

## **Copyright Warning & Restrictions**

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be “used for any purpose other than private study, scholarship, or research.” If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use” that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

**Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation**

Printing note: If you do not wish to print this page, then select “Pages from: first page # to: last page #” on the print dialog screen

The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.

## **ABSTRACT**

### **APPOINTMENT PLANNING AND SCHEDULING IN PRIMARY CARE**

by  
**Babak Hoseini**

The Affordable Care Act (ACA) puts greater emphasis on disease prevention and better quality of care; as a result, primary care is becoming a vital component in the health care system. However, long waits for the next available appointments and delays in doctors offices combined with no-shows and late cancellations have resulted in low efficiency and high costs.

This dissertation develops an innovative stochastic model for patient planning and scheduling in order to reduce patients' waiting time and optimize primary care providers' utility. In order to facilitate access to patients who request a same-day appointment, a new appointment system is presented in which a proportion of capacity is reserved for urgent patients while the rest of the capacity is allocated to routine patients in advance. After the examination of the impact of no-shows on scheduling, a practical double-booking strategy is proposed to mitigate negative impacts of the no-show. Furthermore, proposed model demonstrates the specific circumstances under which each type of scheduling should be adopted by providers to reach higher utilization.

Moreover, this dissertation extends the single physician's model to a joint panel scheduling and investigates the efficiency of such systems on the urgent patients' accessibility, the physicians' utilization, and the patients' waiting time. Incorporating the newsvendor approach and stochastic optimization, these models are robust and practical for planning and scheduling in primary care settings. All the analytical results are supported with numerical examples in order to provide better managerial insights for primary care providers.

**APPOINTMENT PLANNING AND SCHEDULING IN PRIMARY  
CARE**

by  
**Babak Hoseini**

**A Dissertation  
Submitted to the Faculty of  
New Jersey Institute of Technology  
in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy in Industrial Engineering**

**Department of Mechanical and Industrial Engineering**

**January 2017**

Copyright © 2017 by Babak Hoseini

ALL RIGHTS RESERVED

**APPROVAL PAGE**

**APPOINTMENT PLANNING AND SCHEDULING IN PRIMARY CARE**

**Babak Hoseini**

---

Dr. Wenbo Cai, Dissertation Advisor Date  
Assistant Professor of Mechanical and Industrial Engineering, NJIT

---

Prof. Layek Abdel-Malek, Committee Member Date  
Professor of Mechanical and Industrial Engineering, NJIT

---

Prof. Sanchoy K. Das, Committee Member Date  
Professor of Mechanical and Industrial Engineering, NJIT

---

Dr. Athanassios Bladikas, Committee Member Date  
Associate Professor of Mechanical and Industrial Engineering, NJIT

---

Dr. Junmin Shi, Committee Member Date  
Assistant Professor of Martin Tuchman School of Management, NJIT

## BIOGRAPHICAL SKETCH

**Author:** Babak Hoseini  
**Degree:** Doctor of Philosophy  
**Date:** January 2017

### Undergraduate and Graduate Education:

- Doctor of Philosophy in Industrial Engineering,  
New Jersey Institute of Technology, Newark, NJ, 2017
- Master of Science in Industrial Engineering,  
Isfahan University of Technology, Isfahan, Iran, 2010
- Bachelor of Science in Industrial Engineering,  
Isfahan University of Technology, Isfahan, Iran, 2008

**Major:** Industrial Engineering

### Presentations and Publications:

- B. Hoseini, W. Cai, L. Abdel-Malek, “Improving primary care performance using stochastic carve-out scheduling”, submitted to *International Journal of Production Economics*, 2016.
- B. Hoseini, W. Cai, “Impact of overbooking in appointment scheduling of primary care services”, *INFORMS Annual Meeting*, Nashville, TN, 2016.
- W. Cai, L. Abdel-Malek, B. Hoseini, S.R. Dehkordi “Impact of flexible contracts on the performance of both retailer and supplier”, *International Journal of Production Economics*, vol. 170, pages 429-444, 2015.
- B. Hoseini, W. Cai, “Overbooking strategy in appointment scheduling: benefits and outcomes”, *Dana Knox Student Research Showcase, NJIT*, 2016.
- W. Cai, L. Abdel-Malek, B. Hoseini, S.R. Dehkordi “Flexible Contracting with Minimum Order Quantity Commitment and Expedited Delivery”, *18th International Working Seminar on Production Economics*, Innsbruck, Austria, 2014.

*This dissertation is dedicated to my beloved wife Sharareh for her kindness and devotion, and for her endless support.*



## ACKNOWLEDGMENT

I would like to express my deepest appreciation to my dissertation advisor doctor Wenbo Cai, who has shown the attitude and the substance of a genius. She continually and persuasively conveyed a spirit of perfection in regard to research and an excitement in regard to teaching. Without her supervision and constant help, this dissertation would not have been possible.

I would like to thank my dissertation committee members: Dr. Athanassios Bladikas, Professor Sanchoy Das, Professor Layek Abdel-Malek and Dr. Junmin Shi for their support not just in this dissertation but throughout my university career.

I am also indebted to the staff at the department for their continued support, especially Nicholas A. Muscara and Joseph Glaz for their guidance and support throughout my teacher assistantship.

My sincere thanks also go to Farshid and Simin Rajaei the lovely couples for all their advice and support over these years. I am also thankful to Roozbeh Rajaei for his advice and supports although he was heavily involved with so many activities.

My life as a Ph.D. student at NJIT was enriched by the company of the great friends I found here. I have some of my best life-long memories from the parties, conversations, trips and other activities that I enjoyed in the company of: Ehssan, Iman, Nastaran, Pegah, Abbas, Ali, Pezhman, Shahrouz, Abolfazl, Christeen and many others that I cannot do justice in this short note.

I would like to thank my family and friends for their endless supports, motivations, and inspirations. Specifically, I would like to thank my Mom Mastaneh and my brothers Siamak and Soheil for all that they have done and continue to do in support of my endeavors

At the end, I want to thank my Dad. I will be ever grateful for teaching me how to treat people and exploring the real perspectives of life for me. I am tremendously sad and sorry that he has not lived to see me graduate.

## TABLE OF CONTENTS

| Chapter  | Page |
|--|------|
| 1 INTRODUCTION . . . . .                               | 1    |
| 1.1 The United States Health Care Issues . . . . .     | 1    |
| 1.2 Research Motivation . . . . .                      | 2    |
| 1.3 Background and Problem Statement . . . . .         | 6    |
| 1.4 Research Objectives . . . . .                      | 10   |
| 1.5 Contribution . . . . .                             | 11   |
| 2 LITERATURE REVIEW . . . . .                          | 12   |
| 2.1 Health Care Planning and Scheduling . . . . .      | 12   |
| 2.2 Primary Care Planning and Scheduling . . . . .     | 14   |
| 2.2.1 Primary Care Performance Metrics . . . . .       | 14   |
| 2.2.2 Appointment Systems . . . . .                    | 17   |
| 2.2.3 Challenges in Primary Care Scheduling . . . . .  | 20   |
| 2.2.4 Interventions to Manage No-shows . . . . .       | 23   |
| 2.3 Scheduling Models and Methodologies . . . . .      | 26   |
| 2.3.1 Queuing Theory and Markov Chain . . . . .        | 26   |
| 2.3.2 Simulation Based Studies . . . . .               | 27   |
| 2.3.3 Stochastic Optimization . . . . .                | 29   |
| 3 SINGLE PHYSICIAN APPOINTMENT SCHEDULING . . . . .    | 31   |
| 3.1 Scheduling Modeling . . . . .                      | 31   |
| 3.2 Analytical Solutions . . . . .                     | 37   |
| 3.3 Computational Solutions . . . . .                  | 42   |
| 3.3.1 Characterizations of Optimal Schedules . . . . . | 43   |
| 3.3.2 Heuristic Schedules . . . . .                    | 49   |
| 3.4 Managerial Insight . . . . .                       | 53   |
| 4 JOINT PANEL SCHEDULING . . . . .                     | 56   |

