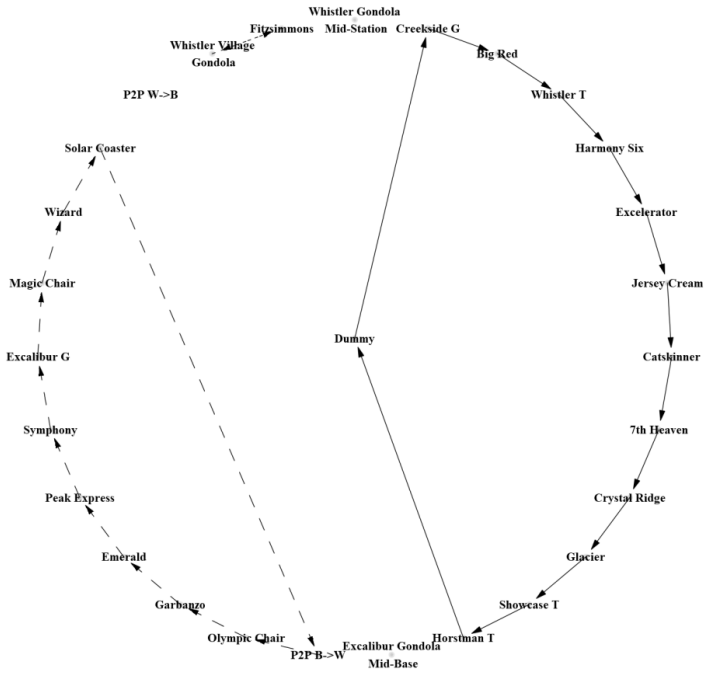


Figure 2-3 Example Subtours in Candidate Solution

Optimal solutions to equation (1)-(5) may lead to the problem of subtours, for which we need to add instances of equation (6) shown below. Figure 3 depicts a preliminary solution with three subtours, the smallest of which simply joins the Whistler Village Gondola to the Fitzsimmons chair and back again (shown just to the left of the image's top-centre.) Another subtour proceeds from the Dummy lift to Creekside Gondola and eventually returns to the Dummy lift after a visit to the Horstman T-bar. The underlying X_{ij} choices are optimal with respect to (1)-(5) but do not form a meaningful route in reality.

Following the DFJ formulation (Dantzig *et al*. 1954), we use 'lazy' constraints to eliminate subtours. We first define:

Z a set of edges (lift-to-lift transitions) comprising a subtour which we wish to eliminate.

We then add constraints:

$$\sum_{z \in Z} X_{I(z)J(z)} \leq |Z| - 1, \quad \forall\, Z \tag{6}$$

Constraints (6) ensure that at least one of the connections in the subtour cannot exist if all the others do, that is, to require a solution that has at least one connection into and one connection out of the subset Z, respectively from and to the complementary set of lifts N\Z.

2.7.3 Time Validation, Route Adjustment/Rejection (Stage 3)

After obtaining a full tour candidate solution to equations (1)-(6) as output from stage 2, we check to see if the solution satisfies the time window restrictions of each successive lift in the tour.

We define the following sets, functions and variables:

W an *ordered* set of edges $i \to j$ chosen within a full tour candidate solution, where we denote the l^{th} element of the set as W_l.

$I(w)$ a function which returns the origin lift $i \in \hat{N}$ from an edge $w \in W$.

$J(w)$ a function which returns the destination lift $j \in \hat{N}$ from an edge $w \in W$.

$T^w_{I(w)}$ the time at which a skier is expected to *depart* from an origin lift *I(w).*

$T^w_{J(w)}$ the time at which a skier is expected to *arrive* at a destination lift *J(w).*

$O^w_{I(w)}$ the *opening* time for an *origin* lift $I(w)$ of an edge w.

$O^w_{J(w)}$ the *opening* time for a *destination* lift $J(w)$ of an edge w.

$C^w_{J(w)}$ the *closing* time for a destination lift $J(w)$.

Noting that $T^w_{J(w)} = T^w_{I(w)} + t_{I(w)J(w)}$ and $O^w_{J(w-1)} = O^w_{I(w)}$, we calculate the starting time of stage w, which begins at lift $I(w) = J(w-1)$, as follows:

$$T^w_{I(w)} = \begin{Bmatrix} O^w_{J(w)} & \text{for } w = W_1 \\ Max\left\{T^{w-1}_{J(w-1)}, O^w_{I(w)}\right\} & \text{for } w = W_l \ (\forall l \neq 1) \end{Bmatrix}$$

The starting time for the first stage is the opening time for the first destination lift (departing from the dummy lift). For the remaining stages, the starting time is the greater of the arrival time from the previous stage, and the opening time of the origin lift in the current stage. The difference, if any, represents an amount by which the start of leg w is delayed. We add the constraints:

$$T^w_{J(w)} \leq C_{J(w)}, \quad \forall\ w \in W \qquad (7)$$

Constraints (7) specify that each ordered transition w must be completed before the closing time of its destination lift.

Table 2 provides example output from stages 1 and 2 of our solution procedure. This is a minimum time solution to equations (1)-(5) plus the set of subtour constraints (6) needed to obtain a full tour. However, we find in stage 3 that this candidate solution critically violates one or more time window constraint equations (7). Whereas the skier's expected early arrivals at the 2nd and 3rd lifts may be avoided simply by starting the tour later, late arrivals beginning at the 15th lift make this candidate solution infeasible with respect to time windows. Thus, we add a constraint (6) to eliminate this full tour and seek the next best candidate solution to the updated set of equations (1)-(6).

Order w	Destination Lift J(w)	Minutes Delay	HH:MM Depart	Minutes Ride & Ski Time	Comment
1st	8120 Creekside G	-	8:15 am	11.41	
2nd	8130 Big Red	3.59	8:30 am	13.84	
3rd	8196 Whistler T	91.16	10:15 am	14.9	
4th	8180 Harmony Six	-	10:29 am	24.88	
5th	8210 Excelerator	-	10:53 am	10.16	
6th	235 7th Heaven	-	11:03 am	21.29	
7th	260 Catskinner	-	11:24 am	12.66	
8th	820 P2P B->W	-	11:36 am	27.28	
9th	145 Olympic Chair	-	12:03 pm	11.97	
10th	8140 Garbanzo	-	12:14 pm	15.38	
11th	8150 Emerald	-	12:29 pm	11.36	
12th	8190 Peak Express	-	12:40 pm	17.39	
13th	8195 Symphony	-	12:57 pm	32.33	
14th	8100 Whistler Village G	-	13:29 pm	14.54	
15th	8110 Fitzsimmons	-	13:43 pm	15.53	Lift Closed
16th	205 Excalibur G	-	13:58 pm	7.54	-
17th	234 Magic Chair	-	14:05 pm	9.07	-
18th	8200 Wizard	-	14:14 pm	12	-
19th	220 Solar Coaster	-	14:26 pm	19.89	-
20th	245 Crystal Ridge	-	14:45 pm	14.41	-
21st	8250 Jersey Cream	-	14:59 pm	13.86	-
22nd	240 Glacier	-	15:12 pm	10.59	Lift Closed
23rd	8165 Showcase T	-	15:22 pm	10.56	Lift Closed
24th	8255 Horstman T	-	15:32 pm	-	Lift Closed

Table 2-2 Ordered Full-Tour Solution Infeasible Due to Time Windows

2.8 Parameterizing the Model

2.8.1 Lift Network

Table A1 in the appendix provides a list of the $|N| = 27$ lift segments at WB, along with the elevations of their entry and exit locations, and their standard ride times T_i. To estimate transition times t_{ij} between lifts, we used time intervals between RFID scans collected at lift

entrances. Our methods were first developed using data from 25 February 2016, then applied to four full days of scan data (see Table 1).

2.8.2 Transitions Times

After adjusting for exit scans and removing anomalies from some missing intermediate scans, we obtained sets of observed $i \rightarrow j$ transition times for various lift pairs. Figure 4 shows boxplots of example transition times observed on a particular day. The set at left are times recorded from 14 different origin lifts leading *to the Wizard Chair 8200*. On the right are times recorded *from the Symphony Chair* 8195 to seven different destination lifts recorded on the day.

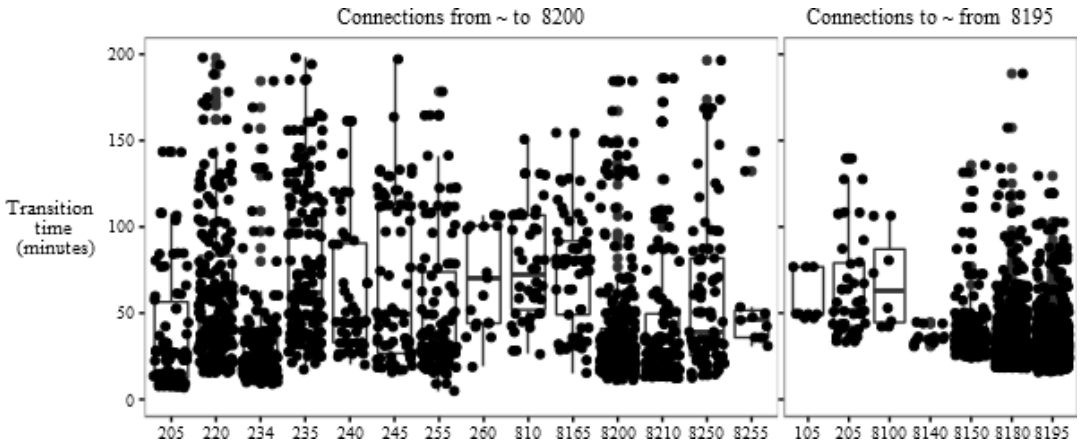


Figure 2-4 Transition Times To Wizard Chair and From Symphony Chair

We eliminated observations of transition times greater than 60 minutes, on the premise that skiers likely stopped for refreshment at some point during those runs.

Our model did not consider lift queues and wait times explicitly, as we had no basis on which to disaggregate them from total transition times. RFID scans were only captured at points just prior to lift boarding and provided no record of any preceding queue. Some extra time to reflect expected delays is embedded in each t_{ij}^{k} parameter estimate by our approach to their derivation described below.

Some feasible lift-to-lift transitions had no observations in our data sets. For this reason, two feasible destinations from 8195 are missing from Figure 4. On the other hand, we see two destinations (including a return to lift 8195 itself) that had well over 2000 observations per day to derive meaningful percentile ski times.

The left side of Figure 5 provides a different view of transitions from 8195. It depicts percentile ski times (RFID scan-to-scan interval minus fixed ride time) from 8195 to its various recorded destination lifts. The 10th and 30th percentile observations are marked as circles and triangles, respectively. Five of these pairs are highlighted again in the right side of Figure 5, as explained in the following paragraph.

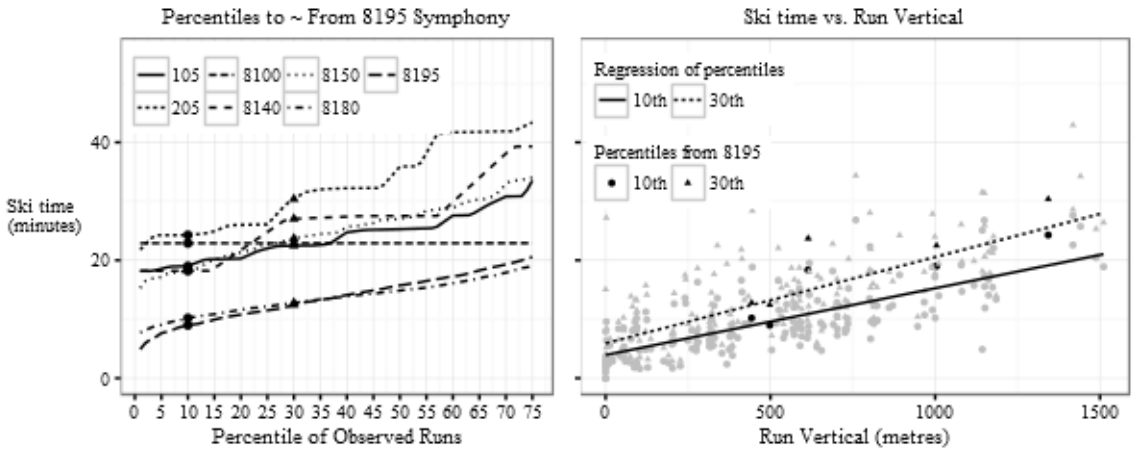


Figure 2-5 Lift-to-Lift Transition Time Percentiles and Time~Vertical Regressions

We used percentiles data only from transitions with at least 25 observations (180 out of 312 at WB, including same lift returns). We formed 100 sets of data by percentiles, each set with coordinate pairs $\{h_{ij} \equiv \text{vertical\_distance}(i,j), y_{ij}^k \equiv \text{percentile\_k\_ski\_time}(i,j)\}$ to which we fit linear models of the form $\tilde{\mathbf{y}}^k \sim \alpha^k + \beta^k \mathbf{h}$. These linear models were used to parameterize ski times $\tilde{y}_{ij}^k = \alpha^k + \beta^k h_{ij}^k \triangleq \tau_{ij}^k$ for each percentile skier to ski each feasible lift transition, regardless of whether and how many of those transitions were observed in the dataset.

The right-hand side of Figure 5 shows two of these linear models, for the 10th and 30th percentiles. The points depicted as dark circles and triangles correspond to five pairs of data on the left, those being the 10th and 30th percentile times for like transitions from 8195, among transitions with 25 or more observations only.

The higher quantile linear model has both a steeper slope and a higher intercept. The latter point is consistent with the difference in ski time between skiers of two different abilities being amplified by the length of a run. The former accounts for skiers of different abilities also being slower (faster) to navigate cordoned lift line-ups, and/or to prepare themselves to

initiate a run after disembarking. This fixed time component may also account for stops on trails that are made more frequently by less advanced skiers.

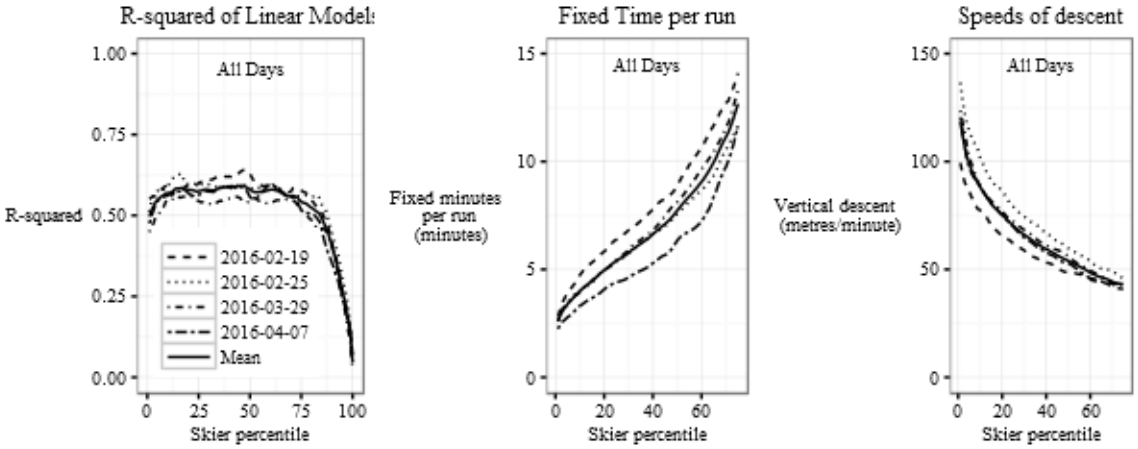


Figure 2-6 Comparison of Ski Time Regression Models Across Skier Percentiles and Days

Our linear models had consistent R-squared values in the range 0.5-0.6, across all percentiles below the 75th. Intercept estimates increased progressively from a base of three minutes for the lowest percentile, fastest skiers. We consider these fixed times in our model as buffers to account for navigation around lift entries and exits, regardless of actual ski time between lifts. The inverses of our linear model slope parameters are shown in the right-hand side of Figure 6, that is, to depict metres per minute (rather than minutes per metre). This is to simplify characterization for a 28th percentile skier, the highest for which we found a solution to the Mega Day problem, as one whose typical rate of descent we estimate to be roughly 65 metres (200 feet) per minute.

We used the four-day mean values of our linear model parameters to calculate τ_{ij}^{k} for all feasible runs at each successive skier percentile k. These calculated ski times, added to the fixed ride time of the origin lift in each case, serve as the objective function coefficients $t_{ij}^{k} = T_i + \tau_{ij}^{k}$ for each execution of our model for $k = 1, 2, ...$

2.8.3 Time Windows

We derived lift time windows empirically from our RFID scans data. Figure 7 depicts the relative volumes of skiers riding each lift, per minute-of-day interval, summed over four days. (Refer to Lift Codes in Table A1 in the appendix for lift names and details). Lift opening and

closing times often varied from day-to-day. We used the means of the four day observations to the quarter-hour, for lift opening times and lift closing times.

The fifth row of Figure 7 shows that Magic Chair (234) remains open well after 15:45 pm by which time all other lifts are typically closed. Fitzsimmons Chair (8110) is generally open for only the first couple of hours of the day (although longer on busy holidays, according to WB staff). The Whistler T-bar (8195), which lies in an upper section of Whistler Mountain and provides access to the back valley side, generally opens late and closes early (11:30 am-14:30 pm, on 25 February 2016, for example). While inclement weather, mechanical maintenance and other factors occasionally require adjustment to lift opening and/or closing times, we treat time windows as deterministic in our model.

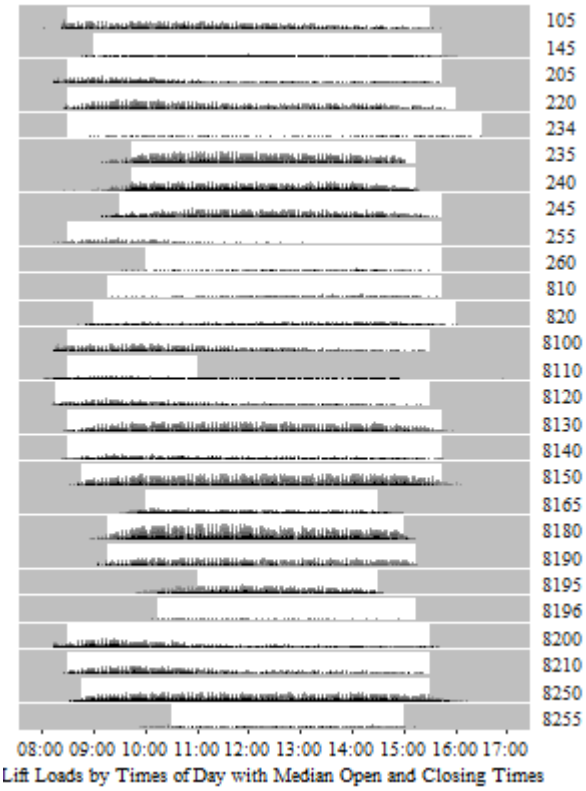


Figure 2-7 Volume of Skiers Riding Each Lift Per 5-Minute Interval (Four Days)

2.9 Results

We executed our model with parameters derived for skier abilities from the 1st to 30th percentiles (unsuccessfully for the 29th and 30th percentiles). In several instances, expected completion time exceeded the objective function value by a sum of delays incurred to adjust

the route in stage 3, due to early expected arrivals at some lifts. In most cases, delays could be merged into a late start, such that the route would still lead to the shortest elapsed time, from start to finish, for the particular skier percentile.

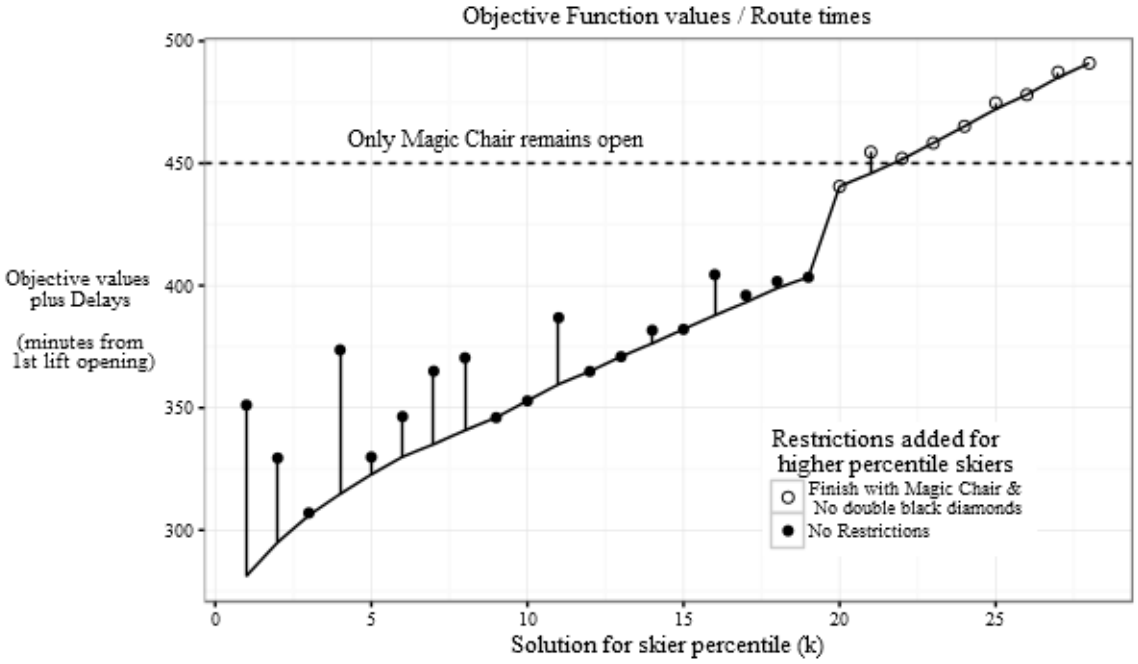


Figure 2-8 Solution Objective Values for 1st-28th Percentile Skiers

Figure 8 shows solution objective values for the 1st to 28th percentile skiers. Circles depict expected completion times after delays for time windows imposed by Stage 3.

Using the IBM ILOG CPLEX 12.3 MIP solver on an Intel i5 1.8 GHz processor, solutions were generally obtained in under two minutes, up until the 20th percentile skier where we did not initially obtain a solution through several hours and more than 3,000 iterations of stages 1-3. We noted at that point that the objective function value (expected finish time without contingency for delays) had exceeded the closing times of several favored final destination lifts in previous lower percentile solutions. Moreover, those earlier solutions often took advantage of three short but very steep transitions (‘double black diamond’ runs).

We thereafter specified that routes for higher percentile skiers should end with the Magic Chair to take advantage of its much later closing time. Simultaneously, assuming that less advanced skiers would likely prefer not to ski any of the double black diamond runs, we removed those from our set of feasible $i \to j$ connections. With these changes we obtained

rapid solutions up to the 28th percentile, albeit with elevated objective function values and expected finish times.

The significant vertical gap in Figure 8 between the 19th and 20th percentile skier finish times stems from the latter being unable to reach a final lift near the top of the mountain on time, following a route that includes Magic Chair at an earlier stage. However, leaving it as their final lift implies that the clock continues to run during a long final transition down to Magic Chair at the base of Blackcomb Mountain. This long transition also starts progressively later and takes progressively longer with each increasing skier percentile.

Aside from suggesting these additional restrictions for higher percentile skiers, our analysis of interim candidate solutions (from stages 1 and 2) provided other interesting insights. We noted that many early candidate solutions incurred time violations at the Fitzsimmons lift between its regular 11:00 am closing time and noon. The implication is that if WB were to keep the Fitzsimmons lift open an extra hour (even if only on weekends), they could vastly increase the accessibility of the Mega Day badge to a wider range of skiing abilities. Secondly, although all final solutions included only 24 lift segments, stages 1-2 occasionally produced candidate solutions which included multiple rides on a single lift. However, none of these turned out to be time feasible full tour solutions.

Solutions for many percentiles shared a number of common lift sub-sequences, yet we obtained 27 distinct routes among 28 different percentile solutions. We explain this result by the fact that transition times increase disproportionately for different skier percentiles. Run b may be twice as long as run a for a 20th percentile skier, but merely 50% longer for a 10% percentile skier, consequently run b and/or its successors may trigger a time windows violation for the slower skier, but not for the faster skier.

2.10 Implementation

2.10.1 Recommended Route

The Mega Day Challenge has considerable marketing appeal, but a skier looking at a trail map and trying to determine the best way to travel through all 24 lifts faces the prospect of committing to a difficult ski day and then perhaps being disappointed because a lift or the resort closes before he or she is able to complete the intended tour.