**TABLE OF CONTENTS**  
**(Continued)**

| <b>Chapter</b>  | <b>Page</b> |
|---|-------------|
| 4.1 Joint Panel Scheduling Modeling . . . . .                 | 56          |
| 4.2 Analytical Solutions . . . . .                            | 65          |
| 4.3 Computational Solutions . . . . .                         | 72          |
| 4.3.1 Characterizations of Joint Panel Optimization . . . . . | 72          |
| 4.4 Efficiency of Joint Panel Scheduling . . . . .            | 78          |
| 4.5 Managerial Insight . . . . .                              | 79          |
| 5 CONCLUSION AND FUTURE EXTENSIONS . . . . .                  | 82          |
| APPENDIX A PROOF OF PROPOSITIONS (SINGLE PHYSICIAN) . . . . . | 84          |
| APPENDIX B PROOF OF PROPOSITIONS (JOINT PANEL) . . . . .      | 89          |
| BIBLIOGRAPHY . . . . .  | 109         |

## LIST OF TABLES

| <b>Table</b>   | <b>Page</b> |
|--|-------------|
| 2.1 Reasons of No-shows . . . . .                            | 21          |
| 3.1 Single Physician Model's Notation . . . . .              | 33          |
| 3.2 Optimal Capacity Planning Decisions in Case 1 . . . . .  | 43          |
| 3.3 Optimal Capacity Planning Decisions in Case 2 . . . . .  | 45          |
| 3.4 Optimal Capacity Planning Decisions in Case 3 . . . . .  | 46          |
| 3.5 Optimal Capacity Planning Decisions in Case 4 . . . . .  | 47          |
| 3.6 Optimal Capacity Planning Decisions in Case 5 . . . . .  | 48          |
| 3.7 Complete Enumeration ( $\gamma = 8\%$ ) . . . . .        | 51          |
| 3.8 Performance of Heuristics ( $\gamma = 8\%$ ) . . . . .   | 51          |
| 3.9 Complete Enumeration ( $\gamma = 30\%$ ) . . . . .       | 52          |
| 3.10 Performance of Heuristics ( $\gamma = 30\%$ ) . . . . . | 52          |
| 3.11 Heuristics Efficiency Comparison . . . . .              | 53          |
| 4.1 Joint Panel Scheduling Model's Notation . . . . .        | 59          |
| 4.2 Optimal Capacity Planning Decisions - Case 1 . . . . .   | 73          |
| 4.3 Optimal Capacity Planning Decisions - Case 2 . . . . .   | 74          |
| 4.4 Optimal Capacity Planning Decisions - Case 3 . . . . .   | 75          |
| 4.5 Optimal Capacity Planning Decisions - Case 4 . . . . .   | 76          |
| 4.6 Optimal Capacity Planning Decisions - Case 5 . . . . .   | 77          |
| 4.7 Joint Panel Scheduling Efficiency Comparison . . . . .   | 79          |

## LIST OF FIGURES

| Figure   | Page |
|--|------|
| 1.1 Overall ranking of OECD countries in health care-2010. . . . .   | 3    |
| 1.2 Overall ranking of OECD countries in health care-2013. . . . .   | 4    |
| 1.3 Life expectancy vs. expenditure per capita among OECD countries. . . . .   | 4    |
| 1.4 Amenable mortality in industrialized countries. . . . .  | 5    |
| 1.5 Different appointment systems. . . . .   | 9    |
| 3.1 General scenarios for optimal schedule. . . . .  | 38   |
| 3.2 Optimal schedules with $T_{O,D}$ . . . . .   | 39   |
| 3.3 Optimal schedules with $T_{D,O}$ . . . . .   | 40   |
| 3.4 Optimal schedules with $T_{D,O}^{j,j+1}$ . . . . .   | 40   |
| 3.5 Optimal schedules with $T_{D,O}^{j,N}$ . . . . .   | 41   |
| 3.6 Flowchart for optimal schedule selection. . . . .  | 41   |
| 3.7 Optimal schedule in Case 1 (Top: $D_s \sim U[0, 4]$ , Bottom: $D_s \sim U[0, 5]$ ). . . . .                                | 44   |
| 3.8 Optimal schedule in Case 2 (Top: $D_s \sim U[0, 3]$ , Bottom: $D_s \sim U[0, 4]$ ). . . . .                                | 45   |
| 3.9 Optimal schedule in Case 3 (Top: $D_s \sim U[0, 4]$ , Bottom: $D_s \sim U[0, 6]$ ). . . . .                                | 46   |
| 3.10 Optimal schedule in Case 4 (Top: $D_s \sim U[0, 4]$ , Bottom: $D_s \sim U[0, 5]$ ). . . . .                               | 47   |
| 3.11 Optimal schedule in Case 5 (Top: $D_s \sim U[0, 3]$ , Middle: $D_s \sim U[0, 4]$<br>Bottom: $D_s \sim U[0, 5]$ ). . . . . | 49   |
| 3.12 Optimal schedule using back loading heuristic for Case 4. . . . .   | 50   |
| 3.13 Optimal schedule using front loading heuristic for Case 4. . . . .  | 50   |
| 3.14 Optimal schedule using evenly spaced heuristic for Case 4. . . . .  | 51   |
| 4.1 Joint panel scheduling vs. single physician scheduling. . . . .  | 57   |
| 4.2 General scenarios for joint panel scheduling. . . . .  | 66   |
| 4.3 Optimal schedule with $T_{O,D} + T_{O,D}$ . . . . .  | 67   |
| 4.4 Optimal schedule with $T_{D,O} + T_{D,O}$ . . . . .  | 67   |
| 4.5 Optimal schedule with $T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1}$ . . . . .  | 68   |

**LIST OF FIGURES**  
(Continued)

| <b>Figure</b> |   | <b>Page</b> |
|---------------|---|-------------|
| 4.6           | Optimal schedule with $T_{D,O}^{j,N} + T_{D,O}^{h,N}$ . . . . .         | 68          |
| 4.7           | Optimal schedule with $T_{D,O} + T_{O,D}$ . . . . .                     | 69          |
| 4.8           | Optimal schedule with $T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h}$ . . . . .     | 69          |
| 4.9           | Optimal schedule with $T_{D,O}^{j,N} + T_{D,O}^{h-1,h}$ . . . . .       | 70          |
| 4.10          | Flowchart for optimal schedule selection. . . . .                       | 70          |
| 4.11          | Different types of symmetric discrete triangular distributions. . . . . | 73          |
| 4.12          | Optimal schedules: single physician vs. joint panel - case 1. . . . .   | 73          |
| 4.13          | Optimal schedules: single physician vs. joint panel - case 2. . . . .   | 74          |
| 4.14          | Optimal schedules: single physician vs. joint panel - case 3. . . . .   | 75          |
| 4.15          | Optimal schedules: single physician vs. joint panel - case 4. . . . .   | 77          |
| 4.16          | Optimal schedules: single physician vs. joint panel - case 5. . . . .   | 78          |

# CHAPTER 1

## INTRODUCTION

### 1.1. The United States Health Care Issues

Health care around the world and specifically in the United States has put a lot of effort into improving knowledge, innovation, and increasing capacity to conduct and manage the current system. Still, there are several major problems in quality, productivity, cost, and equity which are the critical attributes of a system. Shortcomings and inefficiencies in these areas may result in missed opportunities, waste, and harm to patients. Based on several reports [Schoen et al., 2006, Schoen et al., 2009, Smith et al., 2013], America's health care system has become too complex and costly for business as usual to continue. Inefficiencies, low productivity, and quality issues obstruct progress in improving health care system.