We proposed the recommended Mega Day Challenge route based on the shortest time, and time windows feasible, solution for the 28th percentile skier. This route is depicted in Figure 9, with vertical axes showing lift rides in sequence including time windows. We also proposed a recommendation to skiers at Whistler-Blackcomb, based conservatively on the pace of a 20th percentile skier, that "if you can ski from the Roundhouse lodge to the Whistler Village base in under 12 minutes, you should be quite capable of completing the Mega Day Challenge."

Our solution and recommendations were communicated to the WB management team and delivered to skiers through an electronic newsletter. While WB Communications noted that this content received unusually high website user 'engagement" measured by web page browsing time, it is difficult to determine impact, directly. Statistics reveal, however, that in the three months following the newsletter 224 Mega Day badges were achieved, compared with 313 over seven prior skiing months (a 67% increase). Following the acquisition of Whistler-Blackcomb by Vail Ski Resorts during the 2016-17 season, integration of WB+ into the Vail Ski Resorts EpicMix® system will expose a much larger community to the Mega Day challenge and presumably motivate even greater numbers of skiers to undertake and complete the challenge.

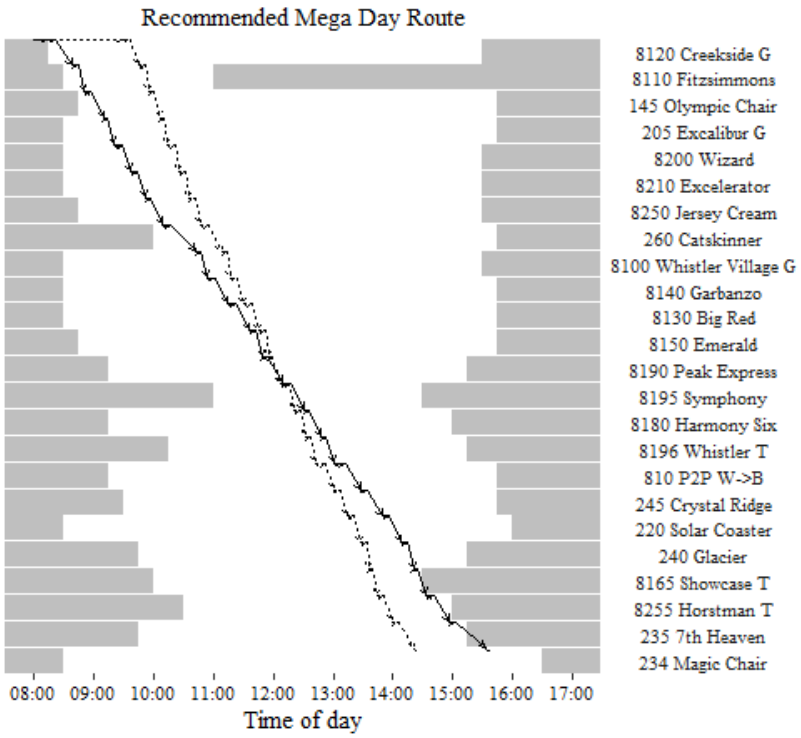


Figure 2-9 Recommended Route for 28th Percentile Skier

The Mega Day route we recommend for a 28th percentile skier is shown in Figure 9 including at the pace of a 1st percentile skier (dotted line) with a delayed start. This route should have the greatest possibility of being completed by the largest number of skiers, since faster skiers can always delay their start time and/or adjust their pace of skiing to follow this route.

2.10.2 Validation by Trial

Author Lyons skied WB on 22 January 2017 and set out to follow a route recommended by our model for a 10th quantile skier. He boarded the first lift at 8:42 am and, notwithstanding a detour caused by temporary closure of one lift, completed the Mega Day Challenge at 15:20 pm (in a total of 398 minutes). This time is quite consistent with times suggested by the model. Data obtained later showed that Lyons was the lone skier to complete a Mega Day challenge among over 15,000 skiers on the mountain that day.

Subsequently, we received an additional data file containing records of all Mega Day Challenge awards completed from inception of the RFID/WB+ system in December 2015 up to late-January 2017. This data showed that 313 Mega Day badges had been earned by 295 distinct skiers up to that point. In 138 of those cases, the Mega Day skier rode only the

minimal number of 24 lifts. In many cases the challenge was completed by groups of 2-4 skiers skiing together with the result that there had been to that date 64 Mega Day expeditions successfully completed.

We determined the routes followed in these 64 instances. These groups started from six different points: Whistler Village Gondola (19), Fitzsimmons Chair (31), Excalibur Gondola (2), Wizard Chair (1), Creekside Gondola (7) and Magic Chair (4). A surprising result was that every single one of these skier-group routes was distinct: no two groups actually followed the same sequence of 24 lifts!

These findings suggest that the Mega Day Challenge is a difficult accomplishment partly because there is no recognized or obvious 'optimal' route, despite the fact that for strong skiers, many routes are possible. Weaker skiers need to select their route more carefully if they are to accomplish the feat before the last lift closes.

2.11 Conclusion

WB's introduction of the RFID system and associated WB+ web-application make extensive skier data available to enhance the skier's experience by providing challenges and offering rewards. This 'gamification' of skiing opens up opportunities in the realm of analytics that mirror the rapid growth in the application of analytics in other predominantly team-based sports over the last two decades.

We undertook this study initially because we were intrigued by the real problem and its relationship to a variety of routing problems addressed by operations research. We were further motivated by discovering how few skiers accomplish this particular feat, partly, we believe, because of the difficulty in imagining a feasible route for their skiing ability. In the research process, we were able to develop a reasonably efficient model and to parameterize the model with real data to produce results consistent with actual records of successful Mega Day challengers. The guidance that we have been able to provide should not only inspire more skiers to attempt the feat, but also give them an appreciation for the type of operations research problem-solving approach upon which it is based.

2.12 Acknowledgement

The authors wish to acknowledge the generous support of Whistler-Blackcomb, and especially their Business IT Support Team, for making data available and for their counsel and feedback, and anonymous Interfaces reviewers whose comments greatly improved this paper.

2.13 References (Chapter 2)

Biggs, N., E. K. Lloyd and R. J. Wilson (1976). Graph theory 1736-1936. Oxford [Eng.], Clarendon Press.

Campbell, A. M., M. Gendreau and B. W. Thomas (2011). "The orienteering problem with stochastic travel and service times." Annals of Operations Research **186**(1): 61-81.

Cook, W. (2012). In pursuit of the traveling salesman: mathematics at the limits of computation. Princeton, N.J, Princeton University Press.

Dantzig, G., R. Fulkerson and S. Johnson (1954). "Solution of a large-scale travelling salesman problem." Operations Research(2): 393-410.

Dash, S., O. Günlük, A. Lodi and A. Tramontani (2012). "A Time Bucket Formulation for the Traveling Salesman Problem with Time Windows." INFORMS Journal on Computing **24**(1): 132-147.

De Biagi, V., B. Friggo and B. Chiaia (2013). "A network-based approach for the study of criticalities in ski-resorts." International Snow Science Workshop Proceedings.

Desrosiers, J., Y. Dumas, M. M. Solomon and F. Soumis (1995). Chapter 2 Time constrained routing and scheduling, Elsevier B.V. **8:** 35-139.

Eksioglu, B., A. V. Vural and A. Reisman (2009). "The vehicle routing problem: A taxonomic review." Computers & Industrial Engineering **57**(4): 1472-1483.

Focacci, F., A. Lodi and M. Milano (2002). "A Hybrid Exact Algorithm for the TSPTW." INFORMS Journal on Computing **14**(4): 403-417.

Jünger, M., T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, L. A. Wolsey and SpringerLink, Eds. (2010). 50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art. Berlin, Heidelberg, Springer Berlin Heidelberg.

Kantor, M. G. and M. B. Rosenwein (1992). "The Orienteering Problem with Time Windows." The Journal of the Operational Research Society **43**(6): 629.

Khan, A. (2010). "SCVNGR Gamifies Skiing At Stowe Mountain Ski Resort." AdWeek.

Letchford, A. N., S. D. Nasiri and D. O. Theis (2013). "Compact formulations of the Steiner Traveling Salesman Problem and related problems." European Journal of Operational Research **228**(1): 83-92.

Öncan, T., İ. K. Altınel and G. Laporte (2009). "A comparative analysis of several asymmetric traveling salesman problem formulations." Computers and Operations Research **36**(3): 637-654.

Pesant, G., M. Gendreau, J.-Y. Potvin and J.-M. Rousseau (1998). "An Exact Constraint Logic Programming Algorithm for the Traveling Salesman Problem with Time Windows." Transportation Science **32**(1): 12-29.

Sean (2013). "How Gamification Changed the Sport of Skiing (And How it Applies to Your Life)." Location Rebel (Blog).

Shealy, J. E., C. F. Ettlinger and R. J. Johnson (2005). "How Fast Do Winter Sports Participants Travel on Alpine Slopes?" Journal of ASTM International **2**(7): 1-8.

Tas, D., M. Gendreau, O. Jabali and G. Laporte (2016). "The traveling salesman problem with time-dependent service times." European Journal of Operational Research **248**(2): 372-383.

Vansteenwegen, P., W. Souffriau and D. V. Oudheusden (2011). "The orienteering problem: A survey." European Journal of Operational Research **209**(1): 1-10.

2.14 Appendix A – Lifts Information

The table below lists the 27 discrete lift segment members of N, along with which of the 24 lift groups G each belongs, their base and top elevations (metres above sea level), standard ride times, and typical opening and closing times.

Index	Lift Name	Lift System G	Base Elev (metres)	Top Elev (metres)	Ride Time (mins)	Lift Open (HH:MM)	Lift Close (HH:MM)
8100	Whistler Village Gondola	1	684	1023	0:05:46	8:30 am	15:30 pm
105	Whistler G O Station	1	1023	1827	0:13:09	8:30 am	15:30 pm
145	Olympic Chair	2	1020	1145	0:06:30	8:45 am	15:45 pm
205	Excalibur Gondola	3	684	765	0:02:43	8:30 am	15:45 pm
255	Excalibur G Mid-Base	3	765	1133	0:05:36	8:30 am	15:45 pm
220	Solar Coaster	4	1250	1862	0:07:16	8:30 am	15:45 pm
234	Magic Chair	5	683	778	0:04:07	8:30 am	16:30 pm
235	7th Heaven	6	1660	2249	0:06:37	9:45 am	15:15 pm
240	Glacier	7	1547	2142	0:07:07	9:45 am	15:15 pm
245	Crystal Ridge	8	1282	1822	0:06:34	9:30 am	15:45 pm
260	Catskinner	9	1539	1860	0:09:10	10:00 am	15:45 pm
810	P2P W->B	10	1825	1878	0:11:07	9:15 am	15:45 pm
820	P2P B->W	10	1878	1825	0:11:07	9:00 am	15:45 pm
8110	Fitzsimmons	11	689	1021	0:06:43	8:30 am	11:00 am
8120	Creekside G	12	661	1302	0:07:52	8:15 am	15:30 pm
8130	Big Red	13	1301	1846	0:09:22	8:30 am	15:45 pm
8140	Garbanzo	14	1021	1676	0:07:43	8:30 am	15:45 pm
8150	Emerald	15	1413	1834	0:06:51	8:45 am	15:45 pm
8165	Showcase T	16	2146	2274	0:03:28	10:00 am	14:30 pm
8180	Harmony Six	17	1584	2102	0:06:07	9:15 am	15:00 pm
8190	Peak Express	18	1771	2172	0:03:46	9:15 am	15:15 pm
8195	Symphony	19	1529	2027	0:07:43	11:00 am	14:30 pm
8196	Whistler T	20	1786	1962	0:05:27	10:15 am	15:15 pm
8200	Wizard	21	688	1252	0:08:27	8:30 am	15:30 pm
8210	Excelerator	22	1131	1635	0:07:01	8:30 am	15:30 pm
8250	Jersey Cream	23	1547	1912	0:05:21	8:45 am	15:30 pm
8255	Horstman T	24	2047	2250	0:06:57	10:30 am	15:00 pm

Table 2-3 Whistler-Blackcomb Lifts Information

Chapter 3

3 Elective Surgery Scheduling to Improve Perioperative Patient Flow

3.1 Abstract

This paper addresses a practical problem of scheduling operating room (OR) elective surgeries to minimize the likelihood of surgical delays caused by unavailability of capacity for patient recovery in a central post-anesthesia care unit (PACU). We segregate patients according to their patterns of flow through a multi-stage perioperative system and use characteristics of surgery type and surgeon booking times to predict time intervals for patient procedures and subsequent recoveries. Working with a hospital in which 50+ procedures are performed in 15+ ORs most weekdays, we develop a constraint programming (CP) model that takes the hospital's elective surgery pre-schedule as input and produces a recommended alternate schedule designed to minimize the expected peak number of patients in the PACU over the course of the day. Our model was developed from hospital data and evaluated by application to daily schedules during a testing period. Schedules generated by the model indicated the potential to reduce the peak PACU load substantially, 20-30% during most days in our study period, or alternatively reduce average patient flow time by up to 15% given the same PACU peak load. During the evaluation we also developed tools for schedule visualization that can be used to aid management both prior and post surgery day, to plan PACU resources, propose critical schedule changes, identify the timing, location and root causes of delay, and to discern the differences in surgical specialty case mixes and their potential impacts on the system.

3.2 Introduction

Demographic changes and political economic conditions have intensified demand for more efficient health care operations, including a call to reduce elective surgery wait-times. Health Quality Ontario (Ontario 2019) for example, an organization established to advise the province regarding performance of its $55 billion annual health care

expenditures, maintains an up-to-date public internet dashboard listing of surgical wait-times, for six key categories of procedures not only at the provincial level, but also by region and individual hospital.

London Health Sciences Centre (LHSC) is a 600-bed regional tertiary care hospital that has been exploring opportunities to increase surgical throughput by establishing some operating room (OR) schedule blocks as Rapid and Standardized Operating Rooms (called 'RASTOR rooms".) Their aim is to reduce wait-times in services where it exceeds the provincial averages, by more than double in some service categories. While these RASTOR rooms involve more numerous and shorter procedures than other ORs, they operate within the same perioperative system which includes a centralized Post Anesthesia Care Unit (PACU).

More rapid patient-procedures in these RASTOR rooms create an imperative that these ORs can function without obstruction from external processes and conditions, as happens when a patient whose surgery has been completed cannot be moved from the OR due to the PACU being at full capacity. This delays the subsequent patient, causes lost OR time for the surgeon, and is a waste of utilities and staff time to support open ORs, often re-incurring added costs later at higher overtime rates. Meanwhile, some of the PACU bed capacity may be occupied inappropriately, due to a downstream patient destination being unable to receive a patient into the next stage of care, either a hospital ward or PACU2 (step-down recovery, just prior to discharge from the hospital.)

This problem already exists with LHSC's current daily volumes of scheduled patients and it is likely to worsen if volumes are increased to address the wait-time performance issue. We note that some recent developments in anesthesia practices (Nilsson, Jaensson et al. 2019) can eliminate the need for PACU recovery in certain cases, however general anesthesia remains dominant and leads to similar recovery times for most RASTOR surgeries, notwithstanding that the procedures themselves are being made shorter, more rapid and frequent. One consideration is whether additional PACU physical capacity is required to accommodate greater patient volumes. Our research considers whether case

scheduling might be better coordinated across ORs, and perhaps integrated with PACU bed management, to improve patient flow.

We addressed these issues by developing a model to translate daily patient rosters (i.e., existing schedules) into individual OR sequences aimed collectively at minimizing the peak PACU patient load. We run our model iteratively, with each iteration seeking a schedule to meet a progressively lower peak PACU load target, and within that context to choose a solution that minimizes total patient flow time and thereby OR makespans. Flow time is defined by (Conway, Maxwell et al. 1967) as the total time a job (patient) spends in a shop (i.e. preparation, OR and recovery) using the simplifying assumption that all patients arrive at the start of their respective OR block opening times. We used information available on days following surgery to evaluate our model, and to visualize events (including delays) across time, location and status of patients, ORs and PACU beds. Our methods take the existing 'pre-schedule' and respond with a best alternative sequencing of procedures for the ORs. We then extract and transform data collected during schedule execution into a visual flow format, with complimentary metrics, for surgeons and perioperative management to monitor and refine their scheduling and staffing decisions.

We believe that we are the first to address the general problem in this paper using CP, and that strengths of this form of programming in terms of intuitiveness and flexibility are well-demonstrated in this setting. In effect, we have developed a platform which can be easily-implemented to suggest beneficial schedule changes based on a very wide range of factors that can be captured and conditions that might be imposed in a complex setting. Notwithstanding such extensions, our analysis shows that a collective and calculated approach to scheduling in multiple ORs can have significant positive effects in terms of reducing by 20% or more the expected peak patient loads in the PACU.

3.3 Background Literature

The body of research concerned with operating room scheduling is substantial. (Blake and Carter 1997), (Cardoen, Demeulemeester et al. 2010), (Guerriero and Guido 2011) , (Gür and Eren 2018) provide extensive literature reviews. During the eight years

spanning the final three of these reviews, the number of papers considered expanded from 115 to 170, or roughly one every two months, as a testament both to the heightened importance of these issues to health care managers and to the variety of specific problems and operating contexts involved. (Cardoen, Demeulemeester et al. 2010) and (Gür and Eren 2018) share a similar structure which classifies papers according to six dimensions, as follows: patient-case characteristics; performance criteria and measurement; decision delineation or planning context; research method or solution technique; handling of uncertainty; and applicability of research.

Patient-case characteristics refer to whether the problem considered includes both elective (pre-scheduled) and/or urgent or emergent (unscheduled) surgeries, and whether distinctions are drawn between inpatients and outpatients. Our research focuses on elective surgeries, as do nearly two-third of the papers reviewed in (Gür and Eren 2018). This is warranted by three levels of capacity allocation the hospital has for unscheduled, emergent and urgent cases with dedicated OR time blocks reserved for priority cases to be inserted into the surgery day (real time allocation). In LHSC, one OR is fully dedicated to priority 1 emergencies as they arrive. Finally, a scheduled elective case may be pre-empted, if necessary to accommodate an emergent case if other options are not available. As for elective cases, our research considers and distinguishes between outpatients and inpatients and also same day admissions that share a common routing inbound and outbound with outpatients and inpatients.

The dominant performance criteria applied in the OR scheduling literature is utilization (especially of ORs, surgeons and nursing teams), followed by waiting times (of patients, surgeons, or as a measure of system throughput), overtime, and load levelling. However, roughly 40% of papers reviewed in (Gür and Eren 2018) included other and/or multiple criteria, as a reflection of many different perspectives about what outcomes are most/least desirable in a given hospital setting. Our model's objective function falls in the category of throughput measure, but indirectly also OR utilization, as by minimizing average patient flow time we are also minimizing the sum of individual OR makespans given their specified workloads. Our iterative approach, on the other hand, falls in the category of load levelling, as it explicitly seeks to do that for the PACU.

Decision delineation or planning context incorporates but extends traditional classifications of planning: strategic, e.g., case mix planning; tactical, e.g., block time allocation; and operational, e.g., case sequencing (Santibáñez, Begen et al. 2007). (Cardoen, Demeulemeester et al. 2010) consider the matter in two dimensions which represent whether decisions are to be made concerning date, time, room or capacity allocation, and whether the operational unit of concern is individual or groupings of surgeons, patients or other units such as the PACU. Our research takes an integrated approach in assuming daily surgeon-patient-room assignments as given (by the hospital's pre-schedule) and seeking the best sequence of these procedures across and within the ORs such that they can be performed unhampered by downstream capacity limitations, in the PACU and/or wards.

Regarding research methodology and solution techniques, (Gür and Eren 2018) identified thirteen categories dominated by Mixed Integer Linear Programming (MILP), Heuristic Algorithms, Simulation, and Integer Programming (IP). Only four out of 170 papers employ a CP solution technique as does our research, the oldest of these being from 2010. Of special interest is (Wang, Meskens et al. 2015) that provide a comparison of MILP to CP for scheduling problems in operating theatres. Our research exploits certain advantages of CP such as enabling a compact and intuitive model formulation, the ease with which expressing multiple and complex constraints can be added, and computational efficiencies.

Whereas other mathematical programming techniques, such as MIP can lead to exact 'optimal' solutions, or improved approximations thereof, we argue that optimality is not the ultimate promise in surgery scheduling, as the process involves many entities (patients, surgeons, support staff and equipment) with uncertainty at every stage from preparation through procedure and recovery. About two-thirds of papers reviewed in (Gür and Eren 2018) employ deterministic models, while acknowledging uncertainty and generally accommodating it with some manner of time buffering. We take a similar approach, firstly by using only specified allowable start times, secondly by hedging long on procedure and recovery duration estimates. To the extent that procedures are not on the long side of their estimates, cleanup time do not present a problem, and where

procedures turn out to be longer than expected, cleanup can be accelerated relatively easily with the help of floating staff in the system.

In classifying reviewed articles according the 'applicability of research', (Gür and Eren 2018) identified a near-equal division of studies that used a theoretical data set versus those based on real data. Our model is parameterized from actual hospital data and is executed with a daily schedule as input using only data available to the scheduler prior to the surgery day. We also measure our model's performance retrospectively, using actual procedure and recovery durations that become available on the day following surgery, whereupon we compare the day's actual PACU load to what would have occurred with the actual task durations but following case sequences proposed by our model.

We note that six out of 115 articles reviewed in (Cardoen, Demeulemeester et al. 2010) address load levelling of the PACU, but in different planning contexts and settings than those we study, generally using different solution techniques, and/or applying theoretical data rather than actual data. The articles that are most proximate to our research include (Abedini, Li et al. 2017), (Bam, Denton et al. 2017), and (Fairley, Scheinker et al. 2018). The first of these applies theoretical data and a deterministic MIP for a master surgery-scheduling problem that considers the possibility of next day surgical blocking by the PACU given the current state of occupancy in various units of the hospital. We instead use actual daily case mixes and seek to avoid blocking of surgeries by the PACU indirectly, without specific knowledge of downstream unit occupancies. Retrospectively, we provide a tool to identify the frequency, timing and location of these events to help characterize and remediate delays, according to what else had happened in-situ or was happening elsewhere in the system to cause the delay. (Bam, Denton et al. 2017) employ a MIP and two-phase heuristic to determine surgeon-to-OR assignments followed by surgical case sequencing. We study a context where surgeons have pre-assigned OR time blocks and a pre-specified list of patient procedures to perform within those blocks. (Fairley, Scheinker et al. 2018) work is similar to this article in subject matter and setting but uses an MIP solution procedure to level daily PACU loads, placing special emphasis on a machine learning model for more accurate prediction of procedure and recovery durations. We assume much less about information available for such accurate

predictions and emphasize instead the potential of intuitive and adaptable CP optimization, combined with process visualization, disaggregation and iterative refinement, in spite of the vast uncertainties in perioperative care.

3.4 Problem Description

Figure 1 provides a global view of the perioperative system through which three basic categories of patients flow: One day stay (ODS), same day admission (SDA), and in-patients (IP). At the highest level, ODS and SDA patients enter the system from outside the hospital through admission and preparation, whereas IP patients arrive from a hospital ward. As for departures, ODS patients leave through Day Surgery from whence they arrived, whereas SDA patients are transferred into a hospital ward, through which IP patients also return to their ward, although either may be transferred instead to an intensive care unit (ICU) bed.

At a more detailed level within the perioperative process, ODS patients follow one of three different paths involving recovery in *either or both* the PACU and Day Surgery, the latter referred to as PACU2 in this context. SDA patients normally recover in the PACU before being transferred to a ward bed when one becomes available. IP patients often recover in the PACU but occasionally bypass it returning directly to their ward bed. Similarly, some ODS patients may bypass the PACU if they require only minor recovery, which they can undergo in PACU2. Because our model addresses only elective cases, we make a conservative assumption that all surgeries will be followed by a PACU stay, as is most often the case. Some patients may be diverted on a given day but these cases will serve as a counterbalance to some of the added PACU load from unplanned emergent and urgent cases that arise.

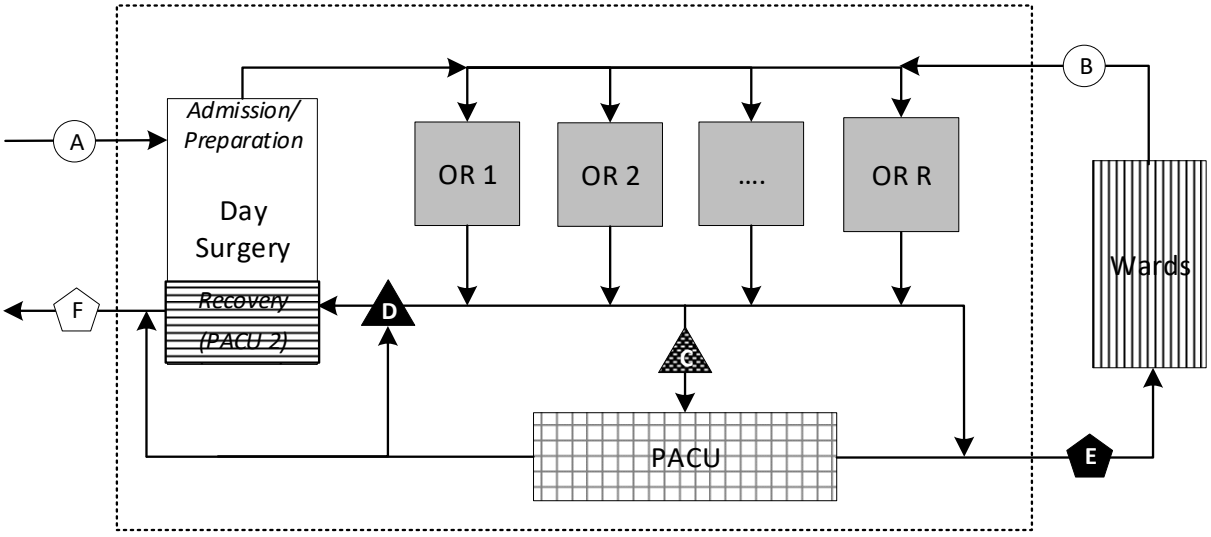


Figure 3-1 Perioperative System Flow

The stage of perioperative flow of greatest concern to management (point 'C' in Figure 1) lies on the most common path followed by patients immediately after surgery. If this pathway is blocked, due to a PACU hold, a patient must remain under anesthesiologist surveillance in the OR, preventing room turnover and delaying the start of the next surgery. In these situations, the hospital incurs the opportunity costs of an idle and expensive room and/or team. Additional costs are incurred if overtime is required at the end of the day to make up for the delay or if surgeries have to be postponed. Surgeons lose some of their limited OR block time to perform revenue generating (and backlog reducing) activities. The health care system and patients awaiting treatment suffer, as fewer elective surgeries performed in a given day, week, month and year translate into larger queues and longer wait-times.

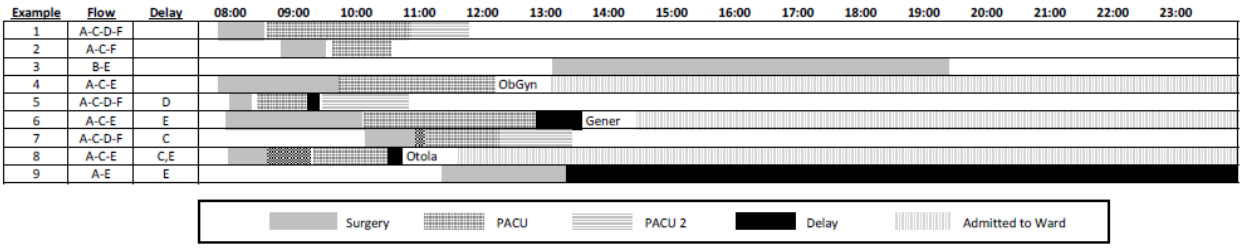


Figure 3-2 Patient Flow Examples

Figure 2 shows several examples of patient flow. The characters in the 'Flow' and 'Delay' columns correspond to critical points in process as depicted in Figure 1, that is,

the path through which each example patient flowed and the location(s) at which they encountered any delays. The patterns within the timelines correspond to the various locations in Figure 1. We are concerned with delays both upstream and downstream from the PACU, as eliminating the latter can effectively eliminate at least some of the former. The downstream delays are represented in solid black, corresponding to the same format as points 'D' and 'E' in Figure 1, and these are delays in moving patients *from* the PACU to either PACU2 or to a ward bed. Checkered black intervals represent upstream delays moving patients *into* the PACU (at point 'C' of Figure 1) coming from any one of the ORs. Brief gaps between intervals correspond to the patient transit times and thereby separated time-stamp gaps in the hospital data.

The hospital typically opens 15-17 ORs on weekdays, with 1-3 operating outside of the core scheduled surgery period of 8:00-18:00. ORs can run overtime, but the management of perioperative care has discretion to cancel a final OR procedure if it is deemed likely to induce substantial overtime. The number of scheduled surgeries per day at LHSC typically falls between 50 and 65 with 8 to 12 urgent or emergent (unscheduled) procedures added on most days.

The PACU is rarely empty as it often begins with an overnight patient trauma patient, or perhaps one from the previous day still waiting for a ward bed as in Ex. 9 of Figure 2. PACU occupancy rises quickly in the morning to roughly the number of open ORs as their first procedures are completed. It continues to rise through the mid-afternoon before declining rather slowly at the end-of-day. This pattern stems from patient recoveries being on average longer than the associated procedure times, or the OR service rate is greater than the PACU service rate, consequently the PACU cannot release patients as quickly as it receives them, even in the best case without any downstream delays. Finally, as new PACU patient arrivals slow toward the end of the day, patients with long recoveries keep the PACU load from dissipating quickly. Finally, the unavailability of a ward bed or the inability to move a patient back to Day Surgery /PACU 2 will cause PACU occupancies to remain above manageable preferred levels.

3.5 Model Development

We develop a scheduling model aimed at mitigating the problem of blocking at points ‘C’ and ‘E’ (in Figure 1) in this complex and multidimensional situation. Blocking at point ‘D’ can be addressed with simpler heuristics as presented in the discussion section of the paper. We begin by stating several assumptions, conventions and simplifications that we use to provide for a tractable, generalized model that can be refined to accommodate a variety of additional constraints, as required.

3.5.1 Assumptions

First, we assume that procedure and recovery durations can be predicted with reasonable accuracy and are available well-enough in advance to inform the scheduling of procedure sequences and start times. We note that (Fairley, Scheinker et al. 2018) provide a compelling case for machine learning techniques to improve such time estimates, however, to take advantage of these methods information systems and internal communication require more details about patients and procedures than are currently available at LHSC. As one example, we found that ASA scores (a measure of patient physical condition established by the American Society of Anesthesiologists (ASA) could substantially improve recovery time estimates, but these are only determined just prior to surgery by the attending anesthetist and so are not useful in OR scheduling.

Second, we disregard unscheduled (trauma/priority) cases as they are difficult to anticipate and not known at the time of scheduling. We assume that managing scheduled cases to minimize likelihood of PACU blockages will better accommodate unscheduled cases as they arise. We also assume that all patients will recover in the PACU, despite this not being required in roughly 10% of cases, and thus serving to counterbalance the effect of not including unscheduled cases in our PACU load estimates.

Third, we assume that surgeons, OR teams, and patients, are indifferent toward case sequence and timing; so long as the procedure will be done in the designated OR during the block surgery time allocated to the surgeon. However, we note that our model can be adapted to address specific needs or preferences regarding case sequencing if necessary.

Fourth, our objective is minimizing the total completion time (flow time) of all surgeries scheduled for the day. This is an appropriate collective goal although it may not be individually optimal for every OR and surgeon involved.

Finally, we follow the convention that the hospital wishes to maintain scheduled start times only at :00 and :30 of the hour, and that OR cleanup (not including closure and setup) can generally be accomplished in 20 minutes or less.

3.5.2 Hospital Data

We obtained data regarding more than 27,000 surgical procedures performed at the hospital from 1 April 2017 to 31 March 2018 and developed two regression models for predicting procedure and recovery durations based on information available at time of schedule release (normally 2:00 pm on the afternoon of the day preceding surgery.) We reserved four weeks of this data for testing our models and to provide a proof of concept. We later used the model repetitively over three weeks of surgical activity from 25 March to 12 April 2019 to pilot its use with actual daily schedules.

3.5.3 Procedure and Recovery Durations

We developed linear regression models to predict procedure and recovery times. In both regression models (Figure 3), the independent variables included: initial OR time as booked by the surgeon (30-minute increments), patient age, and surgical discipline (binary indicator variables). After some surgical discipline binary variables were removed from the recovery model due to lack of statistical significance, the model for procedure durations has an R-squared of 0.83 whereas the model for recovery durations is much less explanatory with R-squared of 0.19. These two models provide a reasonable starting point for anticipating arrivals of patients into the PACU based on their procedure start times, and to predict their length of stay in the PACU. We note that the standard error of both predictions is in the realm of 40 minutes, reflecting the substantial variability among cases but this is a similar magnitude of error as reported elsewhere in the literature (Wright, Kooperberg *et al.* 1996),(Larsson 2013).

```
Call:
lm(formula = Duration ~ Booking + Cardiac + Cardio + Dent + General +
    Med + Neuro + ObGyn + Opthalm + Ortho + Otolary, data = LMOpsRev)

Residuals:
    Min      1Q  Median      3Q     Max
-317.69  -19.40   -3.11   15.95  334.31

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -12.925780   2.039300  -6.338 2.38e-10 ***
Booking       0.898066   0.003433 261.636  < 2e-16 ***
Cardiac       9.681865   2.585051   3.745 0.000181 ***
Cardio        9.153573   2.134385   4.289 1.81e-05 ***
Dent         17.777139   2.457676   7.233 4.91e-13 ***
General      10.017199   2.158234   4.641 3.49e-06 ***
Med          11.706073   2.085638   5.613 2.02e-08 ***
Neuro         6.881878   2.260027   3.045 0.002330 **
ObGyn        10.270732   2.354016   4.363 1.29e-05 ***
Opthalm      20.174514   2.729100   7.392 1.51e-13 ***
Ortho        11.087332   2.338748   4.741 2.15e-06 ***
Otolary       8.354000   2.643552   3.160 0.001580 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 39.77 on 17015 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.8288,    Adjusted R-squared:  0.8287
F-statistic:  7488 on 11 and 17015 DF,  p-value: < 2.2e-16
```

```
Call:
lm(formula = Recovery ~ Booking + Age + Cardiac + Dent + General +
    ObGyn + Opthalm + Otolary, data = LMOpsRev)

Residuals:
    Min      1Q  Median      3Q     Max
-165.30  -33.71  -10.28   23.72  221.70

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  60.60807    1.00670  60.205  < 2e-16 ***
Booking       0.19654    0.00428  45.922  < 2e-16 ***
Age           0.38198    0.01675  22.799  < 2e-16 ***
Cardiac     -12.43342    2.09038  -5.948 2.77e-09 ***
Dent        -29.68442    1.77294 -16.743  < 2e-16 ***
General       5.02272    1.09591   4.583 4.61e-06 ***
ObGyn       -10.93568    1.58121  -6.916 4.81e-12 ***
Opthalm      31.33003    2.32251  13.490  < 2e-16 ***
Otolary      16.69701    2.17008   7.694 1.50e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 48.98 on 17017 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.1941,    Adjusted R-squared:  0.1937
F-statistic: 512.4 on 8 and 17017 DF,  p-value: < 2.2e-16
```

Figure 3-3 Linear Regression Models for Procedure and Recovery Durations

Except for emergent cases, surgical ORs cycle through four states while they are open: setup, surgery, closure, and cleanup; where a patient's stay in the OR coincides with all phases but the cleanup. Although we had access to hospital data that would allows us to distinguish between setup, surgery and closure times, we treated them together as one procedure time estimate, both in formulating and applying the linear models and to predict procedure times. One exception is that if a PACU delay was indicated at the end of the procedure, we excluded closure time from the procedure duration and used it instead as a measure of PACU delay, even though some of this time would have in fact been required to close the procedure. Where no such delays occurred, as happened in the majority of ORs on most days, the sum of procedures times divided by total OR open time provides a direct measure of OR utilization.

3.6 Constraint Programming (CP) Model

We chose to model this problem using CP for several reasons including those highlighted in (Hanset, Meskens et al. 2010) and (Wang, Meskens et al. 2015). By comparison, OR scheduling with Mixed Integer Programming (MIP) typically requires a three-dimensional representation of patients, rooms and aggregated time periods (e.g. 5-minutes, or more) to reduce the combinatorial size of the problem for more efficient computation. CP enables a more granular solution and offers intuitive variable constructs such as intervals and sequences, along with unique constraints that take advantage thereof, to express the problem very compactly in terms of 'real-world' variables. Our model introduces the concept of 'occupancies' that are intervals (of individually specified

length) indexed simply by patient and room. Our core problem is not so complex that it requires greater computational efficiency which MIP may offer in a large, minimally-constrained instance, as we are generally able to find solutions in a minute or less. However, CP offers the possibility to more easily accommodate many different constraints that we envision will be required for any substantial implementation, such as requiring that two surgeries be scheduled back-to-back or that a specific surgery in a particular OR be scheduled first due to a unique set-up requirement. (Wang, Meskens et al. 2015) provide evidence that as many such constraints are added to the problem, CP in fact becomes more computationally efficient than MIP.

We developed a CP model as detailed below using IBM's ILOG CPLEX 12.9 CP Optimizer. Variable inputs to the model include a set of planned procedures for a given day with initial start times and booking durations as requested by the surgeon, along with procedure and recovery durations estimated by the linear models above. In addition to this procedure information, the model takes as input a set of OR block times which specify the surgical discipline to which the OR is allocated, and the opening and closing times within which the set of procedures must be scheduled.

3.6.1 Preliminaries

Model definitions are presented under the following groupings, after which we provide a model formulation, and some discussion:

1. Tuple Sets (*Patients* P, *Rooms* R)
2. Deducible Parameters and Relationships (*Occupancy Durations* X^{pr}, *Assignment Matrix* M)
3. Decision Variables
   a. Intervals (*Occupancies* O^{pr}, *Patient PACU Occupancies* U^{p}, *OR Occupied Times* V^{r})
   b. Sequences (*OR Loads* L^{r}, *Patient Paths* Π^{p})
4. Interval Parameters (*OR Open Windows* W^{r}, *Surgical Day Time* T)
5. Setting Specific Parameters (*Allowable OR Start Times* Y^{r}, *OR Cleaning Times* K^{r})
6. Intermediate Functions and Expressions (Patient PACU Pulse pU_t, Patient Flow Time F^{p})
7. Objective Function

Tuple sets provide for simple extraction of model input data from the hospital's pre-schedule. These data are processed into deducible parameters related to patient assignments-to and durations-in specific ORs and subsequently in the PACU (if required.) The ultimate decision variables O^{pr} fully specify a solution that is a feasible set of occupancy intervals (start-to-end times) of all patients in their respective ORs as well as in the PACU, as required. However, the time locations of these occupancy intervals are partially a function of complementary decisions regarding occupancy interval sequences. These interval sequences are determined for both sets of procedures within each room, and each patient's flow through the two-stage system. Additional parameters help to constrain the problem according to local practices, and additional functions translate any solution candidate set of interval decisions into metrics against which the objective is measured and optimized.

3.6.2 Definitions

Note: The intervals decision variables and interval parameters defined below each possess the following properties:

.start	an integer value lying with a specified range '.in ..' (see below)
.end	an integer value lying in the same specified range '.in ..' : $.end \geq .start$
.size	an integer value representing the difference $= .end - .start$
.optional	a boolean value (default: False) if interval not required in a solution
.in ..	a specified range constraining the above, $.startMin...endMax$

3.6.3 Tuple Sets

$\overline{R}$	a set of discrete ORs (operating rooms)
R	the union of the set $\overline{R}$ with the PACU (a special room): $R = \overline{R} \cup PACU$

r For each $r \in R$ we introduce a tuple of input data with the following

.id	key field $r = 1, 2, .., \lvert R \rvert$
.room	OR # or PACU
.service	name of the surgical group (e.g. Ortho)
.open	(.start) of *first* prescheduled surgical procedure
.close	(.start + .booking) of *last* prescheduled surgical procedure

P a set of patients on each of whom a surgical procedure is to be performed

p For each $p \in P$ we introduce a tuple of input data with the following properties:

.patient	key field $p = 1, 2, .., \lvert P \rvert$
.service	name of surgical group performing the service
.room	OR # or PACU
.start	prescheduled start time (e.g., 480 mins = 08:00)
.booking	a pre-scheduled booking window (e.g., 30, 60, 120 mins)
.procTime	the expected duration of the procedure
.recTime	the expected duration of the patient's recovery
.precedence	a real number lying in $e^{-q}: \; q \in [0,1]$

3.6.4 Deducible Parameters

X^{pr} an integer representing the time a patient p is expected to spend in room r

$$X^{pr} = \left\{ \begin{matrix} p.procTime & , \text{if } r: \; r.room == p.room \\ p.recTime & , \text{if } r: \; r.room == PACU \\ 0 & , otherwise \end{matrix} \right\}, \; \forall p \in P, \; \forall r \in R$$

M a matrix of dimensions $\lvert P \rvert \times \lvert R \rvert$ with binary values:

$$M^{pr} = \left\{ \begin{matrix} 1 & , \text{if } X^{pr} = O^{pr}.size > 0 \\ 0 & , \text{if } X^{pr} = O^{pr}.size = 0 \end{matrix} \right\}$$

3.6.5 Decision Variables

3.6.5.1 Intervals

O^{pr} an interval decision variable such that $O^{pr}.size = X^{pr},\ \forall p \in P,\ \forall r \in R$

note: $M^{pr} = 0 \Leftrightarrow O^{pr}.optional = \text{True}$

U^{p} an interval decision variable representing the time a patient p will occupy a bed in PACU

V^{r} an interval decision variable representing the time during which room r is expected to be open, beginning at $r.open$ and ending with the latest $O^{pr}.end$ in that room

S^{p} an interval decision variable represent a patient's total length of stay

3.6.5.2 Sequences

L^{r} a sequence decision variable on every O^{pr} of *type* r,
such that $L^{r}.start = \min(O^{pr}.start)$ and $L^{r}.end = \max(O^{pr}.end)$

Π^{p} a sequence decision variable on O^{pr} of *type* p,
such that $\Pi^{p}.start = \min(O^{pr}.start)$ and $\Pi^{p}.end = \max(O^{pr}.end)$

We refer to L^{r} as the load on r, and Π^{p} as the path of patient p.

3.6.6 Interval Parameters

W^{r} an interval parameter representing the range $[r.open\ ..\ r.close]\ \forall r \in R$

3.6.7 Setting Specific Parameters

T a fixed range from 480..1440 (08:00 am-24:00 pm)

Y^{r} a step function which restricts possible start times of procedures

k	a minimum allowable clean-up time gap between procedures
K^r	sets of $\{p_i, p_j, k\}$ specifying minimum time between procedures $i \rightarrow j$ in r
Q	a maximum allowable number of patients in the PACU at any time

3.6.8 Intermediate Functions and Expressions

$pU_t = \sum_p pulse\ [U^{pr}],1$, $\forall t \in T$	the sum of patients in the PACU at points-in-time over minutes of the day.
$F^p = \mathrm{O}^{pr}.end - W^r.open$, $\forall p \in P,\ r : M^{pr} = 1$	flow time for each patient, from the opening of their OR until the expected completion of their surgery completion, given the procedure start time in a proposed schedule solution.
$B^r = \min_{p \in P: X^{pr}=1} \{startOf(O^{pr})\}$	the earliest procedure start time in a room
$C^r = \max_{p \in P: X^{pr}=1} \{endOf(O^{pr})\}$	the latest procedure end time in a room

3.6.9 Objective Function

$\text{minimize} \sum_{p \in P} F^p$	the sum of patient flow times

Alternatively,

$\text{minimize} \sum_p (F^p \times p.precedence)$	the sum of patient flow times weighted by their respective precedence scores

3.6.10 Formulation

$$\min \sum_{p \in P} F^{p} \tag{1}$$

s.t.

$\forall p \in P$

$$span(S^{p}, all(r \in R, O^{pr})) \tag{2}$$

$$last(\Pi^{p}, O^{p|R|}) \tag{3}$$

$$noOverlap(\Pi^{p}) \tag{4}$$

$$lengthOf(S^{p}) = \sum_{r \in R} O^{pr}.size \tag{5}$$

$$startOf(U^{p}) = O^{p|R|}.start \tag{6}$$

$\forall r \in R, r \neq |R|$

$$noOverlap(L^{r}, K^{r}) \tag{7}$$

$$B^{r} >= r.open \tag{8}$$

$$C^{r} \leq 1200 \text{ (i.e. 20:00 pm)} \tag{9}$$

$\forall r \in R, r \neq |R|, \ \forall p \in P$

$$forbidStart(O^{pr}, Y^{r}) \tag{10}$$

$$pU_{t} \leq Q, \ \forall t \tag{11}$$

By comparison with the two part mixed integer program of (Fairley, Scheinker et al. 2018) that has a similar objective of minimizing the squared PACU load, the advantages of more compact and intuitive CP formulations is clear, subject to some explanations for those unfamiliar with the special types of CP constraints used above. Note that the first five constraints apply to all patients, whereas the next three apply to all rooms excluding the PACU.

(2) specifies that decision intervals S^p, which represent each patient's total length of stay from start of procedure to end of recovery, must span their two occupancies, in the OR and PACU.

(3) specifies that the last (2nd) occupancy a patient's path must be in the PACU, numerically the $|R|^{th}$ member of R.

(4) specifies no overlap between a patient's occupancies in the OR and PACU.

(5) complements both equations (2) and (4) specifying further that the size of S^p should exactly match the sum of the patient's occupancies. A solution must have no delay between them.

(6) specifies that the decision interval U^p for each patient begin with PACU occupancy.

(7) specifies that intervals in sequences L^r cannot overlap and furthermore must be separated by at least a minimum clean time (for two patients with occupancies $O^{p_i r} \rightarrow O^{p_j r}$ in a sequence L^r.

(8) specifies that no room procedure can start before that room's block time opening B^r.

(9) specifies the latest any procedure can be expected to end is 20:00 pm (minute 1200.)

(10) requires that patient procedure scheduled starts are only at allowable times in Y^t.

(11) specifies a peak number of patients in the PACU at any time cannot exceed an integer amount Q. This is initially set to a high number to ensure that it will not be binding and thus to discover the unbounded peak PACU expected from a solution with the best objective function value. Q is lowered incrementally through successive iterations of the model until its lowest possible value (with a feasible schedule solution) is achieved.

3.7 Iterative Solution Approach

We use our model iteratively to generate a recommended schedule. Our iterative method works as follows.

1. Solve the problem with an overly generous PACU capacity limit Q^M and the objective of minimizing total patient flow time, which also translates into minimizing total OR makespans.
2. Determine the resulting expected peak PACU patient load, $\max\{pU^t\}$. Call this amount Q^U being the threshold below which capacity has a negative effect i.e. restriction on flow time optimization. Set $Q=Q^U$ before proceeding to the next step 3.
3. Invoke a PACU capacity constraint one less than the peak determined in the previous step, that is, set $Q = Q - 1$ and re-solve for the objective of minimizing total patient flow time, subject to: $\max\{pU^t\} \le Q$.
4. If a feasible solution is <u>not</u> found in the most recent step 3, accept the feasible solution found in the second-most recent step as the recommended schedule, and stop. (In effect, after making one too many progressively constrained solve attempts, return to the last successful one.)
5. Otherwise, a feasible solution was found in step 3, so repeat steps 3. and 4. until reaching a stop.

3.8 Discussion of Model Features

One of the features of our model is the creation of a binary matrix M representing whether a patient requires service in each of the rooms r. For each patient row of the matrix there will be two columns with entry 1, one for a specific OR and the other for their PACU stay. This allows us to define the set of occupancy intervals O^{pr} for all patient and room combinations, while setting the irrelevant ones as optional and of duration $.size = 0$.

Another is a set of intervals that serve as parameters for the block time windows that, despite being called a room in our model, may be one of multiple blocks within a single OR during a surgery day, each with its own time window W^r.

We also introduce three active and interrelated interval decision variables for each patient. One represents specifically their stay in PACU U^p. Two other intervals O^{pr} represent the patient's expected occupancies in two rooms, one of them being the same PACU stay and therefore exactly overlapping U^p. By defining two equivalent, parallel and concurrent intervals we can constrain them independently and/or in combination. For example, a patient's path includes two intervals beginning with their surgical procedure in an OR, which must be followed immediately by their recovery in the PACU. The procedure alone is part of an operating room's sequence whose intervals cannot overlap, whereas the recovery alone is part of the PACU load whose intervals are allowed overlap, as there are often numerous patients in the PACU concurrently.