The United States health care system is unique among OECD (Organization for Economic Co-operation and Development) countries for its dependency and reliance on the private sector for financing, purchasing, and delivery of health care services. In the U.S., 80% of hospitals are non-profit, 2% are government owned and 18% are for-profit [Rosenthal, 2013]. Comparing the health care system with other industries highlight the full extent of the aforementioned shortcomings.

Automobile manufacturers produce thousands of safe and standard vehicles; through implementing systems, such as Just In Time (JIT) or Total Quality Management (TQM) they not only increased the service rate but also the total profit. Although health care system have to accommodate many competing priorities and the human factor, unlike those in other industries, the health care systems can learn from these industries how to achieve target cost and expand choices.




Financial institutes, such as banks, provide customers financial statements and records which are updated in real time as one of their customer satisfaction values. Similarly, patients would be better served by an agile health care system that is consistently reliable and that constantly improves.

## **1.2. Research Motivation**

Recently, rising health care costs in the United States is becoming an important issue. The United States spends far more money on health care per capita than any other country. Since 2004, the U.S. spent almost 50% more than the second highest country, Norway, on its health care [Davis et al., 2010, Davis et al., 2014]. For the year 2014, PwCs Health Research Institute (HRI) projected a medical cost trend of 6.5% increase [PWC, 2013]. The Congressional Budget Office (CBO) expects total spending on health care to increase 14% of the gross domestic product (GDP) by 2039, twice the 7% average seen over the past 40 years [Elmendorf, 2014]. According to the federal budget spending estimation for fiscal years 2012-2016, the health care budget increased almost \$367 billion from 2012 to 2016, which is a 42% increase.

The health care system is not only under pressure because of high expenditure but also the poor quality of care. According to World Health Organization (WHO), the ultimate goal of primary health care is “better health for all”. One of the key elements in achieving this goal is to organize health services around people’s needs and expectations. The United States health care system is not only the most expensive in the world but also has a poor ranking among OECD countries, with respect to the aforementioned critical attributes. As shown in Figures 1.2 and 1.2, the United States has the highest expenditure and lowest ranking in most of the performance metrics. This poor performance drastically affects Americans’ healthiness in terms of life expectancy, infant mortality rate, diabetes prevalence, cancer incidence and many other metrics. For example, Figure 1.3 demonstrates how life expectation

| Country Rankings                        |           |  |         |         |          |         |         |         |
|---|-----------|--|---------|---------|----------|---------|---------|---------|
|   | 1.00-2.33 | AUS  | CAN     | GER     | NETH     | NZ      | UK      | US      |
|   | 2.34-4.66 |  |         |         |          |         |         |         |
|   | 4.67-7.00 |  |         |         |          |         |         |         |
| <b>OVERALL RANKING (2010)</b>           |           | 3  | 6       | 4       | 1        | 5       | 2       | 7       |
| <b>Quality Care</b>                     |           | 4  | 7       | 5       | 2        | 1       | 3       | 6       |
| Effective Care                          |           | 2  | 7       | 6       | 3        | 5       | 1       | 4       |
| Safe Care                               |           | 6  | 5       | 3       | 1        | 4       | 2       | 7       |
| Coordinated Care                        |           | 4  | 5       | 7       | 2        | 1       | 3       | 6       |
| Patient-Centered Care                   |           | 2  | 5       | 3       | 6        | 1       | 7       | 4       |
| <b>Access</b>                           |           | 6.5  | 5       | 3       | 1        | 4       | 2       | 6.5     |
| Cost-Related Problem                    |           | 6  | 3.5     | 3.5     | 2        | 5       | 1       | 7       |
| Timeliness of Care                      |           | 6  | 7       | 2       | 1        | 3       | 4       | 5       |
| <b>Efficiency</b>                       |           | 2  | 6       | 5       | 3        | 4       | 1       | 7       |
| <b>Equity</b>                           |           | 4  | 5       | 3       | 1        | 6       | 2       | 7       |
| <b>Long, Healthy, Productive Lives</b>  |           | 1  | 2       | 3       | 4        | 5       | 6       | 7       |
| <b>Health Expenditures/Capita, 2007</b> |           | \$3,357  | \$3,895 | \$3,588 | \$3,837* | \$2,454 | \$2,992 | \$7,290 |

**Figure 1.1** Overall ranking of OECD countries in health care-2010.

Source: Calculated by the Commonwealth Fund based on 2007 International Health Policy; 2008 International Health Policy of Sicker Adults; 2009 International Health Policy Survey of Primary Care Physicians, “*OECD Health Data*”, Nov 2009, Paris.

versus the expenditure on health care in the United States is lower when compared to other countries. In another study, [Nolte and McKee, 2008] compare rates of “amenable mortality”, which are deaths from certain causes before age 75 that are potentially preventable with timely and effective health care among industrialized countries. Figure 1.4 shows the total ranking of industrialized countries with respect to amenable mortality. The most important cause for such disparity in health care is the failure to emphasize primary care within health care system. Countries that have focused on primary care have better health outcomes at lower cost than the United States [Siferd and Benton, 1992]. Studies show countries with weak primary care infrastructures have poorer performance on major aspects of health [Starfield and Shi, 2002, Sandy et al., 2009].

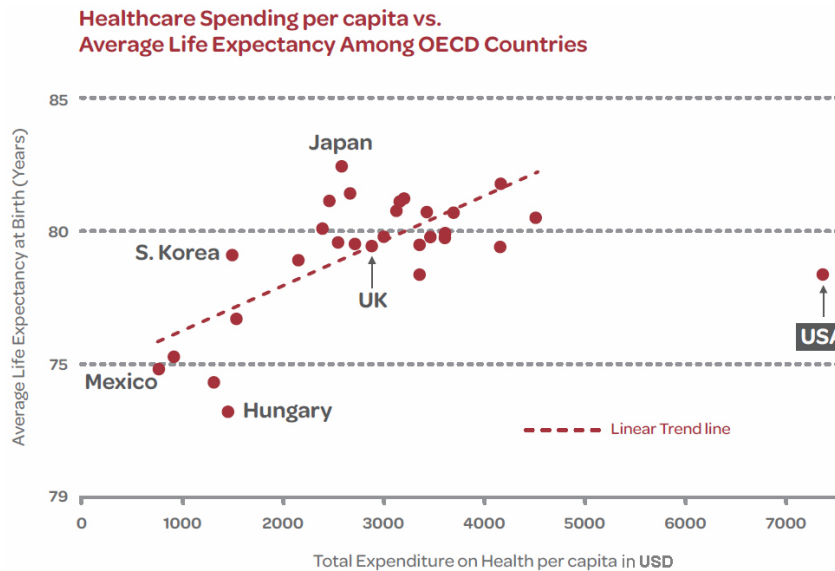
The United States currently has a gap between the number of primary care providers and the number that would be needed to deliver primary care to its’ population. Indeed, between the current health care system and the health care system that we could have lies not just a gap, but a chasm [Bloom, 2002]. According to another study, by 2035 more than 44,000 primary care physicians will be needed and