The hospital schedules surgeries beginning at 08:00 am (minute 480 of the day) on every weekday except Wednesday, on which they begin at 09:00 am. The latest scheduled OR booking ends at 20:00 pm (minute 1200 of the day), although 18:00 pm is when most ORs are expected to close each day. Any schedule solution must have interval decision variables O^{pr} that lie within specified OR time windows W^r. The upper bound of T, midnight (minute 1440) is given to accommodate lengthy expected recoveries especially if arriving toward the end of the OR day (maximum 1200.) The extra length of T allows for extended PACU occupancies but has no impact on the model's objective of minimizing the average OR makespans.

The decision expressions F^p represent *flow time* for each patient. That is the length of time from opening of the OR in which their surgery is performed until the expected completion of their surgery, assuming specified start time for their procedure in a solution. We note that it is a simple change in the CP Model to calculate patient flow time using the expected completion of their recovery. However, our choice has important and more direct effect on the objective function. The primary effect is to produce a schedule following a shortest processing time (SPT) heuristic (Conway, Maxwell *et al.* 1967), all else being equal, that is OR sequences L^r are initially scheduled in order of ascending duration within each room. However, as we incrementally constrain the peak PACU load, some value in the objective function is foregone, to allowing shift of long surgery intervals forward and short surgeries backward within the room sequences, in order to eliminate PACU peaks.

Practically speaking, we are not concerned whether patients finish sooner or later on average in the day, as we understand that to them the day itself is/has been their only serious concern (with exceptions discussed below.) A more important effect of the objective function $\min \sum_p F^p$ is to minimize total flow time of patients as it is certainly important to them how many procedures can be performed in a given time, or conversely, in how little time a set of procedures can be performed, thus reducing the incidence of overtime and perhaps presenting opportunities for additional surgeries.

To accommodate a requirement that procedures be scheduled to start only at specific times, as is currently standard practice at LHSC, we created the step function Y^r. We first chose a start frequency of 30 minutes (which could be changed as desired), and a range of start time epochs from 0 to 23 (such that 30 times 23+1= 12 hours from the start of the surgery day.) We then generated an ordered set of tuples $\{< 0,480 >< 481,100 >< 0,510 >< 100,511 > ...\}$ which formed the basis of a step function Y^r which alternates from 0 (off) to 100 percent or 1 (on), every 30 minutes throughout the day. These tuples can be interpreted as 0 until 08:00 am, then 1 until 08:01 am, then 0 until 8:30 am, and so on. Constraint (9) prevents any OR occupancy $O^{pr}, r \neq |R|$ from

beginning whenever the step function has a value 0, therefore allowing it to begin only at one time (minute) every 30 minutes.

To account for minimum changeover times between procedures in each OR, that is to allow for cleaning, we first generated a set of triplets comprised of two patient identifiers, effectively the procedures ending and beginning in the OR changeover, followed by a constant minimum OR clean-up time $k = 5$. We chose this value despite the expected clean-up time requirement being between 10 and 20 minutes. We rationalize this in combination with the step function Y^r that allows scheduled starts only at 30-minute intervals. If one assumes a random uniform distribution of procedure end times over the minutes of any hour, imposing a minimum five-minute gap will result in a schedule gap of anywhere between 5 and 34 minutes, before any subsequent scheduled start (at :00 or :30.). Our assumption is that surplus and deficit clean times will offset each other in many cases, in the worst case causing only minor delays with no greater impact than already anticipated, stochastically, due to surgeries whose durations turn out to be unpredictably short or long. However, we recognize in the latter case that short scheduled clean-up times could compound problems from surgery time overruns.

We noted previously an alternate form for the objective function $\min \sum_p (F^p \times p.precedence)$. This followed an initial review of our model with hospital perioperative management, they requested that some additional factors be incorporated in the model. The first related to a problem of ward beds often becoming available only late in the day, such that it is preferable for SDA patient procedures to be scheduled later, especially for services whose wards are most notorious in this regard. Secondly, since there are often delays in preparation of very early ODS and SDA patients, as when they arrive late in the morning, preference should be given for inpatient surgeries to be scheduled in the early part of the day as they can be prepared quite early being already in the hospital. The management team also suggested that younger patients should have surgical priority early in the day for two reasons. The first is that young pediatric patients are less tolerant of long flow times from arrival to completion, and the second is that recovery times tend to vary in proportion with age, so that younger PACU patients first

means quicker bed turnover in the earlier part of the day for PACU, which is helpful for mitigating patient load surges during the later peak hours.

Based on these inputs, we added the final property (.precedence) to the patient tuple described above, and we developed a simple method for translating patient factors into a single ordinal variable, regardless of whether these factors themselves may be binary, categorical, interval or ratio in nature. For type of flow (ODS, SDA, IP) we begin by assigning the values (0,-1,1) respectively, because we would like to advance IP procedures to earlier in the day and defer SDA procedures later. However, for patient age, we count infants as 1 and patients 100 or older as -1, then for any patient we use (1-age/50) as a score in the range between (-1,1). Referring to the resulting flow and age scores as f and a, we determine an aggregate precedence score by the expression $\ln(e^{f+a})$ which then varies between [-2,2], although it could easily be normalized and/or weighted differently for each underlying factor. While hospital perioperative management could not yet commit to specific choices about how these factors should be calculated and so weighted, they acknowledged the potential of this method to apply a multi-criteria based objective in the model.

We note that a time interval for physical movements from the ORs to the PACU was not included in our model, as these are generally accomplished within a couple of minutes at LHSC, by physical design of the perioperative suite.

3.9 Results

Figure 4 depicts the results of the iterative CP model optimized solution in terms of expected patient loads in the PACU. It shows the average PACU patient loads by minute of the day across all sample days, expected if following the day's pre-schedule versus expected if following the CP optimized schedules on each day.

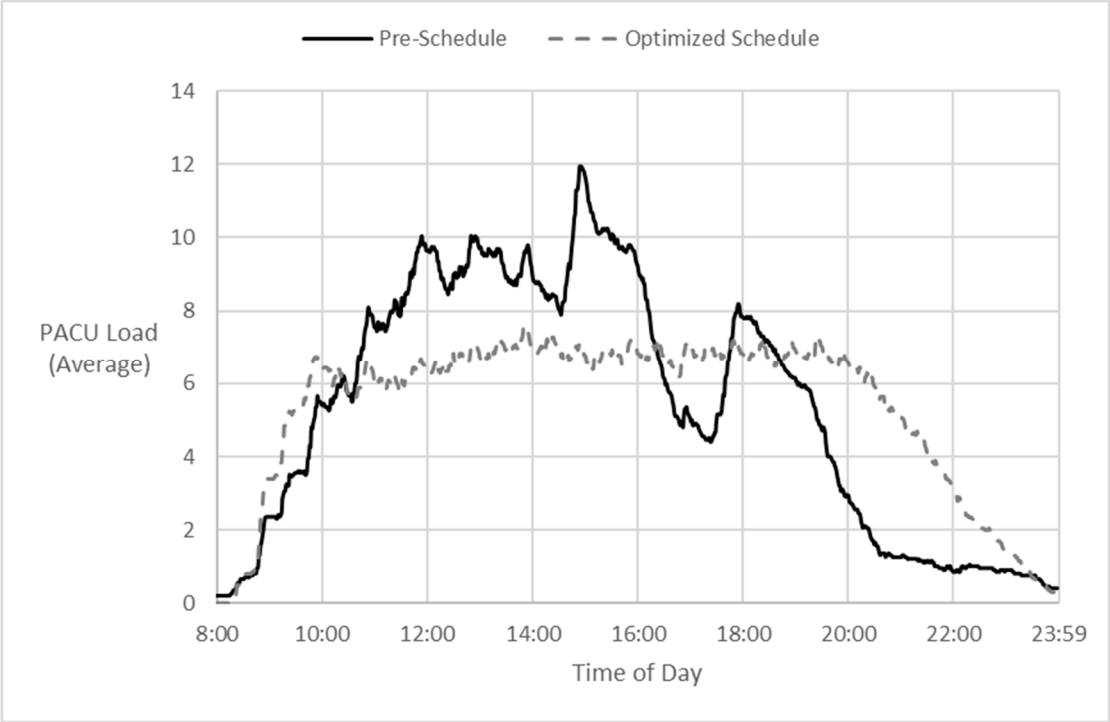


Figure 3-4 Pre-Schedule vs. Model - Forecast Averages: Patient Load by Time of Day

We modeled the expected peak PACU patient load given the established schedule for each day, using our estimated procedure and recovery durations, and assuming on-time starts. A surprising result was that in one-third of the days analyzed, the expected PACU peak from the 'Pre-schedule' was higher than what was expected from our unconstrained 'CP optimized' schedule. On only 40% of the days did the minimized patient flow time of the unconstrained CP optimized schedule impose an increase in the expected peak PACU load beyond what it was already with the pre-schedule. Yet in those cases the CP model could reduce expected patient flow time by an average of 15%. Overall, our model achieved an average 11% reduction in patient flow time at PACU load peaks that were equal to or less than those expected by following the pre-schedule, similar to the particular case depicted in the right side of Figure 5.

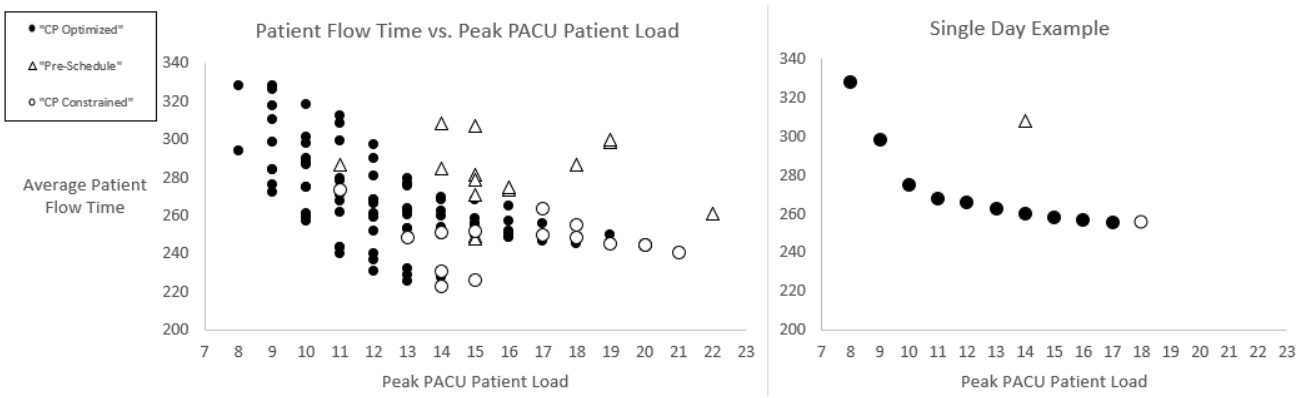


Figure 3-5 Efficient Frontier: Average Patient Flow Time vs. Peak PACU Patient Load

Figure 5 displays two scatterplots of average patient flow times versus peak PACU patient loads as predicted by our model. We show averages rather than the total flow time objective to normalize days with different numbers of patients. Comparison is made between pre-schedules as established by hospital surgeons, and CP optimized schedules as proposed by our model. On the right side of Figure 5 is an example of a single day for which the lowest average patient flow time of 260 minutes is expected from the unconstrained Q, 'CP Optimized' schedule (versus 310 minutes with the 'Pre-Schedule'.) However, that schedule is expected to result in a peak PACU load of 18 patients at some point in the day. In contrast, the pre-schedule is expected to result in an average patient flow time nearly 20% higher (worse), but with a lesser (better) peak load of 14 patients in the PACU. Solid circles in Figure 5 represent a series of intermediate schedule solutions that constitute the frontier of lowest possible patient flow times and PACU peak loads. We obtained these by successively lowering K in the $pU < K$ constraint (10), as described in the Model Development section above.

We see in Figure 5 that a load equal to the pre-schedule expectation (14 in this example) is achieved with little increase in the average patient flow time (only 10 minutes higher than for the unconstrained Q optimized solution.) Furthermore, a one-third reduction in peak PACU patient load is apparently achievable with resulting average patient flow time that is equal to that expected with the pre-schedule. A much greater (~45%) reduction in PACU peak load appears to be achievable with only a modest increase in patient flow time. The left side of Figure 5 shows a similar trend over our sample of testing days.

We conducted an *ex post* analysis using actual procedure and recovery durations in both the Pre and CP optimized schedules. Figure 6 below shows two paired charts of PACU loads for a typical surgery day. The left side refers to the pre-schedule, and the right side refers to the CP optimized schedule. On both charts, a *solid line* depicts the expected count of PACU patients as *predicted* by our model. The *dashed line* on the left depicts *actual* events (considering only scheduled patients) in the PACU that day. The dashed line on the right is slightly different as we can only simulate what would have happened if the CP optimized schedule had been adopted – using actual durations for both procedure and recovery, but with different timing.

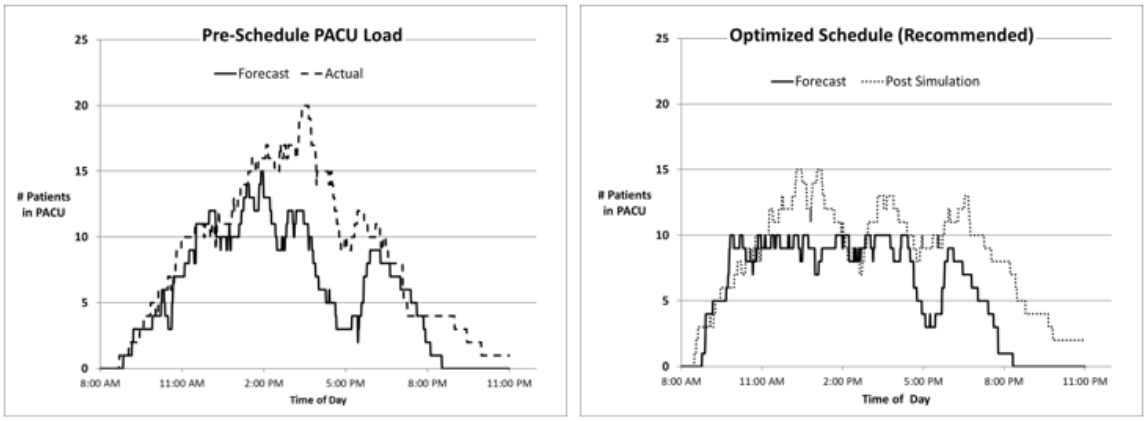


Figure 3-6 PACU Loads: Pre-Schedule & Optimized - Forecast vs. Actual/Calculated

3.10 Visualization Model

Unique durations were previously observable in the patient flow timelines of Figure 2 however, the nine discrete examples of flow-delay shown in that figure represent fewer than half of the possible permutations when also considering actual versus scheduled procedure start times. (These features are included in the visualization model provided to the hospital, but not in Figures 2 or 7, due to the complexity of depicting these additional elements in gray-scale.) As one example, we found several instances of OR delays in moving patients into the PACU that were arguably a result of the procedure starting earlier than scheduled, since these delays would not have occurred and would not have

been recorded had the surgery started at the scheduled (later) time. We mention the latter point to emphasize that the assignment of responsibility for delays is not always straightforward, and that best outcomes are achieved when parties commit to specific timing of their activities (Millstein and Martinich 2014).

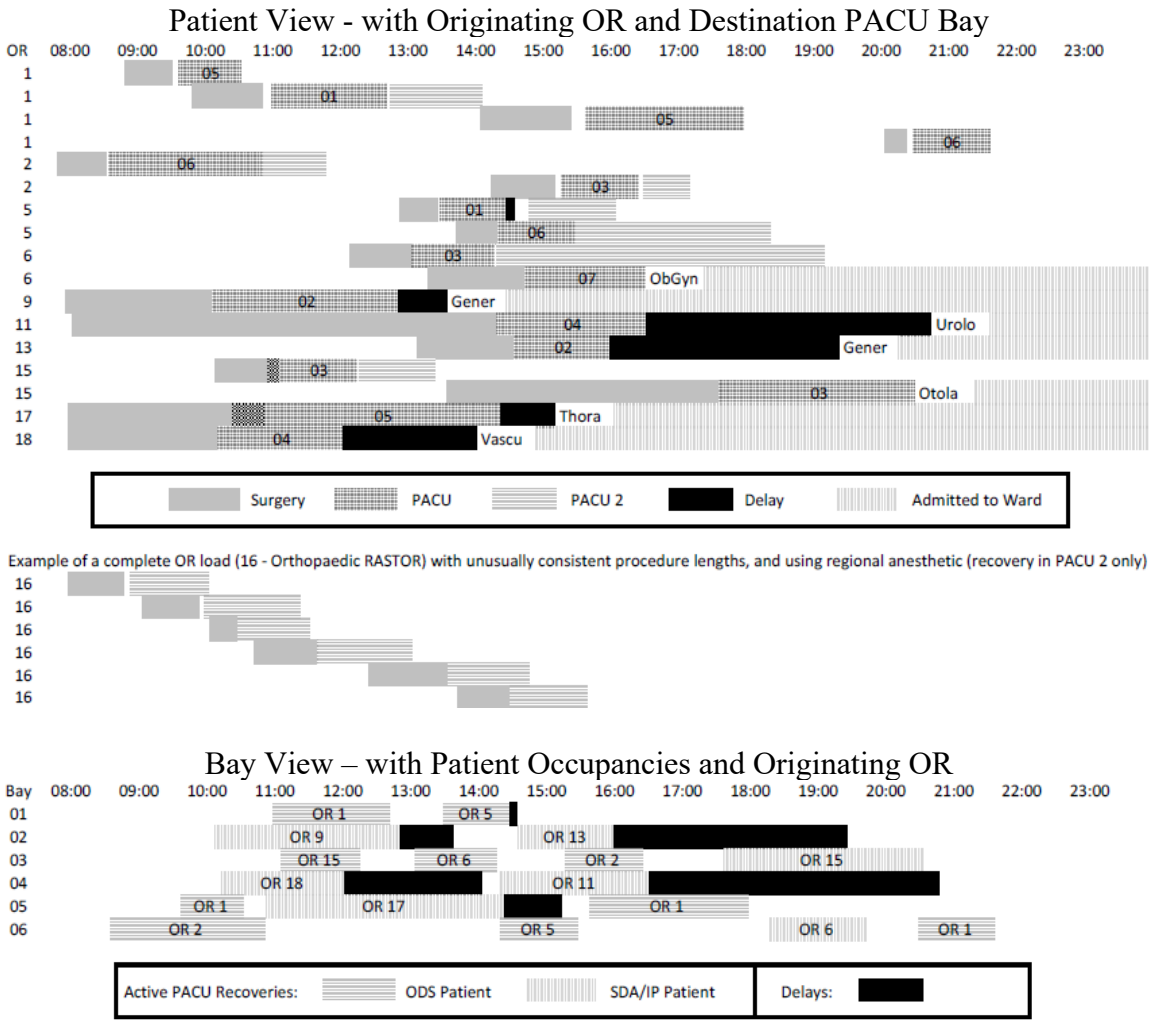


Figure 3-7 Patient Flow Visualization: Patient-OR View and PACU-Bay View

The aim of the patient flow visualization is to provide a systematic view of the process *in spite* the uniqueness of a great many cases across many ORs. Effective management requires clear understanding of both 'the forest and the trees', and appropriate aggregated levels between these extremes. The top of Figure 7 shows timelines for 16 patients passing through the first six PACU Bays shown at the bottom. They represent roughly a quarter of scheduled procedures on a sample day. By cross-referencing the two views, it

is straightforward to trace delays in PACU to a patient's destination ward (or PACU2), and to match OR delays with concurrent activity in the PACU.

The final six cases shown in the top of Figure 7 are examples of ODS patients who did not pass through the PACU, but rather had only minor recovery in PACU2. It happened on this sample day that OR 16's caseload was a series of similar orthopaedic procedures (a RASTOR room as described earlier), all performed under a regional anesthetic rather than a general one and therefore not requiring PACU recovery but PACU 2 only. This example reinforces the value of characterizing ORs by their case loads to understand the impact of these loads on the PACU i.e. the central stage of the system.

Whereas some OR blocks contain homogenous, predictable procedure lengths and flow patterns, others involve procedures that vary greatly in duration and patient type. These provide more opportunities to affect a different PACU load by re-sequencing cases, whether using something like our CP optimization model, or heuristics such as longest or shortest durations 'first', and/or SDA patients and oldest ones 'last'. The incremental PACU impacts may vary markedly depending on the portion of patients receiving a regional versus general anesthetic (although when both are practical the choice is subject to patient agreement and that is not always predictable.)

3.11 Other PACU Considerations

Although some perioperative units have explored the practice of pre-assigning PACU bays (Dexter 2007) prior to patient arrivals, this was not under consideration by LHSC at the time of our study. Management felt strongly that bay assignments were best done dynamically based on actual arrivals, especially considering the unpredictability of either arrival times or recovery lengths. Moreover, certain rules applicable to PACU nursing operations that would need to be incorporated in any assignment algorithm. PACU bays were opened in pairs, with a single nurse assigned to both bays, for example 07 and 08. A nurse should not receive two new patients within a span of 30 minutes. Also, a nurse should not have two SDA or IP patients at once, although one of either in addition to an ODS patient was acceptable (given that the latter were typically less complicated patients.) Some pairs of PACU bays were reserved for paediatric patients, with very

young patients requiring a fully dedicated nurse, sometimes with an additional nurse's assistance.

Nevertheless, the PACU 'Bay View' of the patient flow visualization reveals important facts about the efficiency with which bays are utilized. For example, in the bottom of Figure 7, the last patient assigned to PACU bay 05 and the last two patients in bay 06 could have all been accommodated in bay 01, thereby allowing the third pair of bays (05 and 06) to be closed seven hours earlier by keeping the first pair of bays (01 and 02) open only two hours later. The four instances of extended wait for ward beds in PACU bays 02 and 04 should also warrant investigation, as the early vascular surgery SDA patient in OR 18 might better have been scheduled later in the day.

3.12 Discussion

The perioperative process includes a high degree of uncertainty and substantial inherent variety, consequently there is no magic solution to ensure the smoothest and most efficient patient flow, particularly through the centralized PACU stage. However, despite the uncertainties, we believe much of the case variety that directly impacts timing and flow can be captured in modeling and optimization techniques that are reasonably accessible to practitioners and use operational data that is already being used in heuristic scheduling approaches.

Our choice of patient flow time as the performance metric (objective) for our CP optimization model may be challenged on the notion that patients who are waiting months for surgery might not care about how long a procedure takes on the day of surgery. However, this is a commonly applied metric in *static* scheduling problems when it is assumed, as in our case, that all jobs arrive simultaneously at the beginning of the scheduling period. Furthermore, by choosing as objective minimizing the total patient flow time we are at the same time maximizing global OR utilization i.e., by completing the given set of procedures within the least overall OR time.

Whether individual surgeons who control their own OR schedules can be convinced of the benefits of a more collective scheduling approach remains to be seen. Autonomy is a difficult thing to forego without a clear understanding of, and/or certainty about, the personal impact of allowing someone else to decide the order in which patients should be treated. Surgeons have individual preferences, such as whether two similar procedures should be sequenced contiguously or not, and whether longer procedures are better performed earlier or later in the day. As a modeling platform, CP lends itself to the addition of constraints incrementally and with relative ease, as these specific surgeon needs arise.

The models we developed enable, within roughly one hour, translation of schedules (in PDF, portable document format) into data input files for the CP model, running of the model and output of a proposed alternate CP optimized schedule for review, including for each procedure the direction and length of time of any proposed change in scheduled start time. In some cases, only a handful of changes in a few ORs are enough to reduce a projected peak in the PACU load by one or two patients. It was left to the perioperative management team to negotiate sequence changes with surgeons where they found them most compelling. A dominant factor in these decisions was whether a sequence change supported the precedence of patient types (young people will be treated earlier, and IP patients precede SDA patients) without blocking upstream patients from entering the PACU. It is for this reason that we provided the option of incorporating a precedence weighting factor to individual patient flow times in the alternate objective function (1b), which we note can be easily tuned according the relative importance of age versus flow, and other factors can also be added. The aim was to find a balance between the CP model recommendations based mainly on procedure times with qualitative factors of importance in patient sequencing.