**COUNTRY RANKINGS**

Top 2\*  
Middle  
Bottom 2\*

|   | AUS     | CAN     | FRA     | GER     | NETH    | NZ      | NOR     | SWE     | SWIZ    | UK      | US      |
|---|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| <b>OVERALL RANKING (2013)</b>             | 4       | 10      | 9       | 5       | 5       | 7       | 7       | 3       | 2       | 1       | 11      |
| <b>Quality Care</b>                       | 2       | 9       | 8       | 7       | 5       | 4       | 11      | 10      | 3       | 1       | 5       |
| Effective Care                            | 4       | 7       | 9       | 6       | 5       | 2       | 11      | 10      | 8       | 1       | 3       |
| Safe Care                                 | 3       | 10      | 2       | 6       | 7       | 9       | 11      | 5       | 4       | 1       | 7       |
| Coordinated Care                          | 4       | 8       | 9       | 10      | 5       | 2       | 7       | 11      | 3       | 1       | 6       |
| Patient-Centered Care                     | 5       | 8       | 10      | 7       | 3       | 6       | 11      | 9       | 2       | 1       | 4       |
| <b>Access</b>                             | 8       | 9       | 11      | 2       | 4       | 7       | 6       | 4       | 2       | 1       | 9       |
| Cost-Related Problem                      | 9       | 5       | 10      | 4       | 8       | 6       | 3       | 1       | 7       | 1       | 11      |
| Timeliness of Care                        | 6       | 11      | 10      | 4       | 2       | 7       | 8       | 9       | 1       | 3       | 5       |
| <b>Efficiency</b>                         | 4       | 10      | 8       | 9       | 7       | 3       | 4       | 2       | 6       | 1       | 11      |
| <b>Equity</b>                             | 5       | 9       | 7       | 4       | 8       | 10      | 6       | 1       | 2       | 2       | 11      |
| <b>Healthy Lives</b>                      | 4       | 8       | 1       | 7       | 5       | 9       | 6       | 2       | 3       | 10      | 11      |
| <b>Health Expenditures/Capita, 2011**</b> | \$3,800 | \$4,522 | \$4,118 | \$4,495 | \$5,099 | \$3,182 | \$5,669 | \$3,925 | \$5,643 | \$3,405 | \$8,508 |

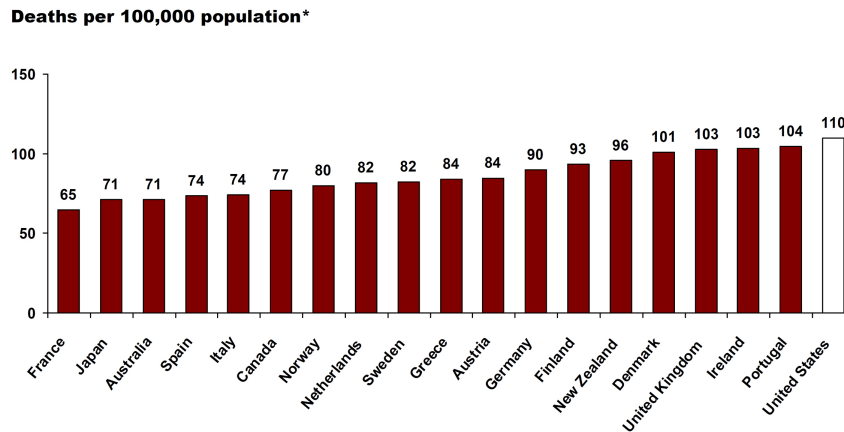
**Figure 1.2** Overall ranking of OECD countries in health care-2013.  
 Source: Calculated by the Commonwealth Fund based on 2011 International Health Policy; 2012 International Health Policy of Sicker Adults; 2013 International Health Policy Survey of Primary Care Physicians, “*OECD Health Data*”, Nov 2013, Paris.



**Figure 1.3** Life expectancy vs. expenditure per capita among OECD countries.  
 Source: Calculated by the Commonwealth Fund based on 2007 International Health Policy; “*Health at a Glance*”, 2009

with current primary care production rates, the United States will face a shortage of 33,000 primary care physicians [Petterson et al., 2015]. Despite numerous publications and reports on deficiencies within the primary care system, the U.S. needs more effort to strengthen primary care.

### Mortality Amenable to Health Care, 2002-03



**Figure 1.4** Amenable mortality in industrialized countries.

Source: Adapted from E. Nolte and C. M. McKee, *Measuring the Health of Nations: Updating an Earlier Analysis, "Health Affairs"*, Jan 2008, 27(1):5871.

According to [Starfield, 1998], a good primary care system should consist of first-contact (initial outreach to primary care practitioner), maintaining continuity of care over time, comprehensiveness and be coordinated with various aspects of the health care system. Institute of Medicine (IOM) Committee on the Quality of Health Care in America declares two main goals for primary care in the U.S. The first goal is providing care that is responsive to individual patients' preferences, needs, and values and the second goal is reducing waits and sometimes harmful delays for both recipient and practitioner.

Although, the recent Affordable Care Act in 2010 tried to expand the coverage among the uninsured population. The question is: who will deliver primary care to the new cohort of patients?

To answer this question, one needs to acknowledge the crisis at hand and be open to the opportunities that are brought forth by this challenge. It is because of this formidable task that the interaction and the attention of policy makers are an absolute need. According to [Sepulveda et al., 2008], if the primary care foundation of the health care system is not fortified, access and cost shaving may be impossible.

Since primary care is considered as a day-to-day care, the practitioner acts as the principal responsible party for the care of patients within the health care system. In recent years, delays in acquiring appointments and long office waiting times have negatively impacted patients' satisfaction, and therefore, become problematic for primary care providers. Based on a survey, 27% of insured adults younger than 65 with health problems, had difficulty getting a timely appointment from their physicians. Moreover, reports show 43% of adults, with an urgent condition, have received care with noticeable delay [Murray and Berwick, 2003].

It is evident that the main cause of the current crisis in primary care in the U.S. is the difficulty in matching supply with demand. However, the Affordable Care Act (ACA) consists of many provisions and recommends strategies to improve access to primary care service centers by increasing the supply of practitioners. These strategies should be considered as a longer-term effort to boost the primary care workforce and may be insufficient due to the fact that a significant increase in new primary care practitioner will take decades to meet trending demand. Consequently, primary care requires more effort to become accessible and efficient. Therefore, reducing waiting times and governing pertinent goals of primary care, such as continuity of care and accessibility results in an increase of quality of care while reducing costs.

### **1.3. Background and Problem Statement**

There is a meaningful difference between primary health care and primary care. According to the World Health Organization (WHO) and the United Nations International Childrens Emergency Fund (UNICEF), primary health care (PHC) is “essential health care based on practical, scientifically and socially acceptable methods and technology made universally accessible through their full participation and at a cost that the community and country can afford”. However, primary care (PC) is “Delivery of a complex set of services, which include the first contact and the

maintenance care, and assumes responsibility for the referral to distinct services in response to the client needs and cultural values” [Barnes et al., 1995]. Consequently, primary care is a subset of primary health care.

Primary care is substantiated within the health care system in which health promotion, disease prevention, health maintenance, counseling, patient education, diagnosis and treatment of acute and chronic illnesses, are provided in a variety of health care settings, such as office, clinics, critical care, short/long-term care and home care. Primary care, considered as a same day care, is provided by a primary care practitioner, such as family physician, general practitioner, geriatrician, pediatrician, nurse practitioner and physician assistants, who act as the first contact and the principal point of continuing care for patients. The practitioners also coordinate other specialists’ services as needed.

In the 1990s, long waits and delays in hospitals and doctors’ offices were considered as status quo [Goitein, 1990]. Many primary providers had difficulty in providing timely care which unnecessarily coerced them into hiring more physicians and staff to reduce backlogs and waiting time, which resulted in an inefficient system. Waiting time consists of two subcategories, direct waiting time and indirect waiting time. *Indirect waiting time* is the difference between the time that an appointment is requested and the time of the actual appointment. *Direct waiting time* (in-office delay) is the difference between a patients appointment time and the time when he/she is actually seen by the physician.

In order to address long waiting times, many primary care providers across the United States implemented different appointment systems for patient scheduling. In the crude model or *traditional model*, all the available appointments are reserved ahead of time, usually two weeks to a month before the allotted date, and urgent patients are squeezed in by double-booking. In this appointment system, the capacity of a provider is fully utilized due to the overabundance of bookings. Additionally,

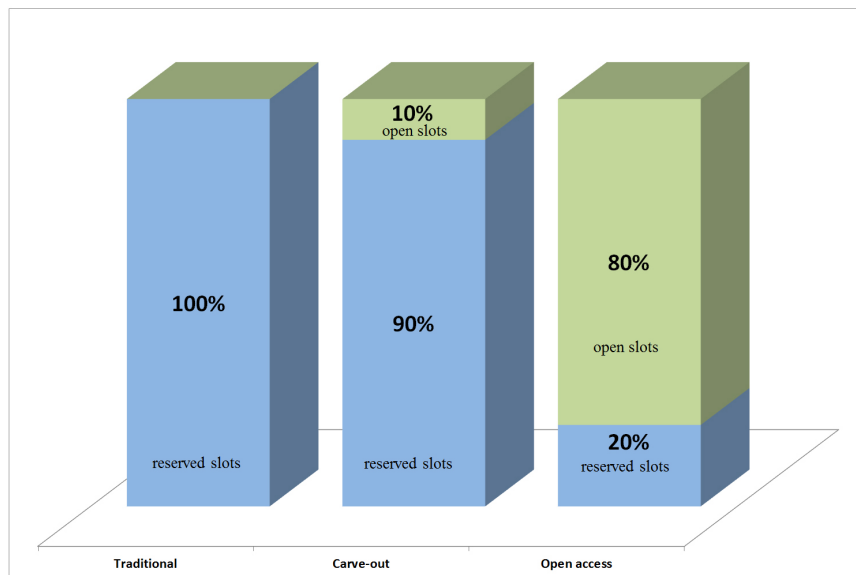
the provider should use a complex matrix appointment format to schedule routine physicals and follow-up appointments. It is of note that such systems contain an inherent disruption of care due to the incompatibility of the schedule. The practice of overbooking and disruption of urgent patients' needs contribute to further prolong waits, which propel patients to seek medical treatment via urgent care, i.e.; emergency department or various urgent care clinics resulting in further disruption of care and increase in costs within the health care system. The slogan for these systems is, "Do last month's work today" [Murray and Tantau, 2000].

In the late 90's and after development of demand prediction methods, a new scheduling system was introduced by researchers called *carve-out model*. In this approach, a proportion of a whole capacity remains open in order to service the urgent patients each day, and the rest of allotted times are booked in advance, like the traditional model. Although, this model is better than the traditional model, it still contains a considerable amount of problems. The patient who calls for a non-urgent appointment will be pushed to the next open slot (potentially next week). Moreover, when a patient calls for an appointment on a certain date that is full, the provider asks the patient to call back again on the requested date for an urgent appointment, which results in no set appointment for that patient and no guarantee of the patient calling back. In addition, due to high uncertainty of demand, there is low precision in estimating number of open slots which lead to a waste of capacity or potential over loading capacity.

The third generation of appointment systems for primary care, developed in the beginning of the 20 century is called *open access*. Open access, also known as advanced access or same-day scheduling, is a method of scheduling in which almost all patients can receive an appointment on the day they call, almost always with their personal physician. Theoretically, this approach leads the provider to have no waiting time and improve continuity of care. In this approach rather than booking style in which

allotted time slots were reserved a week in advance, most of the slots are left open in order to serve the same-day patients. The slogan for this approach is straightforward “Do today’s work today” [Murray and Tantau, 2000]. This approach brings various benefits for the primary care patients and providers. It increases the availability of appointments for urgent patients in the same day. In addition, accessibility for routine patients, who are seeking appointment, is maintained without further indirect waiting time. This approach also increases the likelihood of a patient see his/her own personal physician, which improves continuity of care and patients’ satisfaction.

Open access approach improves customer service and quality levels and address the two major goals of primary health care defined by World Health Organization, accessibility and continuity of care, but implementation of this system is not easy and if a provider implements this system inefficiently or improperly, for example with shortage in staff, irregular physician availability and etc., the consequences can become more inefficient and potentially disruptive to both patients and practitioners [Phan and Brown, 2009, Mehrotra et al., 2008]. Figure 1.5 demonstrates the three aforementioned appointment systems.



**Figure 1.5** Different appointment systems.



In addition, designing a good appointment system is just one feature of the appointment planning system in which problematic scenarios, such as no-shows, cancellations, and walk-ins should be taken into account. The provider may want to use different strategies to handle these issues to minimize risks and maximize profit. The appointment planning is the process of reconciling supply (available appointment) and demand (routine patient and urgent patients), where the provider should choose the general setting of an appointment system, such as number of available slots, maximum allowable overtime hours, acceptable waiting times, etc.

While planning is limited to the general setting of the appointment system, appointment scheduling refers to assigning patients to available slots based on daily demand. However, the daily demand is uncertain; in primary care, most of the patients require services that can be performed within a fixed length of time which allows providers to divide their time into evenly devised time slots. In contrast, in a primary care environment, there are unscheduled encounters like no-shows, unpunctuality, late cancellation, and walk-ins which cause more uncertainty to patients scheduling. Therefore, many strategies have been developed to handle these uncertainties, such as block-booking, patient classification (age, disease type, socioeconomic level, etc.) and reserving slots, to name a few. In addition, the schedule itself can be static or dynamic, which affects the scheduling manner and assumptions. In the static case, the schedule is set before the beginning of the session usually the morning of, while in dynamic case, the schedule can be revised based on new appointment requests, walk-ins, and delays.

#### 1.4. Research Objectives

This dissertation accomplished the following research objectives are as follows:  
**Research Objective 1:** develop a joint newsvendor and stochastic model that maximizes the social welfare, which include the provider's utilization, staffs' overtime

cost, and patients' waiting times for a carve-out scheduling appointment system.

**Research Objective 2:** find the optimal number of slots reserved for both routine and urgent patients. Due to the large computation time associated with computing the optimal solutions, we also develop heuristics to estimate these results.

**Research Objective 3:** derive a daily schedule in which each available slot will be assigned to either an urgent patient or a routine patient, and whether the double booking is allowed if it's reserved to a routine patient.

**Research Objective 4:** obtain the conditions under which a certain schedule is optimal for both the single-physician scheduling and the joint panel scheduling.

**Research Objective 5:** examine the effect of no-shows and walk-ins on appointment scheduling.

## 1.5. Contribution

This dissertation intends to make three major contributions to the literature as follows: (i)develop an innovative model that considers a stochastic number of routine and urgent patients, no-shows, and overbooking in a carve-out scheduling system. To the best of our knowledge, our work is the first that considers both the planning and scheduling aspects of both the single physician and the joint panel systems. (ii)derive the conditions under which a certain schedule is optimal, given the number of reserved slots for routine and urgent patients, and the maximum number of slots allowed for double-booking. (iii)develop heuristics to find the optimal number of reserved slots for routine and urgent patients, and the maximum number of slots allowed for double-booking to largely reduce the run-time.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1. Health Care Planning and Scheduling

Providing of high quality and affordable health care is one of the greatest challenges for most countries. Increasing the efficiency of health care resources (doctors, nurses, and medical equipment) can help meeting patients' needs with reducing patients' waiting time, improving continuity of care, and health outcomes as a whole.

Many studies have investigated efficiency improvement method through capacity planning [Green, 2004, Hick et al., 2004, VanBerkel and Blake, 2007, Exadaktylos et al., 2008], staff scheduling [Siferd and Benton, 1992, Dowsland, 1998, Jaumard et al., 1998, Bourdais et al., 2003, Centeno et al., 2003] and patient oriented scheduling, such as ambulatory care, surgeries and outpatient scheduling.