Rapid translation of a pre-schedule to a CP optimized schedule enabled management to spot improvement opportunities more quickly (moreover, to justify them as having been proposed objectively), as opposed to managers having to scan lists of five dozen procedures to identify these opportunities. Similarly, on the day following surgery as actual time-stamp data became available for extraction, summary reports including the

patient flow visualizations like Figure 7 were constructed. These reports facilitated timely and efficient reflection on the events of the previous day and allowed managers to identify and communicate where and when problems occurred, decisions that turned out to be positively or negatively impactful, and to assess the accuracy with which estimates had been made. The patient flow visualizations provided to the hospital include identifiers within the timelines that depict the model's predictions of procedure times and durations (and where absent, indicating unscheduled cases whose impact could be assessed.) In many cases the time predictions proved to be reasonably accurate, even when shifted in time due to other events. A rationale was sought in cases where time predictions proved to be inaccurate, including questions regarding how the predictive models could be improved. We note that aside from the CP model which was developed in and for use with IBM ILOG CP Optimizer, other model components were developed in Excel VBA (Visual Basic for Applications), which was readily available to hospital management and staff and thereby facilitated implementation.

3.13 Opportunities for Further Research

We are not aware of any previous attempts to model the problem of assigning PACU bays while respecting constraints such as one-nurse-per-two-PACU-bays, described in the previous section. We believe that these factors could be incorporated in an extended CP model, although doing so would amplify complexity without necessarily bolstering confidence in the accuracy of recovery time predictions. If such a model were to be dynamic, and applicable in or near 'real-time' using updated data to plan PACU bay assignments forward in the day (subject to change), that could certainly enable higher occupation and patient flow, and thus the possibility the of running multiple RASTOR rooms on some days, and thereby address the surgery backlog and wait-time problems in the system.

One of the challenges in developing solutions to perioperative patient flow problems is they are often setting-specific. LHSC chooses to schedule procedure start times at :00 and :30 whereas another works in five-minute intervals. LHSC Victoria campus does a lot of orthopaedic and obstetrics-gynecology and general and paediatric procedures whereas its University Hospital campus handles a higher portion of neurology and cardiology

patients. The compatibility of such different surgery block cases warrants exploration, including an assessment of the value of simple heuristics to the benefit of the system in whole.

3.14 Conclusion

In this paper, we have extended the view of the perioperative process to distinguish between patients according to a variety of paths they follow, as in-patients, one-day surgeries and same day admissions. We have also introduced a constraint program to develop coordinated OR schedules aimed at minimizing peak patient loads in the PACU, to better ensure that OR delays will not be incurred due to the PACU reaching full capacity and being unable to accept new patients from the ORs.

The model required a mechanism to predict procedure and recovery times and these are open to further refinement through improved data collection and methods such as machine learning. We have developed and described a model for patient flow visualization which can help perioperative care managers and surgeons to quickly locate problems in time and stage of the process, to better understand interactive effects of schedule sequence decisions, to propose additional practical factors that can be incorporated the optimization model, and to provide a simple, common foundation for ongoing review and refinement of OR scheduling practices.

Our model has contributed to organizational learning and improved communication and cooperation among the many parties involved in scheduling the ORs at LHSC. In a letter supporting our research, the management team wrote: “The Ivey research team has provided us important insight to the possibilities and significant potential benefits of adopting a more methodical and coordinated approach to daily OR scheduling. We look forward to the next phase of development in what we believe can be a valuable and implementable tool in streamlining the costs and timeliness of surgical operations, not only within LHSC but throughout the health care system, if successful. In the meantime, we strongly support not only the research team’s continued work on our behalf, but also the sharing of their methods and conclusions (ongoing as they are) with members of operational research community focused on similar health care challenges.”

3.15 Acknowledgement

The authors would like to thank the perioperative care management and staff of London Health Sciences Centre (LHSC) Victoria Hospital campus for the support and collaboration in this research.

3.16 References (Chapter 3)

Abedini, A., W. Li and H. Ye (2017). "An Optimization Model for Operating Room Scheduling to Reduce Blocking Across the Perioperative Process." Procedia Manufacturing 10: 60-70.

ASA. "Physical Status Classification System." Retrieved 30 May 2019, from https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system.

Bam, M., B. T. Denton, M. P. Van Oyen and M. E. Cowen (2017). "Surgery scheduling with recovery resources." IISE Transactions 49(10): 942-955.

Blake, J. T. and M. W. Carter (1997). "Surgical process scheduling: a structured review." Journal of the Society for Health Systems 5(3): 17-30.

Cardoen, B., E. Demeulemeester and J. Beliën (2010). "Operating room planning and scheduling: A literature review." European Journal of Operational Research 201(3): 921-932.

Conway, R. W., W. L. Maxwell and L. W. Miller (1967). Theory of scheduling, Addison-Wesley Pub. Co.

Dexter, F. (2007). "Bed Management Displays to Optimize Patient Flow the OR to the PACU." Journal of PeriAnesthesia Nursing 22(3): 218-219.

Fairley, M., D. Scheinker and M. L. Brandeau (2018). "Improving the efficiency of the operating room environment with an optimization and machine learning model." Health care management science: 1-12.

Guerriero, F. and R. Guido (2011). "Operational research in the management of the operating theatre: a survey." Health Care Management Science 14(1): 89-114.

Gür, Ş. and T. Eren (2018). "Application of Operational Research Techniques in Operating Room Scheduling Problems: Literature Overview." Journal of healthcare engineering 2018: 5341394-5341315.

Larsson, A. (2013). "The accuracy of surgery time estimations." Production Planning & Control 24(10-11): 891-902.

Millstein, M. A. and J. S. Martinich (2014). "Takt Time Grouping: implementing kanban-flow manufacturing in an unbalanced, high variation cycle-time process with moving constraints." International Journal of Production Research 52(23): 6863-6877.

Nilsson, U., M. Jaensson, K. Dahlberg, K. Hugelius, h. Institutionen för and u. Örebro (2019). "Postoperative Recovery After General and Regional Anesthesia in Patients Undergoing Day Surgery: A Mixed Methods Study." Journal of PeriAnesthesia Nursing 34(3): 517-528.

Ontario, H. Q. (2019). "System Performance - Wait Times for Surgeries and Procedures."

Santibgáñez, P., M. Begen and D. Atkins (2007). "Surgical block scheduling in a system of hospitals: an application to resource and wait list management in a British Columbia health authority." Health Care Management Science 10(3): 269-282.

Wang, T., N. Meskens and D. Duvivier (2015). "Scheduling operating theatres: Mixed integer programming vs. constraint programming." European Journal of Operational Research 247(2): 401-413.

Wright, I. H., C. Kooperberg, B. A. Bonar and G. Bashein (1996). "Statistical modeling to predict elective surgery time. Comparison with a computer scheduling system and surgeon-provided estimates." Anesthesiology **85**(6): 1235-1245.

Chapter 4

4 The Effect of Revenue Versus Profit Maximization on Firm Profits

4.1 Abstract

This paper addresses the impact on profits from the use of revenue maximizing heuristics. We consider the problem of choosing optimal prices and quantities in two markets, each having its own demand characteristics and cost structure of service. We explore conditions in which pricing for revenue maximization (RM) leads to worse profit outcomes than pricing aimed directly at profit maximization (PM). We consider three types of cost: a per unit manufacturing (or procurement) cost, a per unit transportation or delivery cost specific to each market, and a percentage of selling price or sales commission for each market.

Using linear price-response functions for each market we derive and compare optimal prices and supply quantities under the two objectives (RM, PM). Beginning with independent markets and deterministic demand without a capacity constraint we progress to the problem of dual markets with constrained supply and stochastic demand. We reveal interesting patterns in closed-form solutions derived for all but the stochastic constrained case, for which we propose instead a heuristic method that can be modeled and solved Microsoft Excel. We conclude with an illustrative example of these various results and their interrelationships.

We find that PM prices are always greater than RM prices in unconstrained markets, but the reverse may be true under a capacity constraint to sustain demand in the less costly to serve and less price sensitive market, also the higher-priced and more profitable one in our numeric example. We find that percentage of sales costs play an interesting role in combination with optimal unconstrained prices for determining PM prices at any constrained capacity level. The difference in profits from PM vs. RM is greatest when neither problem unconstrained, that occurring first for the RM problem. This brings RM outcomes, although not equally individual market decisions, closer in line with PM outcomes where capacity becomes a binding also on the latter problem. Stochastic

revenue maximization extensions that take unit production cost into account can narrow the expected profit difference in the unconstrained case, but not accounting for other cost types increases it under a capacity constraint.

4.2 Introduction

This article focusses on solutions to a fundamental and practical pricing and inventory problem, that of a firm selling a common product through multiple markets, and thereby needing to choose optimal prices and quantities for each of those markets. (Mankiw, Kneebone et al. 2011) provide an extensive review of this research area while Phillips (2005) provides a management-oriented overview of the associated scientific pricing concepts.

The problem exists in many different settings. Walmart provides the example of a physical store retailer with a parallel online sales channel with different costs of product shipping and handling, as well as selling. Airlines, hotels, and rental car companies generally sell through their own online channel and other online channels like Expedia. These channels involve different costs of sale and/or delivery and/or product costs (e.g. bundled benefits.) Many product categories include direct sales complemented by some sort of self-service channel for experienced customers.

There has been abundant research on maximization of revenues across channels with different demand characteristics. A similar bounty of research has been concerned with cost efficiency of activities to produce and deliver product to different markets, given demand, which is of course driven by pricing. (Bell and Chen 2017) describe these as two different protocols for the allocation of decision responsibility for costs, one being a Revenue Management (RM^2) protocol and the other a Supply Chain Optimization (SCO) decision protocol. They observe that there is broad recognition that the use of either SCO or RM is sub-optimal, and despite some well-published attempts, integration of SCO and

[2] We use RM either for revenue maximization methods, or revenue management practice/protocol as in this case.

RM has proven difficult to accomplish technically and has also run into management difficulties due to departmentally siloed responsibilities and incentives. What (Bell and Chen 2017) refer to as an integrated SCO-RM decision protocol we refer to as Profit Maximization (PM).

The motivation for our research is to enlighten the conversation about when is revenue management sub-optimal in terms of profit maximization. How dependent is the PM versus RM profit difference on cost factors and/or channel demand characteristics? How should optimal prices and quantities be determined in a given circumstance, and how great is the difference in profits from using RM price and quantity decisions versus PM decisions?

In this paper we develop solutions to a deterministic problem and incrementally extend them highlight important relationships between price and quantity decisions for different objectives i.e. under different decision protocols, for different firm cost structures and market characteristics, and we show how these decisions change if product capacity is constrained.

Optimal decisions are presented in deliberate patterns that informs comparison between different scenarios and between markets in terms of the problem parameters i.e. component costs and market characteristics. These formulae together with heuristic methods we provide can be easily applied by managers to their own business circumstances, or perhaps consider the extent to which their current pricing practices may warrant a re-calibration.

4.3 Background and Related Literature

Revenue management (RM), or scientific pricing, as a business function traces back to the airline industry of the 1970s (Littlewood 1972). On the heels of highly successful implementations reported by American Airlines (Smith, Leimkuhler *et al.* 1992) over the decade that followed, the practice began to gain acceptance across a broad range of industries, including automotive, retail, telecommunications, financial services and manufacturing. Today, as virtually every business has unprecedented access to customer

and transactional data to inform pricing decisions, the practice is no longer only the purview of large corporations, but medium and even small ones. In parallel with cost efficient operations to produce and deliver product, the RM principle is simple enough: focus attention on realizing as much expected revenue as possible from customers. Proponents of RM argue that because the benefit of realizing higher prices (or higher volumes at lower prices) flows directly to the bottom line, implementation of scientific pricing is often one of the highest return investments available to a company (Phillips 2005).

Recognizing this opportunity, a large community of software and service providers has grown into USD $20 billion industry representing roughly one-quarter of global investments in enterprise resource planning (ERP) systems overall. Moreover, even small companies are being enabled by an emergence of several cloud-based platforms and solutions for pricing and revenue management (ReportsnReports.com 2019).

Fueling the development of commercial pricing and revenue management solutions is a rich field of academic literature on the subject including nearly 900 articles during 17 publication years of the *Journal of Revenue and Pricing Management*, in addition to considerable coverage in other leading operational research journals. Some of the many sub-speciality areas that have been studied include airline and hotel yield management and overbooking, retail markdown management, and auctions. The subject matter of this article follows most directly from the convergence of inventory control and price theory as described in a seminal article of that name (Whitin 1955). The latter half of the paper deals with demand uncertainty in price and quantity decisions, relating to the well-known 'newsvendor' problem for which (Petruzzi and Dada 1999) remains among the most important contributions.

Our specific concern stems from the fact that RM pricing science emerged and is most practiced in industries with high capital costs, low variable costs, and time perishable inventory where incremental revenue flows quite directly to net income. A key assumption in the RM literature is that all costs are fixed or are sunk and consequently the cost side of the firm is considered independent of pricing decisions. (Bell and Chen

2017) point out that the assumption that costs are not affected by price changes may be tenuous in multi-channel distribution systems or for firms serving multiple markets (hereafter "channels"), since changing relative prices across channels will result in sales migrating between them. Excluding costs from pricing decisions is to assume that cost of sales is identical in all channels, whereas clearly sales commissions and distribution costs are channel-dependent.

Other articles related to dual channel pricing include (Lovell and Wertz 1981), (Karakul 2008), (Yue and Liu 2006), (Zhang, Bell et al. 2010) , (Chen, Fleischhacker et al. 2015), (Xiao and Shi 2016) and (Huang, Ding et al. 2018). However, these studies focus on different settings and considerations such as price fencing between markets, supply chain dynamics, and multi-period decisions for revenue maximization, without regard for market or channel specific delivery and selling costs.

Our research integrates these cost factors into the single period price and quantity decision problem. Its objective is to isolate the impacts of these factors on optimal solutions and expected outcomes. The traditional emphasis of the RM literature and practice has been on top-line revenue performance, whereas our aim is to explore the potential bottom-line profit benefits of a more integrated approach to pricing in distinct markets.

4.4 The General Model

4.4.1 Parameters and Notation

We assume two channels with linear demand functions. We think of one channel being the "home" channel (for example, the Air Canada or Marriott hotel website) and the second channel being a proxy for a set of secondary channels (for Air Canada or Marriott this could be sites such as Expedia and Travelocity) that can be modelled as a single aggregate demand function and individually identical cost structures from our supplying firm's perspective.

Define, $d_i(p_i) = a_i - b_i \cdot p_i$, where $i = 1, 2$ and $d_i(p_i)$ is demand in market i at price p_i, and a_i and b_i are channel specific parameters. Parameters a_i are a measure of market

sizes, the numbers of customers who would demand the item at $p_i = 0$. The parameters b_i represent the reduction in number of customers demanding product with each unit increase in price, effectively price elasticity in each market. The primary home channel is typically less price elastic than the secondary channel.

Given two such markets or channels through which a firm sells its product and given the associated cost structures, a firm must choose how much to produce and at what price to sell in each market.

We define the decision variables p_i, q_i for $i = 1, 2$ as the prices and quantities to supply to the two channels. With deterministic demand, pricing decisions imply a specific demand as a function of price $d_i(p_i)$, leading to stocking quantity decisions q_i of the same amount.

$$a_i - b_i p_i = d_i(p_i) = q_i, \text{ for } i = 1, 2 \quad (1)$$

We define the firm's cost structure as follows:

m_i the cost to manufacture each unit of product that will be sold in channel i.

t_i the cost of transporting or delivering each unit of product to channel i.

s_i the selling cost or commission in channel *I* as a percentage of selling price in that market or sales channel with $0 \le s_i < 1$.

We assume that $m_i,\ t_i,\ s_i \ge 0$

Table 4-1 defines four types of objective functions according to which of the above costs are considered in the maximization problem. Our aim is to compare the results of optimizing price and quantity decisions for maximizing revenue $R(\cdot)$ versus maximizing profit $\Pi(\cdot)$. Where appropriate we also consider two other intermediate objectives or decision protocols besides RM and PM. The first we call Contribution Maximization (CM), with an objective outcome denoted by $\chi(\cdot)$, and the second we call Net Sales Maximization (SM), with an objective outcome denoted by $\gamma(\cdot)$.

Decision Protocol and Related Objective		Costs Considered
RM	Revenue $R(\cdot)$ Maximization	Does not consider any costs.
PM	Profit $\Pi(\cdot)$ Maximization	Considers all costs m_i, t_i, s_i
CM	Contribution $\chi(\cdot)$ Maximization	Considers only product cost m_i
SM	Net Sales $\gamma(\cdot)$ Maximization	Considers only selling percentage cost

Table 4-1 Pricing Decision Protocols and Objectives

CM is a natural choice for a manufacturer or wholesaler allocating demand to two channels under freight terms F.O.B. origin, wherein the channel receiving the goods is responsible for transportation and selling costs hence the only cost relevant to the supplier is the manufacturing cost (m_i). SM is a natural choice for a sales agent or agency without responsibility for costs of production or delivery, and for whom only the sales commission percentages (s_i) may be relevant to the pricing decision.

Various scenarios and decision protocols are analyzed in this article where each can be assessed in terms of any one of the objective functions $R(\cdot)$, $\chi(\cdot)$, $\gamma(\cdot)$, $\Pi(\cdot)$, however our emphasis is primarily on the first and last of these, being driven by RM and PM, respectively.

The following subscripts, superscripts and accents are used in this article.

An asterisk * denotes an optimal decision (or function evaluated at optimal decision values) and may be combined with other accents.

> An asterisk with no other accent identifies a solution or outcome in the most basic decision scenario, that is RM to maximize revenue (through optimal decisions p_i*) where demand is considered deterministic and there is no capacity constraint.

A bar accent $\bar{\cdot}$ denotes a PM decision when all costs (m, t and s) are considered.

For example, $\overline{p}_i^{\,*}$ is a profit maximizing price in market i for the deterministic demand case.

A circumflex accent $\hat{\cdot}$ denotes a CM decision where only the cost m_i is considered.

We note that $\Pi(\cdot) = \chi(\cdot)$ if $s = t = 0$.

A double-dot accent $\ddot{\cdot}$ denotes a SM decision where only the sales commission or percent-of-sales cost component s_i is considered.

We note that $\Pi(\cdot) = \gamma(\cdot)$ if $m = t = 0$.

A tilde accent ~, either alone in the revenue maximization case or compounded with other accents within other decision scenarios, identifies a stochastic decision.

As a superscript, k denotes a variable or function associated with a capacity constrained scenario. As a parameter, k represents capacity in units.

Similarly, as a superscript, u identifies a variable or function as being associated with an unconstrained scenario.

4.5 Deterministic Model

4.5.1 Independent Markets, Unconstrained

Initially we assume no capacity constraints and we seek to simultaneously determine optimal prices and quantities in each market.

4.5.1.1 Revenue Maximization (RM)

Solutions to maximize revenue are derived from the objective function:

$$\underset{p_1,p_2}{Max}\{R(p_1,p_2)\} = \sum_{i=1,2} p_i \cdot (a_i - b_i p_i) \tag{2}$$

Optimal decisions for this deterministic, revenue optimization problem are well-known:

$$p_i^* = \frac{a_i}{2b_i}, \quad q_i^* = \frac{a_i}{2} \tag{3}$$

We note that in these solutions as in many others that follow the factor ½ arises in optimization of an objective function involving $p \times q(p)$. For equation (2) which is our fundamental (RM) reference case, the optimal solution involves pricing with the aim of selling to ½ of each market, each at prices that are ½ the ratio of that market's size to its price sensitivity parameter b_i.

4.5.1.2 Profit Maximization (PM)

The deterministic profit maximization problem has the following cost function:

$$C(p_1, p_2) = \sum_{i=1,2} (m_i + t_i + p_i s_i) \cdot q_i = \sum_{i=1,2} (p_i s_i + m_i + t_i) \cdot (a_i - b_i p_i) \tag{4}$$

In this case, profit maximizing price decisions are solutions to the problem:

$$\begin{aligned} \underset{p_1, p_2}{Max}\, \Pi(p_1, p_2) &= R(p_1, p_2) - C(p_1, p_2) \\ &= \sum_{i=1,2} [(a_i - b_i p_i) \cdot p_i] - [(a_i - b_i p_i) \cdot (p_i s_i + (m_i + t_i))] \\ &= \sum_{i=1,2} (a_i - b_i p_i) \cdot (p_i (1 - s_i) - (m_i + t_i)) \end{aligned} \tag{5}$$

The optimal prices and quantities in this PM problem are as follows:

$$\bar{p}_i^* = \frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1 - s_i)}, \quad \bar{q}_i^* = \frac{a_i}{2} - \frac{b_i (m_i + t_i)}{2(1 - s_i)} \tag{6}$$

Remark 1: We note that $\bar{p}_i^* \geq p_i^*$ and $\bar{q}_i^* \leq q_i^*$, assuming $m_i, t_i, s_i \geq 0$. Later in this paper we show that these inequalities hold as well between PM and RM prices and quantities for the stochastic unconstrained scenario.

From equations (3) and (6) we see that $p_i^* = \bar{p}_i^* \Leftrightarrow m_i + t_i = 0$, that is, revenue maximizing prices equal profit maximizing prices if $m_i = t_i = 0$ even if $s_i > 0$. We also note that price increases for profit maximization should be greater in whichever market is more costly-to-serve, that is, whichever has the larger $(m_i + t_i)/(1 - s_i)$. Revenue maximization, conversely, involves setting prices below their profit maximizing

counterparts to induce greater demand. As the PM to RM price difference is not dependent on the *market* parameters that determining prices in proportion to $\frac{a_i}{b_i}$, RM may induce greater demand by undue price reduction in the least profitable channel (were costs to be taken into consideration) whereas PM avoids this pitfall.

4.5.2 Profit Loss from Unconstrained RM vs. PM

Substituting revenue and profit maximizing prices (3) and (6) into the revenue and profit functions (2) and (5) leads to the result in proposition 2.

Remark 2: The difference between profits under PM ($\bar{\Pi}$) and profits under RM (Π) for an unconstrained market without a capacity constraint is given by:

$$\Delta\Pi_i = \bar{\Pi}_i - \Pi_i = \sum_{i=1,2} \frac{b_i(m_i + t_i)^2}{4(1 - s_i)} \tag{7}$$

This difference is additive for each market or channel when there is no overall constraint on supply. The profit loss does not depend on the market size parameter a_i but can be significant if there is: high price sensitivity b_i; a substantial cost of procurement/manufacture and delivery to the market $m_i + t_i$; and/or significant sales commissions s_i. However, if unit manufacturing and delivery costs $m_i + t_i$ are small, sales commissions alone result in little profit difference, between revenue or profit maximization.

Equation (7) helps to explain the prevalence of RM in industries with high net revenues (small m_i) even if requiring high incentives (s_i) for net sales to cover large fixed cost i.e. capital investment. The effects of equation (7) are difficult to generalize, for example, to a business where unit production and delivery costs represent a high portion of selling price. The absence of the market size parameter a_i in equation (7) means that in a large market with low prices (due to high price sensitivity b_i) the profit difference depends critically on unit costs. When these costs are small in absolute terms, even if comprising a large portion of the product's price, there will be little difference in profit as the term $(m+t)^2$ will be small. On the other hand, if these costs are small relative to price but

large in absolute terms, as in a large market with high prices (low b_i), the difference in profit may be large enough to be worthy of attention and correction.

4.5.3 Intermediate Objective Functions

A general manager should understand the effects of distinct market demand characteristics and cost structure on overall profit when choosing prices. But at lower levels of the organization the objective associated with price and quantity decisions may vary. A production manager might concern themselves primarily with CM, considering only the difference between revenue and cost of product, regardless of delivery or selling costs. On the other hand, a sales manager might be concerned with maximizing net sales i.e. SM, that is concerned only with commissions, regardless of product or delivery costs.

We note that CM and SM are special cases of PM when the only costs considered are m_i or s_i respectively, but analysis of the three generalizes to all other combinations $(t_i>0,\ m_i=s_i=0)$, $(m_i=0,\ t_i>0,\ s_i>0)$, $(m_i>0,\ t_i=0,\ s_i>0)$, $(m_i>0,\ t_i>0,\ s_i=0)$.

4.5.3.1 Net Sales Maximization (SM)

Net Sales maximization is identical to PM when in (6) we have $m_i = t_i = 0$. The resulting optimal decisions are identical to RM solutions in (3).

Remark 3: Optimal price $\ddot{p}_i^*$ and quantity $\ddot{q}_i^*$ decisions for net sales maximization are identical to optimal solutions for revenue maximization in the deterministic demand case: $\ddot{p}_i^*=p_i^*$ and $\ddot{q}_i^*=q_i^*$. (8)

We arrive at this conclusion beginning with a cost function (which is absent from the pure revenue maximizing RM calculation) that includes sales commissions as the only cost component:

$$\ddot{C}(\ddot{p}_1,\ddot{p}_2)=\sum_{i=1,2}\ddot{q}\cdot\ddot{p}_i\cdot s_i=\sum_{i=1,2}(a_i-b_i\ddot{p}_i)\cdot\ddot{p}_i\cdot s_i \tag{9}$$

The problem becomes:

$$\underset{\ddot{p}_1,\ddot{p}_2}{Max}\{\gamma(\ddot{p}_1,\ddot{p}_2)\} = \sum_{i=1,2} \ddot{p}_i \cdot (1-s_i)\cdot(a_i - b_i\ddot{p}_i) \tag{10}$$

Let $p'_i = \ddot{p}_i \cdot (1-s_i)$. Then equation (10) becomes:

$$\underset{\ddot{p}_1,\ddot{p}_2}{Max}\{\gamma(\ddot{p}_1,\ddot{p}_2)\} = \underset{p'_1,p'_2}{Max}\{\gamma(p'_1,p'_2)\} = \sum_{i=1,2} p'_i \cdot \left[a_i - b_i \cdot \left(\frac{p'_i}{1-s_i}\right)\right] \tag{9'}$$

From the first order condition on (10') we have:

$$a_i - 2b_i \cdot \left(\frac{p'^*_i}{1-s_i}\right) = 0 \rightarrow p'^* = \frac{a_i \cdot (1-s_i)}{2b_i} \rightarrow \ddot{p}^* = \frac{a_i}{2b_i} = p^*_i \tag{8'}$$

Thus, a firm whose only significant costs are sales commissions can be indifferent between optimizing for revenue versus for profit. This could explain a lack of broad attention to the subject, historically, since pricing decisions often reside in a revenue-driven business function such as sales or marketing. This insight also provides a rationale for compensating sales people, as is often done, based on gross revenue rather than net sales revenue, especially when low unit variable costs are involved.

4.5.3.2 Contribution Maximization (CM)

For net revenue maximization, the cost function includes only m_i multiplied by decision quantities.

$$\hat{C}(\hat{p}_1,\hat{p}_2) = \sum_{i=1,2} \hat{q}_i \cdot m_i = \sum_{i=1,2} (a_i - b_i\hat{p}_i)\cdot m_i \tag{11}$$

This results in a simple adjustment of both prices and quantities from RM to CM, increasing optimal prices by ½ of the now considered costs m_i. From (6) with $t_i = s_i = 0$ we obtain:

$$\hat{p}_i{}^* = \frac{a_i}{2b_i} + \frac{m_i}{2}, \quad \hat{q}_i{}^* = \frac{a_i}{2} - \frac{b_i \cdot m_i}{2} \tag{12}$$

For the case of $m_1 = m_2$ representing a common unit product cost, the CM upward price adjustments relative to RM prices are the same in both markets and hence a larger

percentage increase is applied to the lower-priced market (having the smaller $\frac{a_i}{b_i}$.) If the per unit cost of production m_i is a large fraction of the RM price $p_i^* = \frac{a_i}{2b_i}$, the choice of decision protocol adopted by management i.e. RM, CM or PM, will be most impactful.

Profit difference between PM and CM can be obtained from the last two rows, final column of Table 1.

4.5.4 Summary of Deterministic, Unconstrained Results

Table 2 provides optimal decisions for different objectives discussed above for unconstrained cases assuming deterministic demand. Subscripts i are not shown in the table, as results apply to each market. Aggregate revenues and profits are the sums of the corresponding expressions for both markets.

Objective	***Price***	***Demand***	***Revenue***	***Profit***
Revenue (RM)	$p^* = \frac{a}{2b}$	$q^* = \frac{a}{2}$	$\frac{a^2}{4b}$	$\frac{a^2(1-s)}{4b} - \frac{a(m+t)}{2}$
Contribution (CM)	*Same as RM*	*Same as RM*	*Same as RM*	Same as RM
Net Sales (SM)	$\hat{p}^* = \frac{a}{2b} + \frac{m}{2}$	$\hat{q}^* = \frac{a}{2} - \frac{bm}{2}$	$\frac{a^2}{4b} - \frac{bm^2}{4}$	$\frac{a^2(1-s)}{4b} - \frac{a(m+t)}{2} + \frac{bm(m(1-s)+2t)}{4}$
Profit (PM)	$\overline{p}^* = \frac{a}{2b} + \frac{m+t}{2(1-s)}$	$\overline{q}^* = \frac{a}{2} - \frac{b(m+t)}{2(1-s)}$	$\frac{a^2}{4b} - \frac{b(m+t)^2}{4(1-s)^2}$	$\frac{a^2(1-s)}{4b} - \frac{a(m+t)}{2} + \frac{b(m+t)^2}{4(1-s)}$

Table 4-2 Optimal Decisions and Outcomes: Unconstrained Deterministic Demand

We note that all solutions to non-RM objectives conform with remark 1. That is, optimal decisions are modifications of equation (3) with increasing prices $p_i^*(+)$ and decreasing

quantities $q_i^*(-)$, where the adjustments (+/-) are simple expressions in the parameters a_i, b_i, m_i, t_i, and s_i.

We introduced interim objectives CM and SM in the discussion for two reasons. Firstly, the SM is notable by its identical results to those of RM. Secondly, the CM is a necessary comparative to PM in the stochastic case discussion to follow, because unless some cost (typically m_i) is considered the stochastic RM problem is unbounded, and CM is a standard proxy for RM in stochastic pricing.

4.6 Capacity Effects

4.6.1 Critical Capacity Levels

Solutions described above assume that there is no limit on supply and therefore independent decisions can be made for each market. We now introduce a capacity constraint where the total quantity of product available for sale in both markets is limited (as would generally be the case for a hotel or an airline flight, and in industries where capacity must be added only in large increments.) We explore the impact of capacity on optimal prices and quantities under RM versus PM.

When facing a capacity constraint, managers should consider by what proportions demand should be sacrificed by a price increase or preserved by less of a price increase, or perhaps none in the more profitable market. Preservation of demand in the less price sensitive market (higher-priced, other parameters being equal) might be undermined by greater fulfillment costs in that market, as was discussed regarding profit maximization solution equations (6). A naïve but seemingly reasonable approach could be to identify an equal percentage price increase in both markets at which aggregate demand could be satisfied. We will show instead the equality that should be preserved is the difference between optimal market prices, more accurately net of selling percentage costs.

We derive capacity-constrained optimal prices presented in this section by methods detailed in Appendix A. Our focus here is on the conditions where such a constraint is binding. If $k \geq \sum_i q_i^*$ neither problem is capacity constrained. Decreasing k will first

constrain the RM problem because it requires higher quantities (induced by lower prices) in its unconstrained solution.

Let k be the maximum available capacity (i.e. total supply constraint.) As per remark 1, for the unconstrained case PM quantities are less than RM quantities, $\sum_{i=1}^{2} q_i(\overline{p}_i^*) \le \sum_{i=1}^{2} q_i(p_i^*)$.

We use k_R to denote the level of capacity at which supply becomes a binding constraint in the RM problem.

$$k_R \equiv \frac{(a_1 + a_2)}{2} = \sum_{i=1}^{2} q_i^* \tag{12}$$

We use k_Π to denote a lower level of capacity at which supply also becomes a binding constraint in the PM problem.

$$k_\Pi \equiv \frac{(a_1 + a_2)}{2} - \frac{b_1(m_1 + t_1)}{2(1 - s_1)} - \frac{b_2(m_2 + t_2)}{2(1 - s_2)} = \sum_{i=1}^{2} \overline{q}_i^* \tag{13}$$

4.6.2 Capacity-Constrained RM Prices

For intermediate values of k such that $k_\pi < k < k_R$, from equation (2) with the added constraint our revenue maximizing problem becomes:

$$\begin{aligned} \underset{p_1, p_2}{Max}\{R(p_1, p_2)\} \quad &= \sum_{i=1,2} p_i \cdot (a_i - b_i p_i) \\ s.t. \ &\sum_{i=1,2} a_i - b_i p_i \le k \end{aligned} \tag{14}$$

As described in Appendix A, with the first order condition $q_2 = k - q_1$, we obtain the following optimal decisions for capacity-constrained RM:

$$p_i^{k*} = \frac{a_i}{2b_i} + \frac{\left(\frac{a_1+a_2}{2} - k\right)}{(b_1+b_2)} = p_i^* + \frac{\left(\frac{a_1+a_2}{2} - k\right)}{(b_1+b_2)}$$

$$q_i^{k*} = \frac{a_i}{2} - b_i \frac{\left(\frac{a_1+a_2}{2} - k\right)}{(b_1+b_2)} = q_i^* - b_i \frac{\left(\frac{a_1+a_2}{2} - k\right)}{(b_1+b_2)} \quad (15)$$

When $k_\Pi < k < k_R = \frac{a_1+a_2}{2}$ the supply restriction increases revenue maximizing prices and decreases quantities, bringing them closer in-line with profit maximizing decisions. The aggregate quantity reduction is absorbed proportionally in each market according its price sensitivity parameter $\frac{b_i}{(b_1+b_2)}$.

Even though the explicit objective remains to maximize revenue, a binding capacity constraint has the effect of increasing profit. As k decreases further, the profit gap between revenue and profit maximization becomes progressively smaller, since RM decisions fall closer in line with PM decisions. Interestingly, the price adjustments in (15) are identical in absolute terms for both markets, so they will be disproportionate within the two markets except for the unusual case that $\frac{a_1}{b_1} = \frac{a_2}{b_2}$. Meanwhile, PM prices remain unchanged within this intermediate capacity range $k_\Pi < k < k_R$ since the supply constraint is not binding in the PM problem.

4.6.3 Capacity-Constrained PM Prices

At any $k < k_\Pi$, from Equation (5) with the added supply constraint $\sum_{i=1,2} a_i - b_i p_i \le k$, and using the method described in Appendix A, we obtain constrained profit maximizing prices:

$$\begin{aligned}
\bar{p}_i^{*k} &= \frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1-s_i)} + \frac{\frac{a_1+a_2}{2} - \frac{b_1(m_1+t_1)}{2(1-s_1)} - \frac{b_2(m_2+t_2)}{2(1-s_2)} - k}{(1-s_i)\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)} \\
&= \bar{p}_i^{\,*} + \frac{\frac{a_1+a_2}{2} - \frac{b_1(m_1+t_1)}{2(1-s_1)} - \frac{b_2(m_2+t_2)}{2(1-s_2)} - k}{(1-s_i)\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)} \\
&= \bar{p}_i^{\,*} + \frac{k_\Pi - k}{(1-s_i)\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)} \\
&= \bar{p}_i^{\,*} + \frac{\left(\sum_i \bar{q}_i^*\right) - k}{(1-s_i)\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)}
\end{aligned} \tag{16}$$

These prices translate into the following deterministic[3] demand quantities:

$$\begin{aligned}
\bar{q}_i^{*k} &= \bar{q}_i^{\,*} - b_i \frac{k_\Pi - k}{(1-s_i)\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)} \\
&= \bar{q}_i^{\,*} - b_i \frac{\left(\sum_i \bar{q}_i^*\right) - k}{(1-s_i)\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)}
\end{aligned} \tag{17}$$

The initial equation in (16) is difficult to derive algebraically and unwieldy in appearance, but can be reduced to simple expressions in the unconstrained optimal decisions $\bar{p}_i^*$ or $\bar{q}_i^*$, the capacity constraint k, the market price sensitivity parameters b_i and any sales percentage costs s_i.

The final term in equation (16) is positive since $k < k_\pi$ for capacity to be constrained in the PM problem. Capacity reduction naturally warrants higher market clearing prices. This is another instance in support of remark 1, since $\bar{p}_i^{k*} \geq \bar{p}_i^*$ and $\bar{p}_i^* \geq p_i^*$ implies $\bar{p}_i^{k*} \geq p_i^*$.

[3] In later stochastic problem discussion, we refer to deterministic demand as riskless demand

We see that, in absolute terms, the individual market price adjustments induced by the total supply constraint differ in each market only by their denominator $(1 - s_i)$ and are equal in both markets when $s_1 = s_2$. Relative to unconstrained profit maximizing solutions in equation (6), prices should be increased most in the market or channel with highest sales commission, while the difference between market prices net of sales commission across channels should remain constant

Remark 4: When unconstrained PM quantities cannot be met, $k < \sum_{i=1,2} \bar{q}_i^*$, prices $\bar{p}_i^{k*}$ should be set to stimulate demand $\sum_{i=1,2} d_i(\bar{p}_i^{k*}) = k$ while maintaining the following price relationship:

$$\bar{p}_1^{k*} \cdot (1-s_1) - \bar{p}_2^{k*} \cdot (1-s_2) = \bar{p}_1^{*} \cdot (1-s_1) - \bar{p}_2^{*} \cdot (1-s_2) \tag{18}$$

To arrive at (18), a simple approach is to let $p_i'' = \bar{p}_i^{k*}(1-s_i)$ representing constrained PM prices net of sales commissions, and also let $p_i' = \bar{p}_i^{*} \cdot (1-s_i)$ representing the unconstrained PM equivalents. Then equation (16) becomes:

$$p_i^{''*} = p_i^{'*} + \frac{\frac{a_1+a_2}{2} - \frac{b_1(m_1+t_1)}{2(1-s_1)} - \frac{b_2(m_2+t_2)}{2(1-s_2)} - k}{\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)} \tag{19}$$

Using the equivalents of (19) for the two left-hand-side terms in equation (18), and after substituting for $\bar{p}_i^{*} \cdot (1-s_i)$, from equation (6) we obtain the desired relationship between PM prices, which is the difference between unconstrained RM prices p_i^* net of sales commissions, plus a constant: $\frac{1}{2}\left(b_1(m_1+t_1) - b_2(m_2+t_2)\right)$

$$
\begin{aligned}
p_1''^{*} - p_2''^{*} &= p_1'^{*} - p_2'^{*} \\
&= \bar{p}_1^{*} \cdot (1 - s_1) - \bar{p}_2^{*} \cdot (1 - s_2) \\
&= (1 - s_1)\left(\frac{a_1}{2b_1} + \frac{m_1 + t_1}{2(1 - s_1)}\right) - (1 - s_2)\left(\frac{a_2}{2b_2} + \frac{m_2 + t_2}{2(1 - s_2)}\right) \\
&= (1 - s_1)p_1^{*} - (1 - s_2)p_2^{*} + \tfrac{1}{2}\big((m_1 + t_1) - (m_2 + t_2)\big)
\end{aligned}
\tag{20}
$$

Remark 4 suggests an alternative method for determining optimal capacity-constrained PM prices beginning with a calculation of their unconstrained versions $\bar{p}_i^{*}$ and associated demand quantities, then increasing prices in both markets while maintaining the relationship in (20) until the associated total demand quantity matches the overall capacity limit.

4.6.4 Generalized Deterministic Optimal RM and PM Prices

Summarizing the above, we find that optimal prices are defined by the following piecewise functions:

$$
p_i^{*} = \begin{cases} \frac{a_i}{2bi} & , \text{ if } k \geq \frac{a_1 + a_2}{2} \\ \frac{a_i}{2b_i} + \frac{\frac{a_1 + a_2}{2} - k}{b_1 + b_2} & , \text{ otherwise} \end{cases}
\tag{21}
$$

$$
\bar{p}_i^{*} = \begin{cases} \frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1 - s_i)} & , \text{ if } k \geq \frac{a_1 + a_2}{2} - \frac{b_1(m + t_1)}{2(1 - s_1)} - \frac{b_2(m + t_2)}{2(1 - s_2)} \\ \frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1 - s_i)} + \frac{\frac{a_1 + a_2}{2} - k - \frac{b_1(m + t_1)}{2(1 - s_1)} - \frac{b_2(m + t_2)}{2(1 - s_2)}}{2(1 - s_i)(\frac{b_1}{2(1 - s_1)} + \frac{b_2}{2(1 - s_2)})} & , \text{ otherwise} \end{cases}
\tag{22}
$$

4.6.5 Profit Loss from Constrained RM vs. PM

Remark 5: The difference in profits between PM and RM when a total capacity constraint is binding on both problems is given by:

$$\Delta\Pi^{kk} = \Pi_{\Pi}^{k} - \Pi_{R}^{k} = \frac{b_1 b_2 \left((a_1 + a_2 - 2k)(s_1 - s_2) + (b_1 + b_2)(m_1 + t_1 - m_2 - t_2)\right)^2}{4(b_1 + b_2)^2 \left(b_2(1 - s_1) + b_1(1 - s_2)\right)} \quad , \forall k \le k_{\Pi} < k_R \tag{23}$$

Where k is the capacity constraint and k_R, k_{Π} are defined in equations (12) and (13).

Substituting optimum prices (21) and (22) into the revenue and profit functions (2) and (5) leads to the result in remark 5.

When only the revenue maximizing problem is constrained, that is when $k_{\Pi} < k < k_R$, the expression for difference in profit is given by:

$$\Delta\Pi^{uk} = \Pi_{\Pi}^{u} - \Pi_{R}^{k} =$$
$$\left(\frac{a_1}{2} - \frac{b_1\left(\frac{a_1+a_2}{2} - k\right)}{b_1+b_2}\right)\left(m_i + t_1 - (1-s_1)\left(\frac{a_1}{2b_1} + \frac{\frac{a_1+a_2}{2} - k}{b_1+b_2}\right)\right) - \left(\frac{a_1}{2} - \frac{b_1(m+t_1)}{2(1-s_1)}\right)\left(m_i + t_1 - (1-s_1)\left(\frac{a_1}{2b_1} + \frac{m+t_1}{2(1-s_1)}\right)\right)$$
$$+$$
$$\left(\frac{a_2}{2} - \frac{b_2\left(\frac{a_1+a_2}{2} - k\right)}{b_1+b_2}\right)\left(m + t_2 - (1-s_2)\left(\frac{a_2}{2b_2} + \frac{\frac{a_1+a_2}{2} - k}{b_1+b_2}\right)\right) - \left(\frac{a_2}{2} - \frac{b_2(m+t_2)}{2(1-s_2)}\right)\left(m + t_2 - (1-s_2)\left(\frac{a_2}{2b_2} + \frac{m+t_2}{2(1-s_2)}\right)\right) \tag{24}$$

When neither problem (revenue nor profit) is constrained at k_R and above, equation (24) reduces to the sum of independent market solutions in equation (6), where the total profit difference is the sum of equation (7) for the two markets:

$$k \ge \frac{a_1+a_2}{2} \Rightarrow \Delta\Pi^{uu} = \frac{b_1}{4}\frac{(m_i+t_1)^2}{(1-s_1)} + \frac{b_2}{4}\frac{(m_i+t_2)^2}{(1-s_2)} \tag{25}$$

In the direction of decreasing capacity, at k_{Π}, as the PM problem becomes constrained in addition to the RM problem being constrained equation (24) reduces to equation (23).

4.6.6 Graphical Summary

Figure 1 provides a graphical depiction of the results above under these parameters:

$a_1 = 450,\ b_1 = 0.5,\ m_1 = 50,\ t_1 = 5,\ s_1 = 0, a_2 = 2400,\ b_1 = 7.5,\ m_2 = 50,\ t_2 = 15,\ s_1 = 0.10$

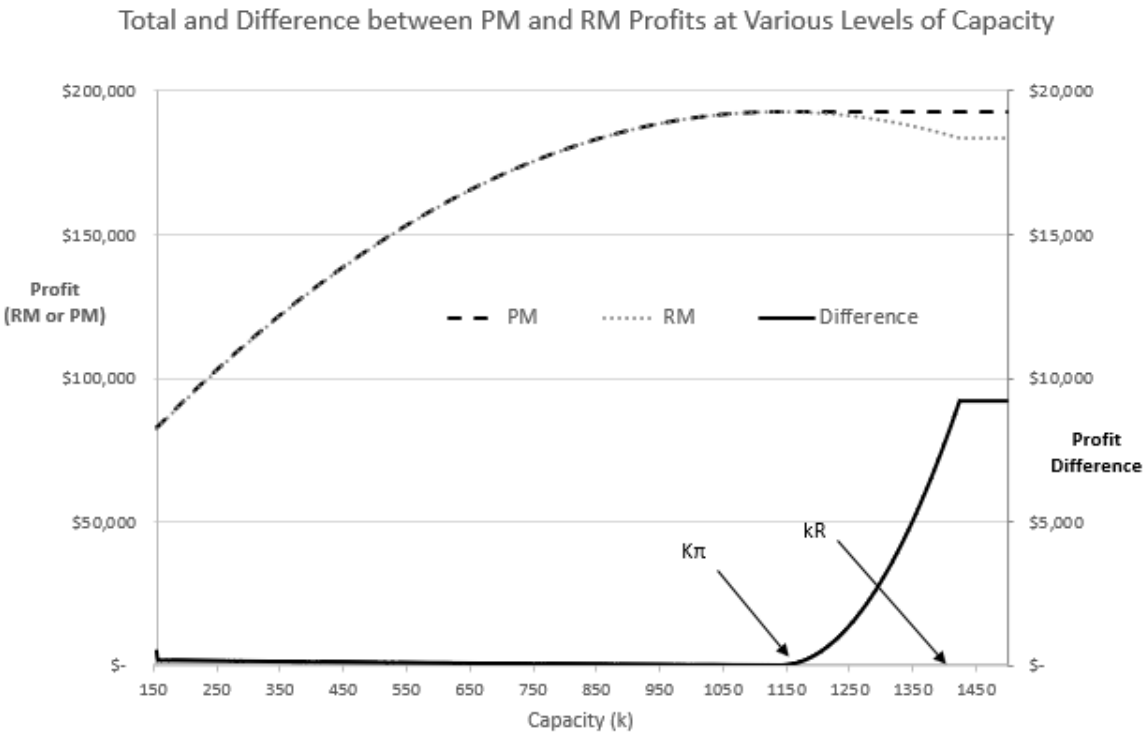


Table 4-3 PM vs. RM - Influence of Capacity on Profit Difference

The horizontal axis in Figure 1 represents production capacity in units. The left vertical axis represents profit either from PM or RM. The right vertical axis represents the profit difference $\Pi^{\Pi} - \Pi^{R}$. We can see that in the region of capacity equal to and greater than k_R, where neither maximization problem is constrained, the difference in profit is constant because additional capacity is not utilized i.e. optimal decisions remain unchanged. The horizontal position of k_R aligns with the peak of an RM revenue curve (not shown due to its higher order-of-magnitude. We show instead the RM profit curve) which is shaped like the PM profit curve in Figure 1 but not peaking at k_{Π} (1140 units in this example), rather at k_R (1425 units in this example).

We do not consider capacity levels near and below 150 units in this example, because prices from equation (16) will in that region be such that there is no demand from market 2 (demand will be consumed by market 1 at prices that are equally or more profit generating.) Furthermore, such capacity is well-outside the range for which the demand curve approximation can be considered reasonable.

The range of greatest interest is between 1140 and 1425 units in our example, that is, between k_{Π} and k_R. This is where the firm is most vulnerable to profit loss from using RM rather than PM. Below k_{Π} we find the very interesting result that profit from RM and PM are nearly equal, although achieved with different prices and quantities, as we see later in the Illustrative Example section.

4.7 The Stochastic Model

We extend our comparison of revenue versus profit maximization to scenarios where demand is uncertain.

Like (Petruzzi and Dada 1999) we model demand in each channel as a linear function, now with an additive error term.

$$\tilde{d}_i(p_i,\varepsilon_i) = d_i(p_i) + \varepsilon_i = a_i - b_i p_i + \varepsilon_i, \text{ for } i = 1,2 \quad \textbf{(27)}$$

where the random elements ε_i are distributed independently according to probability density functions $f_i(\cdot)$ with mean $E[\varepsilon_i] = 0$ and standard deviations σ_i. We assume that the parameters of the demand distributions and the range of prices we will consider are such that the probability of negative demand is essentially zero.

Using method developed in other seminal articles on stochastic pricing and inventory control (Whitin 1955),(Mills 1959), we define new implicit decision variables $z_i = q_i - d_i(p_i)$ to represent the difference between planned inventory quantities q_i and demand quantities predicted by the linear functions $d_i(p) = a_i - b_i p_i$. In effect, z_i is the firm's decision by what amount to either increase or decrease quantities relative to expected demand at a chosen price. This stock adjustment decision hedges in favor (or

against) a surplus (or shortage) in realized demand $\tilde{d}_i(p_i, \varepsilon_i)$, relative to the what we call the 'riskless' demand component $d_i(p_i)$ of equation (27), as originally defined in equation (1).

Again, following the example of (Petruzzi and Dada 1999), we consider $p_i(z_i)$ as a function of the stocking decision. This allows us to reduce the two-variable problem (p_i, z_i) in each channel to a problem in the variables p_i only. We then disaggregate the stochastic profit function into a riskless (deterministic) component and a stochastic loss function component. The latter loss function is comprised in turn of two stochastic functions representing losses from possible inventory surpluses or shortages, respectively.