Patient oriented scheduling is used in a wide variety of environments, such as operation rooms, ambulatory care, primary care, radiology and anywhere who dealing with many pre-scheduled and walk-in patients. [Adan and Vissers, 2002], [Dexter and Macario, 2002], [Harper and Gamlin, 2003], [Ballard and Kuhl, 2006], [Beliën et al., 2006], [Van Houdenhoven et al., 2007], and [Cardoen and Demeulemeester, 2008] present planning and scheduling in operating room and elective or non-elective surgeries. [Cardoen et al., 2010] reviews more than 120 different articles about operating room planning and scheduling in order to analyze their contribution.

Planning and scheduling in primary care are more complex than other scheduling in a health care system. Unlike surgical appointment scheduling or other appointment systems in a health care system, in primary care scheduling, we have more uncertainties, such as random demand on different days, patient no-shows, and

appointment cancellations, while the schedule is constrained by lower flexibility, such as limited overtime availability, limited staff and referral patients.

Planning and scheduling in primary care have attracted the interest of many researchers. [Welch and Bailey, 1952] and [Lindley, 1952] are the pioneers on this topic. After their early studies, many researchers started working on primary care scheduling with different methods, which can be classified into two general categories: static and dynamic models.

In static models the schedule for the day should be ready prior to the beginning of that day, which is a common practice in most appointment-based systems in health care. Most of the publications focused on static scheduling, such as [Liu and Liu, 1998], [Bosch and Dietz, 2000],[Lau and Lau, 2000], [Swisher et al., 2001], [Denton and Gupta, 2003], [Robinson and Chen, 2003], [Muthuraman and Lawley, 2008], and [Chakraborty et al., 2013], in which they primarily focused on the punctual patients with independent and identically distributed service times, who are scheduled for a single session with a single server.

On the other hand, in dynamic case, the schedule can be revised continuously based on the current status of the system. The literature on dynamic scheduling is sparse. [Klassen and Rohleder, 1996], [Klassen and Rohleder, 2004], [Patrick et al., 2008], [Liu et al., 2010], [Erdogan et al., 2015], [Hahn-Goldberg et al., 2012], and [Huang and Zuniga, 2012] present methods to dynamically schedule patients. The application of dynamic case is limited in situations where patients are already admitted to a hospital or a clinic and scheduling should be done for laboratory or other examinations.

This dissertation focuses more on planning and scheduling in primary care setting. Therefore, various pertinent details and literature review will be presented respectively.

## 2.2. Primary Care Planning and Scheduling

Scheduling in primary care is often concerned with matching demand to the available resources to provide a service. The purpose of primary care scheduling is to find an appointment system in order to optimize performance metrics (cost, waiting time, idle time, continuity of care, etc.) in a clinical environment under demand uncertainty. Therefore, one can classify literature based on appointment system design, objectives of the study (performance metrics) or based on the applied methodology.

### 2.2.1 Primary Care Performance Metrics

**Minimizing patients' waiting time and practitioners' idle time.** Patient waiting time is an important factor of quality of service offered by primary care providers. While long waiting times are a major source of patient dissatisfaction and can cause a negative effect on patient treatment, the doctors' and staffs' idle time is a loss of sale for the provider. Therefore, hiring more doctors and staff is not an optimal solution to reduce the patients' waiting time and appease providers' needs of cost control.

The early studies on patient waiting time and staffs' idle time use queueing models started by [Bailey, 1952]. He attempts to balance waiting and idle times through a mathematical queueing model. His procedure suggests giving patients appointments at regular intervals, each equal to the average consultation time. [Jansson, 1966] shows how the sum of waiting time and idle time can be minimized by proper choices of the constant inter-arrival time and the initial numbers of customers in the queue. Later [Fetter and Thompson, 1966] introduce seven variables which affect the balance between patient waiting time and staff idle time i.e. appointment interval, service time, patients' arrival pattern, the number of no-shows, the number of walk-ins, physicians' arrival pattern, and interruptions in patient services respectively. [Birchall et al., 1983] show that their queueing model can reduce the idle time of the

clinic staff without increasing patient waiting times. [Brahimi and Worthington, 1991] develop a queueing model which is successfully implemented to the problem of designing an appropriate appointment system for the out-patient department at the Royal Lancaster Infirmary at that time. After these early studies [Fries and Marathe, 1981], [Babes and Sarma, 1991], [Ho and Lau, 1992], [Paul, 1995], [Aharonson-Daniel et al., 1996], [Liu and Liu, 1998], [Lau and Lau, 2000], [Robinson and Chen, 2003], [Muthuraman and Lawley, 2008], and [Hassin and Mendel, 2008] develop different models to minimize patients' waiting time in primary care systems.

[Harper and Gamlin, 2003] develop a detailed simulation model of an Ear, Nose and Throat (ENT) outpatient department to identify critical factors that influence patient waiting times and the buildup of queues in the clinic. They show how patient waiting times can be significantly reduced through improved planning of the schedule and management of the clinic. [Denton and Gupta, 2003] demonstrate that the optimal schedule in a practice that is not overloaded depends both on the ratio of the cost of handling an urgent overflow patient to the cost of delaying a patient and on the arrival dynamics of the urgent and routine patients. [Kaandorp and Koole, 2007] derive a local search procedure to minimize weighted average of expected waiting times of patients, the idle time of the doctor and the tardiness (lateness) as objective. [Green and Savin, 2008] present two queueing models to provide guidelines to enhance identifying an appropriate balance between physician capacity and patient panel size that are consistent with manageable patient backlogs. [Chen et al., 2010] conduct a cross-sectional study in a community hospital in China and demonstrate how the adoption of an appointment system can reduce waiting time.

[Zhu et al., 2012] analyze an appointment scheduling systems in specialist outpatient clinics (SOC) in Singapore to detect important factors causing long patient waiting time or clinic overtime. Results of their study show overloaded session, late start of the session, unevenly distributed slots, irregular calling sequence, and unused

session time are the most important causes of delaying patients. In a very similar study which is conducted in an Indonesian public hospital, [Mardiah and Basri, 2013] show how applying “doctor on call” system may lead to high patients’ waiting times.

**Minimizing cost of the system while maximizing providers’ profit.**

As described in Section 1.2, with an ever-growing primary care cost, the pressure to remain profitable has undoubtedly intensified for many office-based primary care physicians. This challenge will become even more intensified as patient demand increases, the supply of available primary care practitioners decreases and the focus on the health care sector shifts to overall cost reduction. Therefore, health care providers have the option to increase their profits by seeing more patients and/or by cutting their costs through more efficient utilization of resources to maintain their profitability.

Many literature consider waiting cost, idle time cost and overtime cost as a function which should be minimized. In the contrast, other studies consider different methods to increase patient visitation in a day using overbooking, short time window, delegation of work, utilizing overtime, and group medical appointments.

[Shonick and Klein, 1977] develop a model to overbook patients such that a maximum number of patient visitation is maintained. [Kim and Giachetti, 2006] develop a stochastic overbooking model to determine the optimal number of appointments in order to maximize expected total profits of the provider. [LaGanga and Lawrence, 2007] develop a stochastic model to capture the trade-off between overbooking and relative costs (waiting time and overtime costs). [Muthuraman and Lawley, 2008] develop a myopic scheduling policy to reduce patient waiting time and staff overtime and increase revenue. [Patrick, 2012] develop a Markov decision process model that shows a short booking window can reduce costs and increase staff utilization. [Qu et al., 2012] introduce a mean-variance model in order to find the

optimal percentage of open slots for an advance-access clinic to maximize the average number of patients seen and minimize variability in the number of patients seen.

If the reader is interested, there is an extensive review of appointment scheduling literature by [Cayirli, 2003] which reviews more than 80 papers concerning this topic.