4.7.1 Independent Markets, Unconstrained

Expected profit in each market $E[\tilde{\Pi}_i(z_i, p_i)]$ is written as a combination of the riskless profit function $\Psi_i(p_i)$ and a loss function $L_i(z_i, p_i)$ which accounts for the probability weighted costs of all possible surpluses $\tilde{d}_i(p_i) < q_i$ and shortages $\tilde{d}_i(p_i) > q_i$.

$$E[\tilde{\Pi}_i(z_i, p_i)] = \Psi_i(p_i) - L_i(z_i, p_i) \tag{28}$$

The riskless profit function is further defined as:

$$\Psi_i(p_i) = \left[p_i(1-s_i) - (m+t_i)\right] \cdot d_i(p_i) \tag{29}$$

This is equivalent to the deterministic single channel profit function in equation (5).

The loss function is further defined as:

$$L_i(z_i, p_i) = (m_i + t_i)\Lambda_i(z_i) + (p_i(1-s_i) - (m_i + t_i))\Theta_i(z_i) \tag{30}$$

It incorporates two random functions representing potential inventory surplus $\Lambda_i(z_i)$ and shortage $\Theta_i(z_i)$, multiplied by their respective costs, and in the latter case based on price.

Given a chosen price $\tilde{\bar{p}}_i^*$ and inventory level $d_i(\tilde{\bar{p}}_i^*)+z_i^*$, the surplus function equation (31) is the expected inventory surplus from the choice of z_i^* i.e. weighted sum of surplus probabilities time amount by which demand will be exceeded,
$(z_i-\varepsilon_i)\cdot \Pr\left\{D_i = d_i(\tilde{\bar{p}}_i^*)+\varepsilon_i \le \tilde{\bar{q}}_i^* = d_i(\tilde{\bar{p}}_i^*)+z_i^*\right\}$

$$\Lambda_i(z_i)=\int_{-\infty}^{z_i}(z_i-u)f_i(u)du \tag{31}$$

The shortage function equation (30) provides the analogous expectation of inventory deficiency.

$$\Theta_i(z_i)=\int_{z_i}^{\infty}(u-z_i)f_i(u)du \tag{32}$$

Our method for determining stochastic unconstrained decisions $\tilde{\bar{p}}_i^*$ and $\bar{z}_i^*$ is dependent on first establishing prices uniquely as functions of z_i as described in Appendix B

$$\tilde{\bar{p}}_l^*(z_i)=\bar{p}_i^*-\frac{\Theta_i(z_i)}{2b_i} \tag{33) / (B7}$$

The solution for z_i^* is an expression like the familiar newsvendor-problem *critical fractile*. However, the calculation is different in our case as it is based on the ratio of *unit profit* to *net price* after commissions, whereas the classic solution uses a ratio of unit *contribution* to *gross price*.

$$z_i^* = F_i^{-1}\left(\frac{\tilde{\bar{p}}_i^*(1-s_i)-(m_i+t_i)}{\tilde{\bar{p}}_i^*(1-s_i)}\right) = F_i^{-1}\left(1-\frac{(m_i+t_i)}{\tilde{\bar{p}}_i^*(1-s_i)}\right) \qquad (34) / (B8)$$

From the above we conclude the following:

Remark 6: For unconstrained stochastic PM price and quantity decisions:

1. prices $\tilde{\bar{p}}_i^*$ should differ from deterministic counterparts $\bar{p}_i^*$ by amounts $-\frac{\Theta(z_i^*)}{2}$, where $\Theta(\cdot)$ is a shortage function as defined in equation (32).

2. stock adjustments $\bar{z}_i^*$ should be chosen such that $F_i(z_i^*) = 1-\frac{(m_i+t_i)}{\tilde{\bar{p}}_i^*(1-s_i)}$ is the probability of actual demand being met by the inventory quantity $\Pr\left\{D_i \le d_i(\tilde{\bar{p}}_i^*) + z_i^*\right\}$.

Some important managerial implications can be drawn from the operand in equation (34). The optimal inventory for any price level p_i is a decreasing function of costs, where the effect of s_i is to multiply the effects of m_i and t_i. High costs clearly warrant holding less 'safety stock'. On the other hand, when manufacturing and/or delivery costs are not included there is motivation to hold more safety stock, regardless of s_i, to take advantage of any possibility for additional revenue by reducing the probability of a shortage to zero. This is also evident from equation (30), in which if no costs are applied to the surplus function there is per unit incremental benefit $p_i(1-s_i)$ from choosing the maximum stock $z_i = \infty$.

Equation (33) is a similar result to that obtained by (Petruzzi and Dada 1999) , however, the first term $\bar{p}_i^*$ on the right-hand-side is not the riskless *contribution* maximizing price of that paper but rather our riskless *profit* maximizing price from equation (6) which incorporates market specific costs m_i, t_i and s_i. Like the result of those authors, the last term on the right-hand-side of equation (33) adjusts prices for uncertainty in a manner that does not depend on any costs. However, the expected profit maximizing stocking factor $\bar{z}_i^*$ is calculated somewhat differently in our PM case, as follows:

$$\bar{z}_i^* = \Theta_i^{-1}\left(2b_i \cdot \left(\frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1-s_i)_i} - \tilde{\bar{p}}_i^*\right)\right) \tag{35}$$

To continue our development, we restrict attention to the case of $f_i(\cdot)$ uniformly distributed on $[-\sigma_i, \sigma_i]$ as it provides for simple exposition, although any non-decreasing hazard rate distribution can be used Combining equations (34) and (35) together with the uniformly distributed shortage function, we obtain the following as in Appendix B:

$$\frac{1}{2\sigma_i}\left(\frac{\sigma_i^2}{2} - \bar{z}_i^*(\sigma_i - \bar{z}_i^*) - \frac{\bar{z}_i^{*2}}{2}\right) = 2b_i \cdot \left(\frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1-s_i)_i} - \tilde{\bar{p}}_i^*\right) \tag{36) / (B10}$$

Excel Model for Stochastic PM in Independent Channels

Although we are unable to derive a closed-form solution for $\bar{z}_i^*(\tilde{\bar{p}}_i^*)$, we can solve for it numerically, in Microsoft Excel, as follows:

1. Determine $\bar{p}_i^*$ from the parameters and equation (6) and enter this value into a cell.
2. Create a cell for $\tilde{\bar{p}}_i$ and enter a trial value (perhaps $\bar{p}_i^*$, initially.)
3. Create a cell for the inverse shortage function $\Theta_i^{-1} = 2a_i(\bar{p}_i^* - \tilde{\bar{p}}_i)$
4. Create a cell containing the critical fractile of equation (33) $\bar{CF}_i = 1 - \frac{(m+t_i)}{\tilde{\bar{p}}_i(1-s_i)}$
5. Create a cell referring to the formula on the left-hand-side of (36) / (B10).
6. Create a control cell for the difference between 3. and 5. Perform a linear numerical search for a price in cell 2 that will set this control cell to zero.

Given the resulting choice of $\bar{p}_i^*$, the riskless inventory quantity $\tilde{\bar{q}}_i^*$ can then be calculated from equation (1), and $\bar{z}_i^*$ can be determined from equation (34).

4.8 Stochastic, Capacity-Constrained Model

The previous discussion applies to each market independently whereas the problem becomes substantially more complex when there is a constraint on the total number of

units that can be produced or procured. It is a common approach to distinguish customer groups (markets in our case) is to assume that demand arrives sequentially. This practice is often employed in airline seat class bookings, and has well-accepted revenue management solutions such as Littlewood's law (Littlewood 1972), (Talluri and Van Ryzin 2004).

The challenge is that expected profit in the first market is comprised of two expectations in equation (37) below, each of which leads to two more similar expectations in the second market, equations (38). For these equations, interpret k as capacity in units, q_1 as the stocking quantity in market 1, d_i as the realized demand in market i, and Q_i as the realized sales in market i, which will be less than or equal to q_i and no greater than d_i.

$$\begin{aligned} E[\Pi_1] = {} & E[\Pi_1 \mid Q_i < q_1] \cdot \Pr\{d_1 < q_1\} \\ & + E[\Pi_1 \mid Q_i = q_1] \cdot \Pr\{d_1 \geq q_1\} \end{aligned} \tag{37}$$

$$\begin{aligned} E[\Pi_2] = {} & E[\Pi_2 \mid Q_2 < k - Q_1] \cdot \Pr\{d_2 < k - q_1\} \cdot \Pr\{d_1 < q_1 : Q_1 = d_1\} \\ & + E[\Pi_2 \mid Q_2 < k - q_1] \cdot \Pr\{Q_2 < k - q_1\} \cdot \Pr\{d_1 \geq q_1 : Q_1 = q_1\} \\ & + E[\Pi_2 \mid Q_2 < k - Q_1] \cdot \Pr\{d_2 < k - q_1\} \cdot \Pr\{d_1 < q_1 : Q_1 = d_1\} \\ & + E[\Pi_2 \mid Q_2 < k - q_1] \cdot \Pr\{Q_2 < k - q_1\} \cdot \Pr\{d_1 \geq q_1 : Q_1 = q_1\} \end{aligned} \tag{38}$$

Analytical solutions developed with this approach proved to be complex, including more than a dozen terms including several dual integrals, thus we do not present them here.

4.8.1 Marginal Analysis (Point Elasticities)

Instead, we approach the problem from a different perspective using point elasticities of demand. Point elasticity of demand is a ratio of the percentage change in demand to a percentage change in price at any given point on the demand curve (or line in our case):

$e_i^d(p_i) = \frac{dq_i}{dp_i} \cdot \frac{p_i}{q_i}$, which for our linear demand models translates to $e_i^d = -b_i\left(\frac{p_i}{a_i - b_i p_i}\right)$ (39)

By substituting deterministic revenue maximizing prices from equation (3) we obtain the result that $e_i^{d*} = 1$ for both channels. It is appropriate that the two are equal, as well as both being equal to one at optimal RM prices. It implies for both channels an equilibrium between revenue growth by higher price (lower quantity) versus by higher quantity (lower price). PM prices, being higher than unconstrained RM prices, have point price elasticities greater than one, $e_i^d(\bar{p}_i^*) > 1$.

Remark 7: Profit optimal prices generate equal marginal profits in both channels and for the capacity-constrained problem of choosing $\tilde{\tilde{p}}_i^k$ the following relationships must hold:

1. $\tilde{\tilde{p}}_1^k \cdot (1-s_1) - \tilde{\tilde{p}}_2^k \cdot (1-s_2) = \bar{p}_1^* \cdot (1-s_1) - \bar{p}_2^* \cdot (1-s_2)$
2. $\sum_{i=1,2} \left(\tilde{\tilde{q}}_i^k(\tilde{\tilde{p}}_i^k) + \bar{z}_i^k(\tilde{\tilde{p}}_i^k) \right) \le k$

Remark 7 is derived from the relationship of point price elasticities (39) to marginal profit, but we develop the concept first for marginal revenue, as it is a simpler exposition. The critical notion is to maintain the relationship between $e_1^d(p_1)$ and $e_2^d(p_2)$ such that they contribute equally to the objective function. Whereas for RM the two should be equal (and =1 if $k \ge k_R$), this is generally not true for PM, and we can determine the relationship for PM from our deterministic results.

From equation (39) we can write $$\frac{dp_i}{dq_i} = \frac{p_i}{q_i \cdot e_i^d} \tag{40}$$

We then write the following for marginal revenue (R_i') in each market:

$$R_i' = \frac{dR_i}{dq_i} = \frac{d}{dq_i}(p_i \cdot q_i) = p_i + q_i \cdot \frac{dp_i}{dq_i} = p_i + q_i \cdot \frac{p_i}{q_i \cdot e_i^d} = p_i \cdot (1 + \frac{1}{e_i^d}) \tag{41}$$

We determine that $R_1' = R_2'$ when $\frac{p_1}{p_2} = \frac{e_1^d}{e_2^d}\left(\frac{e_2^d + 1}{e_1^d + 1}\right)$.

Substituting the final term from equation (41) for each e_i^d we have:

$$1 = \frac{b_1}{b_2}\left(\frac{a_2 - 2 \cdot b_2 \cdot p_2}{a_1 - 2 \cdot b_1 \cdot p_1}\right) \Rightarrow p_1 = p_2 + \frac{a_1}{2b_1} - \frac{a_2}{2b_2} \Leftrightarrow p_1 - p_2 = p_1^* - p_2^* \tag{42}$$

For marginal revenues to be equal in two markets, the difference between the two market prices should be a constant that is equal to the difference in their optimal values.

We apply the same rule to marginal profits, with a slightly different formulation to account for costs. For marginal profit this approach leads to the same result as in our earlier remark 4 for deterministic capacity-constrained PM problems. For marginal profit in each channel ($\Pi_i^{'}$) at any given price p_i, we have a variation of equation (41)

$$\Pi_i^{'} = p_i(1-s_i)(1+\tfrac{1}{\bar{e}_i^d}) - (m_i + t_i) \tag{43}$$

Using the same technique as above, for marginal profits to be equal we require the following:

$$\bar{p}_1^*(1-s_1) - \bar{p}_2^*(1-s_2) = p_1^*(1-s_1) - p_2^*(1-s_2) + \tfrac{1}{2}\left((m_1+t_1) - (m_2+t_2)\right) \tag{44}$$

Substituting $\bar{p}_2^* = \frac{a_2}{2b_2} + \frac{(m_2+t_2)}{2(1-s_2)}$ from (6) into (44) we confirm that $\bar{p}_1^* = \frac{a_1}{2b_1} + \frac{(m_1+t_1)}{2(1-s_1)}$.

Substituting the same into (43), including within the term $\bar{e}_i^d = -b_i\left(\frac{\bar{p}_i^*}{a_i - b_i\bar{p}_i^*}\right)$ we find that marginal profits $\Pi_i^{'} = 0$when $p_i = \bar{p}_i^*$, just as marginal revenues $R_i^{'} = 0$ when $p_i = p_i^*$.

Finally, we note the following result:

Remark 8: Point price elasticities at unconstrained profit maximizing prices are the negative reciprocals of the critical fractiles given for z_i^* in equation (34), that is

$$\bar{e}_i^{d*} = -\left(\frac{\bar{p}_i^*(1-s_i)}{\bar{p}_i^*(1-s_i) - (m_i + t_i)}\right) \tag{45}$$

By setting equation (43) equal to zero we have

$\frac{\bar{e}_i^{d*} + 1}{\bar{e}_i^{d*}} = \frac{(m_i + t_i)}{\bar{p}_i^*(1-s_i)} \Rightarrow \bar{e}_i^{d*}\left(\frac{(m_i + t_i)}{\bar{p}_i^*(1-s_i)} - 1\right) = 1$, from which equation (45) follows directly.

Remark 8 is quite intuitive, as the operand represents the reciprocal of profit gained from an additional unit sold at the specified price, that is $\frac{dq_i(p_i)}{d\Pi_i(p_i, q_i(p_i))}$. From (39) and (45) we obtain $\frac{dq_i}{dp_i}\frac{p_i}{q_i} = \frac{dq_i}{d\Pi_i} \Rightarrow \frac{dq_i}{dp_i}\left(\frac{d\Pi_i}{dq_i}\right)\left(\frac{dp_i}{d\Pi_i}\right) = \frac{dq_i}{d\Pi_{i\,i}} \Rightarrow \frac{d\Pi_i}{dq_i} = 1$ at unconstrained profit maximizing $\overline{p}_i^*$. We should expect this, as otherwise if $\frac{d\Pi_i}{dq_i} < 1$ we could increase profits by increasing quantity (lowering price), or conversely if $\frac{d\Pi_i}{dq_i} > 1$ we could increase profits by decreasing quantity (increasing price.)

4.8.2 Algorithm for PM Under a Capacity Constraint

We propose the following algorithm to identify stochastic profit maximizing prices at capacity k that is below $k_\pi + \sum_{i=1,2} \overline{z}_i^*(\overline{p}_i^*)$:

1. Determine unconstrained *revenue* maximizing prices p_i^*.
2. Determine the constant amount $\Delta ps = p_2^* - p_1^* \frac{(1-s_1)}{(1-s_2)} + \frac{(m_2+t_2)-(m_1+t_1)}{2(1-s_2)}$
3. Starting at *profit* maximizing prices with $\tilde{\overline{p}}_1 = \overline{p}_1^*$ and $\tilde{\overline{p}}_2 = \overline{p}_2^*$, Incrementally increase $\tilde{\overline{p}}_1$ while maintaining the relationship $\tilde{\overline{p}}_2 = \tilde{\overline{p}}_1 \frac{(1-s_1)}{(1-s_2)} + \Delta ps$.
4. At each stage, calculate $d_i(\tilde{\overline{p}}_i^k)$ from (1) and $\overline{z}_i^k(\tilde{\overline{p}}_i^k)$ from (33).
5. Increase prices in this manner until the total quantity calculated in 4. is equal to k.

We used this method to develop solutions to the stochastic, capacity-constrained scenario in the following section.

4.9 Illustrative Example

We consider the same scenario as we did previously with the following parameters, although we note that many other examples are possible and may lead to some characteristically different results, according to the specific market parameters and costs involved:

a_1	b_1	a_2	b_2	m	t_1	t_2	s_1	s_2	σ_1	σ_2
450	0.5	2400	7.5	100	5	15	0	0.10	25	75

Market 1 is a smaller market, but we calculate that 67% of those customers will buy the product at \$300 and roughly 55% will buy at a \$400 price. Market 2 is twice the size, but only 50% of those customers will buy at \$160, and none of these customers will buy at \$320 or above.

In Market 1, the deterministic revenue maximizing price $p_1^* = 450$ with an expected demand of 225 units at that price. In Market 2, the simple revenue maximizing price p_2^* is considerably lower, at \$160, at which expected demand is 1200 units.

For either market, the common unit cost m is \$100, with additional unit delivery costs of \$5 in Market 1 and \$15 in Market 2. There is no sales commission in Market 1 and a 10% sales commission in Market 2. The linear demand forecasts given by $d_i(p_i) = a_i - b_i p_i$ are believed to be accurate within 25 units in Market 1 and 75 units in Market 2.

Tables 3 through 6 provide summaries of deterministic and stochastic, unconstrained and constrained results for our illustrative example. Although our primary interest from the start has been to compare RM and PM, we include CM in the analysis as it is a common proxy for RM in stochastic optimization (without including some cost, typically m_i, stochastic revenue maximization would seek inventory to satisfy the maximum possible demand rather than some reasonably expected amount.)

4.9.1 Deterministic – Unconstrained

Objective	*Price*	*Demand*	*Revenue*	*Profit*
Revenue (RM)	$p_1^* = \$450.00$ $p_2^* = \$160.00$	$q_1^* = 225$ $q_2^* = 1200$	$\$293,250$	$\$183,675$
Contribution (CM)	$\hat{p}_1^* = \$475.00$ $\hat{p}_2^* = \$185.00$	$\hat{q}_1^* = 213$ $\hat{q}_2^* = 1013$	$\$288,250$	$\$192,019$
Profit (PM)	$\overline{p}_1^* = \$477.50$ $\overline{p}_2^* = \$196.11$	$q_1^* = 211$ $q_2^* = 929$	$\$283,092$	$\$192,855$

Table 4-4 Deterministic, Unconstrained Decisions and Outcomes (RM, CM and PM)

Proposition 1 is supported by Table 3, as PM prices exceed RM prices in both markets. We see that PM profit exceeds RM profit, whereas the opposite is true for revenue outcomes. The difference between RM and PM prices is $27.50 in market 1 and a much greater $36.11 in the more costly-to-serve market 2, as discussed following equation (6). The PM price in market 2 is more than 20% greater that the RM price, but with only a 6% higher PM price in market 1. Furthermore, the PM price increase as a share of RM unit profit (price minus all costs) is more than 45% in market 2 and only 6% in market 1. This highlights the concern expressed following (6), that RM relative to PM may promote greater price reduction (demand generation) in the least profitable market.

The PM-RM profit difference in the final column of Table 1 matches proposition 2, equation (7). The differences between prices and quantities in each channel among the three decision protocols are easily confirmed as being the simple adjustment terms to (3) that we showed in (6) and (12). For example, $+\frac{b_2 m_2}{2} = +\frac{7.5 \times 50}{2} \cong +188$ is the difference between RM versus PM quantities in market 2.

Table 3 readily identifies the critical capacity levels $k_R = 1425$ and $k_\pi = 1140$ units, the same as they can be calculated with the parameters by equations (12) and (13). We also see a similar capacity level result which we might call $k_C = 1225$ (rounded) at which the CM problem becomes constrained, noting that $k_C = k_R - \frac{1}{2}\sum_i b_i m_i = 1425 - 200$ in our example.

We calculate that marginal revenues in both channels equal 0 at (p_1^*, p_2^*), and that their point elasticities of demand both equal 1 at the same point. We also calculate that marginal profits equal 0 in both channels at $(\overline{p}_1^*, \overline{p}_2^*)$ and that elasticities of demand at this point are as given in equation (47).

4.9.2 Deterministic - Capacity Constrained

For our capacity-constrained scenarios, we chose a constraint $k = 750$, below all critical levels.

Objective	*Price*	*Demand*	*Revenue*	*Profit*
Revenue (RM)	$p_1^{k*} = \$534.38$ $p_2^{k*} = \$244.38$	$q_1^* = 183$ $q_2^* = 567$	$\$236,297$	$\$175,514$
Contribution (CM)	$\hat{p}_1^{k*} = \$534.38$ $\hat{p}_2^{k*} = \$244.38$	$\hat{q}_1^* = 183$ $\hat{q}_2^* = 567$	$\$236,297$	$\$175,514$
Profit (PM)	$\overline{p}_i^{k*} = \$521.70$ $\overline{p}_i^{k*} = \$245.22$	$\tilde{\overline{q}}_1^* = 189$ $\tilde{\overline{q}}_2^* = 561$	$\$236,211$	$\$175,600$

Table 4-5 : Deterministic, Capacity-Constrained Decisions & Outcomes (RM, CM, PM)

In contrast to proposition 1 for unconstrained optimal prices, we find that at this constrained capacity level $p_1^{k*} > \overline{p}_1^{k*}$ although $p_2^{k*} < \overline{p}_2^{k*}$. This highlights the importance, when facing a capacity constraint, of not overpricing in the most profitable i.e. least costly to serve market 1, as would happen if approaching the problem as one of CM (being a proxy for RM in stochastic cases.)

Interestingly, the capacity-constrained decisions for CM are the same as for RM in this case. We explain this by the fact that we have $m_1 = m_2 = \$50\ (= m)$. The CM optimal price equivalent of both RM equation (15) and PM equation (16) is given by the first equality in equation (48), which if $m_1 = m_2\ (= m)$ leads to the remaining expressions in (48).

$$\hat{p}_i^{k*} = \hat{p}_i^* + \frac{(\frac{a_1+a_2-(m_1b_1+m_2b_2)}{2} - k)}{(b_1+b_2)} = p_i^* + \frac{m}{2} + \frac{\sum_{i=1,2}(\hat{q}^*) - k}{(b_1+b_2)} = p_i^* + \frac{(\frac{a_1+a_2}{2} - k)}{(b_1+b_2)} = p_i^{k*} \quad (48)$$

Using market 1 as an example, from the expression just left of the final equal sign in (48), we have $\$450 + \frac{1425-750}{0.5+7.5} = \$534.38$, while from the expression to the right of the first equal sign we obtain the identical result $\$475 + \frac{1225-750}{(0.5+7.5)} = \$534.38$.

Regarding proposition 4, that the differences *within* each channel between its constrained and unconstrained optimal price net of commissions is given by $(\overline{p}_1^{k*} - \overline{p}_1^*) \cdot (1 - s_1) = (\overline{p}_2^{k*} - \overline{p}_2^*) \cdot (1 - s_2)$, we find that indeed $(\$525.43 - \$477.50)(1) = \$47.93 = (\$249.37 - \$196.11)(0.9)$.

We also verify equation (20), that the difference in optimal prices net of commissions *between* markets, under a capacity constraint is equal to that of their unconstrained counterparts plus a constant. With our example values: $\$521.70 - \$245.22(0.9) = \$301 = \$450 - \$160(0.9) + \frac{1}{2}((50+5) - (50+15))$.

Equation (23) in proposition 5 is confirmed as the profit difference between PM versus RM at $k = 750$:

$$\frac{(0.5)(7.5)((2850 - 2(750))(0 - 0.1) + (0.5 + 7.5)(50 + 5 - 50 - 15))^2}{4(0.5 + 7.5)^2 (7.5(1) + 0.5(0.9))} = \$175,599.33 - \$175,514.36 = \$85.17$$

4.9.3 Stochastic – Unconstrained

Objective	*Price*	*RF Demand*	*Stock +/-*	*E[Net Revenue]*	*E[Profit]*
Contribution (CM)	$\tilde{\hat{p}}_1^* = \$494.66$ $\tilde{\hat{p}}_2^* = \$186.82$	$\hat{q}_1^* = 203$ $\hat{q}_2^* = 999$	$z_1^* = 20$ $z_2^* = 35$	$\$286,859$	$\$189,929$
Profit (PM)	$\tilde{\overline{p}}_1^* = \$496.62$ $\tilde{\overline{p}}_2^* = \$196.25$	$\tilde{\overline{q}}_1^* = 202$ $\tilde{\overline{q}}_2^* = 928$	$z_1^* = 19$ $z_2^* = 20$	$\$282,309$	$\$190,521$

Table 4-6 Stochastic, Unconstrained Optimal Decisions and Outcomes (CM and PM)

Table 5 provides optimal decisions and outcomes for the stochastic unconstrained scenario. Note that *expected* revenues and profits are given in this case. We evaluated expectations in terms of the associated decision perspective i.e. E[Net Revenue] is evaluated considering only costs m_i, and on that basis CM decisions are more

favourable. But we see again that PM leads to greater expected profit, while the opposite is true of expected profit from an RM approach.

We see that optimal prices for both CM and PM in market 2 have been increased substantially more than in market 1. This is to accommodate a lower riskless demand quantity offset in each market by its stock adjustment for uncertainty. However, we note that the adjustment z_2^* in market 2 is a smaller portion of the uncertainty parameter σ_2 than is the comparable portion in market 1. This is due to a combination of (a) using critical fractile in equation (33) with $t_i = s_i = 0$ for CM, versus $t_i > 0,\ s_i > 0$ in PM, and (b) generally lower prices in market 2 due to a smaller ratio $\frac{a_2}{b_2} < \frac{a_1}{b_2}$ such that the ratio $\frac{m_i}{\hat{p}_i}$ is greater, and therefore the critical fractile is also smaller.

4.9.4 Stochastic - Capacity Constrained

Using $k = 750$ once again, Table 6 provides the capacity-constrained decisions and outcomes for the stochastic case. In all cases, the total quantity $\sum_{i=1,2}(q_i^* + z_i^*) = 750$. To arrive at these decisions, we used the method described at the end of the section immediately preceding our illustrative example.

Objective	*Price*	*RF Demand*	*Stock +/-*	*E[Net Revenue]*	*E[Profit]*
Contribution (CM)	$\tilde{\hat{p}}_1^{k*} = \$541.63$ $\tilde{\hat{p}}_2^{k*} = \$251.63$	$\tilde{\hat{q}}_1^{k*} = 179$ $\tilde{\hat{q}}_2^{k*} = 513$	$\hat{z}_1^{k*} = 20$ $\hat{z}_2^{k*} = 38$	$\$286,859$	$\$169,888$
Profit (PM)	$\tilde{\bar{p}}_1^{k*} = \$525.43$ $\tilde{\bar{p}}_2^{k*} = \$249.37$	$\tilde{\bar{q}}_1^{k*} = 187$ $\tilde{\bar{q}}_2^{k*} = 530$	$\bar{z}_1^{k*} = 19$ $\bar{z}_2^{k*} = 14$	$\$282,309$	$\$172,559$

Table 4-7 Stochastic, Capacity-Constrained Decisions and Outcomes (CM and PM)

First, comparing these results to the deterministic constrained case in Table 4, we see that prices are slightly higher for both CM and PM. This is because some of the inventory quantity is required to accommodate the weighted probability of demand exceeding the riskless forecast from equation (2) at any chosen price. Unit profits at this constrained capacity level (requiring higher prices relative to p_i^*) are large for both markets, and

above 50% such that z_i^* are positive ($-\sigma_i < 0 < z_i^* < \sigma_i^*$.) This is also true for the previous, unconstrained stochastic cases in Table 5.

The change from Table 5 to Table 6 in z_2^{k*} is noteworthy as it is additive (+3) in the CM case and subtractive (-6) in the PM case. This is due to a combination of:

(a) market 2 having generally lower price ~$225 vs. ~$500, also

(b) higher t_i and s_i in market 2 , with

(c) substantial m_i relative to p_i at unconstrained optimal prices to begin with, and

(d) interaction of all factors in the critical fractile equation (33) determining $z_i^*(p_i^*)$.

Once again, PM produces higher expected profit than CM i.e. RM net of cost m_i only.

Profit difference PM-CM is less than in the equally constrained $k = 750$ deterministic case in Table 3. Comparing PM to net revenue for consistency, we found profit differences for the various scenarios as shown in Table 7.

	Deterministic		***Stochastic***	
Scenario	U	K	$\tilde{U}$	$\tilde{K}$
Profit Difference	$836	$86	$592	$3581

Table 4-8 PM vs. CM Profit Differences Summary

The deterministic results in Table 7 results reflect our earlier insight (in the graphical example Figure 1) that capacity constraints have the impact of reducing the profit difference by bringing RM prices more in line with PM prices. The stochastic cases are quite interesting. Whereas the unconstrained results show an improvement i.e. reduction in the expected PM-CM profit difference, the constrained case unlike its deterministic counterpart shows an increase in that difference.

We believe that the following is a reasonable explanation for this finding, notwithstanding that it is partially also a consequence of the parameters used in this analysis. The stochastic scenario introduces a loss function such that uncertainty has a net detrimental effect on any decision, but the degree of that effect becomes more

exaggerated at tightly-constrained capacities (750 in our example) where the marginal profits are high, versus zero or very nearly zero marginal profits in the unconstrained quantity range (1140 in our example.) We should expect losses from uncertainty to be greater under capacity constraint where both point price elasticities and marginal profits are higher and thus might be expected to have more volatile impacts on profit.

4.10 Summary

Our aim with this paper has been to examine the effects of costs, capacity and decision objectives on profits for a producer with two distinct and independent channels. We developed and presented price and quantity decisions for a variety of scenarios and objectives, most notably for revenue and profit maximization.

We first examined cases of deterministic linear demand, without and with an overall capacity constraint. We identified memorable closed-form solutions for profit maximizing decisions and other objectives, expressing them in the market and cost parameters and relative to revenue maximizing prices. We showed that PM always warrants an upward price adjustment relative to RM in the absence of a capacity constraint.

We identified two critical capacity levels in the deterministic scenario, one corresponding to aggregate demand from RM prices in both markets, and the other to aggregate demand from PM prices. We developed closed-form expressions for the difference between PM and RM decisions and outcomes above, between, and below these critical capacity levels.

For our stochastic analysis we demonstrated a parallel between methods for maximization of expected contribution (CM) in the literature and a method for stochastic PM. We provided such a method for determining optimal prices and stocking quantities in a simple MS-Excel model, both for deterministic and unconstrained stochastic versions of the problem.

Our illustrative example suggests that in a deterministic setting PM has the greatest positive impact when there is no capacity constraint. Moreover, while PM may not increase profits substantially, it does so with a lower inventory requirement, which could

release capital that a firm might deploy otherwise. Tightening capacity restraints in the deterministic case brings RM (decisions and thereby) profits closer in-line with PM results. Uncertainty prevents PM from having as significant an impact in unconstrained situations, but conversely PM can mitigate losses from uncertainty in constrained capacity situations, especially where price and quantity decisions are much higher and lower than their unconstrained counterparts.

4.11 Opportunities for Future Research

We acknowledge that the linear demand models and uniform uncertainty distributions used in this paper may be subject to criticism. We argue that these are reasonable approaches if properly confined to a realistic range of prices and quantities about which decisions need to be made, and the results provide useful lessons to a practitioner including an easily replicable framework for examining their own pricing and inventory decisions. However, the extent to which these results may change under different demand functions and uncertainty distributions warrants further exploration.

We have limited our analysis to two independent channels. Research to consider the effect of demand leakage between channels could serve as a valuable extension. Extension of the model to three or more channels would also be valuable, and we believe that our results may form the basis for optimizing channel decisions with pairwise decisions about relative prices and quantities, such as we have developed for the dual channel case.

4.12 Conclusion

A compelling case has been made for RM in recent years, and its adoption continues to expand into many new businesses based on wide-spread use and well-publicized success in airline and hotel industries where it is most practiced and thoroughly developed. However, RM is not necessarily a good proxy for profit maximization. Proper accounting for all costs can lead to significantly different decisions than suggested by RM. This is especially true in decisions about pricing across channels with different cost structures. When implementing scientific pricing, managers should understand the impact of their

own market demands characteristics, costs and capacity constraints on optimal pricing decisions, and to that end we hope the analysis and exposition of this paper may be helpful.

4.13 References (Chapter 4)

Bell, P. C. and J. Chen (2017). "Close integration of pricing and supply chain decisions has strategic as well as operations level benefits." Annals of Operations Research **257**(1): 77-93.

Chen, W., A. J. Fleischhacker and M. N. Katehakis (2015). "Dynamic pricing in a dual-market environment." Naval Research Logistics (NRL) **62**(7): 531-549.

Huang, G., Q. Ding, C. Dong and Z. Pan (2018). "Joint optimization of pricing and inventory control for dual-channel problem under stochastic demand." Annals of Operations Research: 1-31.

Karakul, M. (2008). "Joint pricing and procurement of fashion products in the existence of clearance markets." International Journal of Production Economics **114**(2): 487-506.

Lovell, C. A. K. and K. L. Wertz (1981). "Price Discrimination in Related Markets." Economic Inquiry **19**(3): 488.

Mankiw, N. G., R. D. Kneebone and K. J. McKenzie (2011). Principles of microeconomics. Toronto, Nelson Education.

Mills, E. S. (1959). "Uncertainty and Price Theory." The Quarterly Journal of Economics **73**(1): 116-130.

Petruzzi, N. C. and M. Dada (1999). "Pricing and the Newsvendor Problem: A Review with Extensions." Operations Research **47**(2): 183-194.

Phillips, R. L. (2005). Pricing and revenue optimization. Stanford, CA, Stanford Business Books.

ReportsnReports.com (2019). "Revenue Management Market Growing at 18.8% CAGR to 2020 - Cloud-Based Deployments Driving Growth." PR Newswire.

Smith, B. C., J. F. Leimkuhler and R. M. Darrow (1992). "Yield Management at American Airlines." Interfaces **22**(1): 8-31.

Talluri, K. T. and G. Van Ryzin (2004). The theory and practice of revenue management. Boston, Mass, Kluwer Academic Publishers.

Whitin, T. M. (1955). "Inventory Control and Price Theory." Management Science **2**(1): 61-68.

Xiao, T. and J. Shi (2016). "Pricing and supply priority in a dual-channel supply chain." European Journal of Operational Research **254**(3): 813-823.

Yue, X. and J. Liu (2006). "Demand forecast sharing in a dual-channel supply chain." European Journal of Operational Research **174**(1): 646-667.

Zhang, M., P. C. Bell, G. Cai and X. Chen (2010). "Optimal fences and joint price and inventory decisions in distinct markets with demand leakage." European Journal of Operational Research **204**(3): 589-596.

4.14 Appendix A – Deriving Constrained RM and PM Prices

Optimal capacity-constrained RM prices p_i^{*k} in equation (17), and counterpart PM prices $\overline{p}_i^{*k}$ in equation (18) are derived by a similar technique described here in detail for the former. The key is to express market 2 decisions and objective functions in terms of p_1 and make use of the first condition $q_2 = k - q_1$ under any binding capacity constraint. For both RM and PM we have a deterministic stocking decision in market 1 that is:

$$q_1(p_1) = a_1 - b_1 p_1 \tag{A1}$$

For market 2 we can write:

$$q_2(p_1) = q_2(q_1(p_1)) = k - q_1(p_1) = k - (a_1 - b_1 p_1) \tag{A2}$$

Using the inverse function of (A1) we can write:

$$p_2(p_1) = p_2(q_2(p_1)) = \frac{a_2 - q_2(p_1)}{b_2} = \frac{a_2 - (k - (a_1 - b_1 p_1)}{b_2} = \frac{a_1 + a_2 - b_1 p_1 - k}{b_2} \tag{A3}$$

For revenue maximization we have the objective function:

$$\begin{aligned} R(p_1) &= p_1 \cdot q_1(p_1) + p_2(p_1) \cdot q_2(p_1) \\ &= p_1 \cdot (a_1 - b_1 p_1) + \left(\frac{a_1 + a_2 - b_1 p_1 - k}{b_2} \right) \cdot (k - (a_1 - b_1 p_1)) \end{aligned} \tag{A4}$$

We assume a constraint is a binding on a constraint function $G(p_1) = 0$ such that the Lagrangian function $\mathcal{L}(p_1, \lambda) = R(p_1) + \lambda G(p_1)$ has partial derivative $\frac{\partial \mathcal{L}}{\partial \lambda} = 0$, and thus $\mathcal{L}'(p_1, \lambda) = R'(p_1)$ simply equal to $R'(p_1)$. The procedure is algebraically extensive, and for clarity we will begin by naming the four main components of (A4) as $R = f_1 g_1 + f_2 g_2$ and apply the chain rule $R' = f_1' g_1 + f_1 g_1' + f_2' g_2 + f_2 g_2'$ for the first order condition with respect to p_1.

$$R'(p_1) = (a_1 - b_1 p_1) + (-b_1 p_1) + \left(\frac{-b_1}{b_2} \right) \cdot (k - (a_1 - b_1 p_1)) + \left(\frac{a_1 + a_2 - b_1 p_1 - k}{b_2} \right) \cdot (b_1) = 0 \tag{A5}$$

We confirm $R''(p_1)$ is strictly negative for $p_1 > 0$ and the function has a maximum in p_1.

Shifting the first two $-b_1 p_1$ terms to the other side of the equal sign and multiplying by b_2 we obtain:

$$\begin{aligned} 2b_1 b_2 p_1 &= a_1 b_2 + \left(-b_1\right)\cdot(k-(a_1-b_1 p_1)) + \left(a_1 + a_2 - b_1 p_1 - k\right)\cdot(b_1) \\ &= a_1 b_2 + 2b_1(a_1 - k - \tfrac{a_2}{2} - b_1 p_1) \end{aligned} \tag{A6}$$

Collecting p_1 terms on the left side of the equality we have

$$\begin{aligned} 2b_1 p_1 (b_1 + b_2) &= a_1 b_2 + \left(-b_1\right)\cdot(k-(a_1-b_1 p_1)) + \left(a_1 + a_2 - b_1 p_1 - k\right)\cdot(b_1) \\ &= a_1(2b_1 + b_2) + a_2 b_1 - 2kb_1 \end{aligned} \tag{A7}$$

We can write $a_1(2b_1 + b_2) = a_1 b_1 + a_1(b_1 + b_2)$ such that (A7) becomes

$$2b_1 p_1 (b_1 + b_2) = a_1(b_1 + b_2) + b_1(a_1 + a_2 - 2k) \tag{A8}$$

Dividing both sides by $2b_1(b_1 + b_2)$ we obtain:

$p_i^* = \dfrac{a_1}{2b_1} + \dfrac{\frac{a_1+a_2}{2} - k}{(b_1 + b_2)}$ the price identified for market 1 in equation (17) for constrained RM.

The same method leads to a similar conclusion for constrained PM. Equations (A1)-(A3) continue to apply, but instead of equation (4) for revenue objective we maximize the following profit function:

$$\Pi(p_1) = q_1(p_1)\cdot(p_1(1-s_1) - (m_1 + t_1)) + q_2(p_2)\cdot(p_2(1-s_2) - (m_2 + t_2)) \tag{A10}$$

Similar algebraic manipulations as above lead to the result as in equation (18):

$\bar{p}_i^{*k} = \dfrac{a_i}{2b_i} + \dfrac{m_i + t_i}{2(1-s_i)} + \dfrac{\frac{a_1+a_2}{2} - \frac{b_1(m_1+t_1)}{2(1-s_1)} - \frac{b_2(m_2+t_2)}{2(1-s_2)} - k}{(1-s_i)\left(\frac{b_1}{(1-s_1)} + \frac{b_2}{(1-s_2)}\right)}$ in market i for constrained PM.

4.15 Appendix B – Deriving Stochastic Capacity-Constrained Decisions

Our objective is to maximize expected profit:

$$\underset{z,\,p}{Maximize}\ \mathrm{E}[\tilde{\Pi}(z,p)] \tag{B1}$$

Taking derivatives of the expected profit function (28) with respect to decision variables z_i and p_i, and noting that the partial z_i derivative of the riskless profit function $\frac{\partial}{\partial z_i}\Psi_i(z_i,p_i)=0$, we have

$$\frac{\partial E[\tilde{\Pi}_i(z_i,p_i)]}{\partial z_i}=-\frac{\partial}{\partial z_i}L_i(z_i,p_i) \tag{B2}$$

When we fully expand the R.H.S. of equation (33) it becomes

$$\frac{\partial}{\partial z_i}L_i(z_i,p_i)=\frac{\partial}{\partial z_i}(m_i+t_i)\int_{-\infty}^{z_i}(z_i-u)f_i(u)du+\frac{\partial}{\partial z_i}(p_i(1-s_i)-(m_i+t_i))\int_{z_i}^{\infty}(u-z_i)f_i(u)du \tag{B3}$$

We can subtract rather than add the last integral term by interchanging z_i and u such that we have a consistent $g(z_i,u)=(z_i-u)f_i(u)$ within both integral terms of equation (34), $\int_{-\infty}^{z}du$ and $\int_{z_i}^{\infty}du$. Using Leibniz's Rule we find that the first order partial derivative of expected profit with respect to z_i is given by

$$\begin{aligned}\frac{\partial}{\partial z_i}E[\tilde{\Pi}_i(z_i,p_i)]&=-\{(m_i+t_i)[F_i(z_i)-0]-(p_i(1-s_i)-(m_i+t_i))[1-F_i(z_i)]\}\\&=-(m_i+t_i)F_i(z_i)+p_i(1-s_i)[1-F(z_i)]-(m_i+t_i)[1-F_i(z_i)]\\&=p_i(1-s_i)[1-F_i(z_i)]-(m_i+t_i)\end{aligned} \tag{B4}$$

We note also that the second partial derivative $\frac{\partial^2}{\partial z_i^{\,2}}E[\tilde{\Pi}_i(z_i,p_i)]=-p_i(1-s_i)\cdot f_i(z_i)$ is strictly non-positive for any $0\le s_i<1$, confirming that the expected profit function is concave in z_i for any given p_i and has a maximum $E\left[\tilde{\Pi}_i^*(p_i^*,z_i^*)\right]$

Differentiating equation (28) with respect to price we obtain:

$$\frac{\partial}{\partial p_i} E[\tilde{\Pi}_i(z_i, p_i)] = (1-s_i)[a_i - 2b_i p_i - \Theta_i(z_i)] + b_i(m_i + t_i) \tag{B5}$$

Again the second partial derivative $\frac{\partial^2}{\partial {p_i}^2} E[\tilde{\Pi}_i(z_i, p_i)] = -(1-s_i) \cdot 2b_i$ is strictly negative, confirming the expected profit function is concave in p_i for any z_i and has a maximum $E\left[\tilde{\Pi}_i^*(p_i^*, z_i^*)\right]$.

Recall that the profit optimal riskless price is $\overline{p}_i^* = \frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1-s_i)}$.Thus, we can write equation (B5) as

$$\frac{\partial}{\partial p_i} E[\tilde{\Pi}_i(z_i, p_i)] = (1-s_i)[2b_i(\overline{p}_i^* - p_i) - \Theta_i(\mathrm{z}_i)] \tag{B6}$$

We see that $\frac{\partial}{\partial p_i} E[\tilde{\Pi}_i(z_i, p_i)] = 0$ when, as presented in the main paper equation (33)

$$\tilde{\tilde{p}}_l^* = \overline{p}_i^* - \frac{\Theta_i(\mathrm{z}_i)}{2b_i} \tag{33) / (B7}$$

Theorem 1 of (Petruzzi and Dada 1999) provides conditions, satisfied by our use of the uniform distribution and non-negative demand, under which for a fixed z_i optimal price can be determined by a numerical search over possible z_i for the largest one for which the derivative in equation (B4) equals zero. That is,

$$z_i^* = F_i^{-1}\left(\frac{\tilde{\tilde{p}}_i(1-s_i) - (m_i + t_i)}{\tilde{\tilde{p}}_i(1-s_i)}\right) = F_i^{-1}\left(1 - \frac{(m_i + t_i)}{\tilde{\tilde{p}}_i(1-s_i)}\right) \tag{34) / (B8}$$

Furthermore, using the uniform probability distribution, we are assuming that u falls in the range $d(p_i)-\sigma_i < u < d(p_i)+\sigma_i$ and we restrict ourselves to the domain in which $d_i(p_i) > 0$. Thus, the shortage and surplus equations (31) and (32) translate as follows:

$$\left.\begin{aligned}\Theta_i(z_i) &= \tfrac{1}{2\sigma_i}\left(\tfrac{\sigma_i^2}{2} - z_i(\sigma_i - z_i) - \tfrac{z_i^2}{2}\right)\\ \Lambda_i(z_i) &= \tfrac{1}{2\sigma_i}\left(\tfrac{\sigma_i^2}{2} + z_i(\sigma_i + z_i) - \tfrac{z_i^2}{2}\right)\end{aligned}\right\} \text{for } -\sigma_i \le z_i \le +\sigma_i \tag{B9}$$

Combining equations (B7) and (B8) and substituting the first equation of (B9) for the shortage function, we can set:

$$\frac{1}{2\sigma_i}\left(\frac{\sigma_i^2}{2} - \bar{z}_i^*(\sigma_i - \bar{z}_i^*) - \frac{\bar{z}_i^{*2}}{2}\right) = 2b_i \cdot \left(\frac{a_i}{2b_i} + \frac{m_i + t_i}{2(1-s_i)_i} - \tilde{\bar{p}}_i^*\right) \tag{B10}$$

Thus, we can establish a numeric relationship between $\bar{z}_i^*$ and $\tilde{\bar{p}}_i^*$ from the parameters. With a choice of value for either one of these variables we can determine the other.

5 Thesis Conclusion

This thesis details studies of three practical Management Science problems. They describe a variety of important MS methods with application to real-world problems and providing insights only obtainable by rigorous data processing and scientific analysis.

The first study involves a context that is purposefully entertaining to a broad non-MS audience, highlighting a relatable combinatorial problem and how it can be solved mathematically. The Integer Program developed is a non-obvious and could apply in other settings for logistics (e.g. multi-mode transport) and manufacturing (e.g. machine path planning), and especially theme parks (i.e. crowd shepherding,) It includes new analysis exploiting e-ticket technologies, with hope those results triggering further interest in recreational gaming as a subject of sports and entertainment analytics.

The second study addresses a societally important problem of how to reduce elective surgery wait-times by ensuring smooth flow of daily procedures to allow more of them to be executed without additional infrastructure. The different Constraint Programming technique used better-suited to scheduling and unique constraints as the setting demands, while it also provides a managerially-intuitive formulation in time spans and intervals. Process flow visualization tools developed in the study serve much as an elite sports team's game video, to identify which offensive blocking-scheme (schedule) works best against a given defensive set-up (OR patient load.)

The third study is of value to those new to scientific pricing, but also uncovers ground for seasoned revenue managers. Few would argue that integrated RM-SCO=PM profit maximization should be employed, ideally. The paper provides support for revenue maximization as an adequate proxy in some cases, but we see this not true, in general. Presentation of optimal pricing decisions for various scenarios (capacity, uncertainty) focuses on marginal price differences aimed at providing managers with understanding of how cost should factor into pricing decisions. As such the study can be particularly helpful in both new channel pricing, and dual market re-pricing under constrained supply.

I appreciate your interest in this study of three practical Management Science problems.

Curriculum Vitae

Name:	John Lyons
Post-secondary Education and Degrees:	University of Western Ontario London, Ontario, Canada 1980-1984 B.Sc. (Physics) The University of Western Ontario London, Ontario, Canada 1989-1991 M.B.A.
Related Work Experience	Instructor – Management Information Systems McMaster University – DeGroote School of Business 2002 Instructor – Financial Analytics University of Western Ontario – Ivey School of Business 2019
Publications:	Lyons, J. S. F., P. C. Bell and M. A. Begen (2018). "Solving the Whistler-Blackcomb Mega Day Challenge." Interfaces 48(4): 323-339.
Memberships:	Canadian Operational Research Society (CORS) Institute for Operations Research and the Management Sciences (INFORMS)