### **2.2.2 Appointment Systems**

Different appointment systems have been proposed to achieve all main goals of PC (first contact, continuity, comprehensive, coordination with other parts of PHC) and reduce costs. The most popular systems for appointment scheduling are traditional, carve out model, and open access (advance access). Although these three systems have a common goal, they follow different approaches and as mentioned in Section 1.3, each of them has their pros and cons. In the following previous literature on these three systems is discussed in detail.

**Traditional scheduling.** In this system patients call the receptionist and usually after a telephone triage by a nurse are classified to an urgent or non-urgent patient. Most of the time urgent patients will be handled by double booking, however, in some cases if the schedule is saturated the practice may send the patient to another doctor or clinic [Murray and Berwick, 2003] which result in disruption of continuity of care.

[Vissers, 1979] propose a simple general method for determination of a suitable appointment system using already known mean and standard deviation of the consultation time, patient punctuality and physician punctuality. [Ho and Lau, 1992] shows how the optimal schedule can change in different environments (different no-show probability, service time and the number of patient per session). [Shonick and Klein, 1977] and [LaGanga and Lawrence, 2007] both use overbooking strategy in a traditional scheduling system to mitigate the negative effect of no-shows in order to increase providers' productivity. [Zacharias and Armony, 2016] proposed



a model to determine panel size and number of offered appointment slots per day in a traditional scheduling model in order to reduce patients' backlogs and waiting time. [Muthuraman and Lawley, 2008] develop an appointment scheduling using stochastic overbooking model to minimize patients' waiting time and staffs' overtime and increase patients' utility while all the appointments are fully booked prior to the beginning of the scheduling. [Liu and Liu, 1998] investigate the impact of doctors' arrival pattern on minimizing doctors' idle time and patients' waiting time through a traditional scheduling system.

**Carve-out scheduling.** In spite of the traditional system, in carve-out scheduling model, some slots are left open for urgent patients to increase accessibility and reduce waiting times. [Klassen and Rohleder, 1996] present a carve-out system in a dynamic environment to reduce patients' waiting time and doctors' idle time simultaneously and shows the best schedule can be obtained if clients with large service time standard deviations are scheduled toward the end of the appointment session. Furthermore, [Klassen and Rohleder, 2004] demonstrate that in a carve-out model the best strategy to reduce waiting times is to spread open slots evenly over the day. [Gupta and Wang, 2008] propose a methodology to decide how to manage open slots when patients choose between a same-day slot (non-convenient) and a future appointment (convenient). [Dobson et al., 2011] use the carve-out system to reduce the average waiting queue for routine patients and the average number of urgent patients who are not handled during the normal office hours. Their results show that the carve-out system is the best system when the incurred cost for an urgent overflow patient is high. [Wang and Gupta, 2011] develop a framework based on the carve-out model for appointment system which can dynamically use patients' preferences to improve future schedules. [Koeleman and Koole, 2012] developed a method to allocate urgent patient to the open slots in order to reduce waiting times and idle time and overtime. Recently [Cayirli and Gunes, 2014] use the carve-out

system in order to handle walk-in patients. Their results indicate that leaving some slots open in order to handle seasonal walk-ins on a weekly or monthly basis will improve clinic performance in terms of patients' waiting time and doctors' idle time.

**Open access.** Open access scheduling, which is introduced by [Murray and Tantau, 1999] and [Herriott, 1999], uses the same concept of reserving slots but in different manner and as reported can improve health care access, patient satisfaction and reduces health care cost and notably decrease patients' no-shows rate [Murray and Tantau, 2000, Murray and Berwick, 2003, OHare and Corlett, 2004, Pierdon et al., 2004, Armstrong et al., 2004, Bundy et al., 2005, OConnor et al., 2006, Rohrer et al., 2007]. However, shortage in capacity, physician availability, and high demand volatility may result in failure in implementation [Murray et al., 2003, Mehrotra et al., 2008, Phan and Brown, 2009] which show the challenge in implementing open access scheduling.

In the open access scheduling, sometimes labeled as advanced access or same-day scheduling, most of the patients can receive an appointment within 12 to 72 hours regardless of their reasons for the visit [Qu et al., 2007] and routine check-up can be scheduled weeks in advance to accommodate follow-ups. In most of the literature, researchers proposed varieties of quantitative or experimental approach to determine the appropriate open slot percentage and optimal panel size. [Green et al., 2007] suggests that the ratio of the average daily demand for appointments to the average daily capacity should be kept below 25% in order to offer same-day appointments to most of the patients. [Dobson et al., 2011] examine the effect of open-access on two different service quality measures. First, the average number of urgent patients that are not handled during normal hours and secondly the average queue of routine patients. They use a stochastic model to set the appointment scheduling without considering no-show. Their numerical results show when the demand for same-day is not high, the optimal policy depends on overtime cost and cost of patients' waiting

time. [Qu et al., 2007] present a quantitative approach in order to determine the optimal percentage of open slots in an advance-access system. They also investigate the sensitivity of this optimal percentage to provider capacity and no-show rate. [Qu et al., 2011] propose a hybrid open access system with two time horizons (days-ahead and same-day appointments) and show the performance of the whole system, in terms of the maximum expectation and the minimum variance of the patients' consultation time, is improved slightly via the hybrid open access policy.

### 2.2.3 Challenges in Primary Care Scheduling

No-shows, late cancellations, and walk-ins are another sources of uncertainty in appointment scheduling which should take into account as discussed in Section 1.3. In this section, we will focus more on these uncertainties in details.

**No-shows and late cancellations.** Missed appointments and late cancellations (hereafter referred to as no-shows) can create financial, utilization, and continuity issues in primary care. Lots of studies have investigated the causes and effects of no-show on primary care performance metrics. Lots of researchers tried to capture no-shows' effects in appointment scheduling or at least mitigate its negative impact on scheduling by developing different strategies like overbooking, block booking, and ad hoc calling patients which are discussed in Section 2.2.4. Reasons of no-shows are investigated in literature in various aspects which briefly described in Table 2.1. The average percentage of no-shows is reported 5% to 40%. According to [Izard, 2005] the average rate of no-shows nationally in 2000 was 5.5 percent. Average no-show rate reported 30.9% in pediatric clinics [Rust et al., 1995], 33.6% in ENT out-patient clinics [Geraghty et al., 2008], 30.1% in gynecological practices [Dreiherr et al., 2007], 12%-24% in endoscopy [Berg et al., 2013], 29%-42% in addiction treatment settings [Weisner et al., 2001], and 23.1% in academic outpatient

**Table 2.1** Reasons of No-shows

| Reasons                              | Studies  | Results  |
|--------------------------------------|--|--|
| Long indirect waiting time           | [Benjamin-Bauman et al., 1984],<br>[Bean and Talaga, 1994],<br>[Murdock et al., 2002],<br>[Lee et al., 2005],<br>[Gallucci et al., 2005],<br>[Norris et al., 2014] | Chance of missed appointment increases in intervals between scheduling an appointment and the actual appointment date.                                   |
| Prior appointment adherence          | [Dove and Schneider, 1981],<br>[Bean and Talaga, 1992],<br>[Norman and Conner, 1996],<br>[Lee et al., 2005],<br>[Norris et al., 2014]                              | Patients who failed to keep appointments usually had such behavior in the past.  |
| Psychological or behavioral problems | [Goldman et al., 1982],<br>[Neeleman and Mikhail, 1997],<br>[Mitchell et al., 2007],<br>[DuMontier et al., 2013]   | Missed appointment rate is higher than normal no-show rate in primary care specially in substance abuse services and geriatric psychiatry.               |
| Emotional barriers                   | [Lacy et al., 2004]  | Negative emotions about going to see the doctor outweigh the perceived benefit of keeping the appointment.   |
| Social class                         | [Bean and Talaga, 1992],<br>[Neal et al., 2001],<br>[Schechtman et al., 2008]  | Patients of lower socioeconomic classes, the poorly educated, and those who lives in deprived area are most likely to miss their scheduled appointments. |
| Age                                  | [Dove and Schneider, 1981],<br>[Bean and Talaga, 1992],<br>[Smith and Yawn, 1994],<br>[Lee et al., 2005]   | Patients age forty-five and younger, especially pediatric patients, generally have a higher rate of broken appointments.                                 |
| Race                                 | [Schechtman et al., 2008],<br>[Parker et al., 2012],<br>[DuMontier et al., 2013]   | The areas' ethnicity may effect no-show rate (Latinos and African-Americans are at highest risk of missing appointments*)                                |

\*Kaiser Permanent Diabetes Study of Northern California, 2005-2007, [Parker et al., 2012].

practice [Parikh et al., 2010]. [Hixon et al., 1999] reported that no-show rate in 33% of family practice residency clinics is greater than 21%.

Moreover, no-shows not only disrupt the scheduling of the primary care provider but also reduce overall efficiency, increase operating costs while decreasing revenue, and waste resources. Most of the patients, who missed their appointments, often use emergency departments which drives up costs and jeopardize the continuity of care, particularly in patients with chronic disease, such as diabetes, epilepsy, and bronchiectasis to list a few.

**Walk-ins.** Decades ago walk-ins were rarely accepted because the physicians' schedule were almost full, but recently, especially in the U.S., most of the clinics are responsible for the patient's total care. Therefore, walk-ins must be considered and planned for in the appointment scheduling of clinic sessions. Studies show the probability of walk-in without an appointment is higher among lower socioeconomic status [Taylor, 1984, Virji, 1990].

Indirect waiting time for walk-in patients is negligible but direct waiting time may be longer than patients who reserved an appointment in advance. The presence of walk-ins in literature is insignificant. In an early study [Rising et al., 1973] analyze daily arrival patterns in order to schedule more patients during periods of low walk-in demand. Recently [Cayirli et al., 2012] introduce a procedure for appointment systems to reduce negative impacts from walk-ins, later they extended their investigation on developing an appointment system to deal with the feasibility of walk-in [Cayirli and Gunes, 2014].

**Unpunctuality.** Unpunctuality usually refers to the difference between an actual arrival of a patient and time of the reserved appointment. However, based on different studies patients arrive early more often than late [Klassen and Rohleder, 1996, Lehaney et al., 1999]. According to [White and Pike, 1964] average waiting time in a primary care clinic with tardy patients does not greatly differ than having

punctual patients. [Tai and Williams, 2012] demonstrate, through an empirical study, that optimal appointment intervals are highly related to patients' unpunctuality patterns. Recently, [Williams et al., 2014] examine the effects of an intervention to reduce patient unpunctuality.

#### **2.2.4 Interventions to Manage No-shows**

According to Section 2.2.3, the no-show is one of the reasons for low efficiency and poor quality in primary care. Therefore, lots of researchers suggest primary care provider to use interventions to reduce no-show rates or considering a backup plan to mitigate its negative impact. We will discuss different intervention methods in this section.

**Reminders.** Generally, primary care providers continue to rely on several different types of interventions which include telephone reminders, Short Message Service, and emails, which reported to produce relative reductions in the no-show rate [Macharia et al., 1992, Almog et al., 2003, Downer et al., 2005, Geraghty et al., 2008]. Regardless of the type of intervention relied upon, these methods can be time-consuming and costly when not automated.

**No-show fees.** Another common practice among primary care providers is implementing a policy to charge a monetary fine for patients that do not show up for set appointments in order to recover a proportion of potential revenue. This method is quite common in primary care. However, it has several downsides which include:

- Legal difficulty, which may prohibit charging a no-show fee.
- Complicated policies for the collection of a no-show fee.
- Assessing an amount for a monetary fee is difficult.
- It may not be appropriate for the elderly or uninsured patients.
- Losing customers who do not approve the no-show fee.

The impact of the no-show fee is explored in different studies and reports show it can cause a decrease in no-show rate [Mäntyjärvi, 1994, Lesaca, 1995]. Recently, in Family Medical Associates of Raleigh in North Carolina, this policy is implemented and the no-show rate has dropped from 12-15 percent to 6 percent [Lowes, 2005]. In another report, in primary care clinic of the Geneva University Hospitals, the rate of missed appointments was evaluated at 22% in 2007 while charging the no-show fee (20 EUR for each missed appointment) did not reduce the no-show rate [Perron et al., 2010].

**Overbooking.** The overbooking strategy has the potential to allow primary care providers to see more patients and increase the expected profits. The challenge in overbooking is how to balance the costs of no-shows with the costs of overtime because overbooking may cause an excessive overtime to serve the overbooked patients. Overbooking is an accepted practice in the airline industry, hotel industry, and other services. In Several reports show how this strategy can increase profit and decrease the relative costs of no-shows [LaGanga and Lawrence, 2007].

In addition, overbooking is taken under consideration in lots of literature as a powerful strategy to mitigate negative impacts of no-shows. [Muthuraman and Lawley, 2008] present an overbooking scheduling policy for outpatient clinics in order to minimize patients' waiting time and staffs' overtime. [Kim and Giachetti, 2006], [Zeng et al., 2010], and [Samorani and LaGanga, 2011] all develop a stochastic mathematical overbooking model to maximize providers' expected profits. Lately, [Zacharias and Armony, 2016] study an overbooking model for scheduling for patients with different no-show probabilities.

[Robinson and Chen, 2010] compares traditional and open access appointment scheduling policies for a single physician. Their results indicate that open-access schedule outperforms the traditional schedule in the wide majority of cases and traditional scheduling policy will be preferred only when the no-show probability

is small (less than 5%) or the cost of patients' waiting is low. In their proposed model for the traditional scheduling, they assume that the number of patients scheduled for each day is exogenously given and there are no unscheduled walk-in patients.

In another study, [LaGanga and Lawrence, 2012] develop an analytic appointment scheduling model with overbooking that balances the benefits of increased revenues and service with the expected costs of customer waiting and provider overtime. Their results show that appropriate scheduling in healthcare clinics can significantly improve patient service and provider productivity. However, they assume that the physician only served routine patients who reserved their appointment ahead of time.

According to aforementioned studies, overbooking can significantly improve clinic performance by increasing patient access and improving clinic productivity, but it may cause complexity in scheduling and can increase patients' waiting time and providers' overtime inevitably. Moreover, overbooking is more helpful when the no-show rate is high and is not a recommended strategy in an environment with low no-show rates.

**Block-booking.** Block-booking is another strategy to reduce the negative impacts of no-shows which assign patients to different blocks, instead of slots, in order to maintain utilization of the practitioner in case of no-show. The simplest format of this strategy is *Single-block* rule which assigns all patients to one block at the beginning of the session and serve them base on FIFO system. The primitive studies used single-block scheduling as their basis, such as [Bailey, 1952] and [Lindley, 1952] knowing that it may lead to excessive waiting times for patients. In another case called *Multiple-block* groups of patients are assigned to each appointment slot. [White and Pike, 1964] and [Soriano, 1966] both use multiple-block rule for their proposed appointment systems. [Fries and Marathe, 1981] analyze a single physician variable-size multiple-block system and show how this flexibility can reduce average



waiting times and idle times. [Liu and Liu, 1998] develop a block appointment system for primary care clinic considering doctors' lateness.

[Shonick and Klein, 1977] and [Millen and Shonick, 1978] develop mathematical models and suggest to reserve enough number of patients in a block to ensure that even if some patients do not show up, the physician's time would be at all times fully utilized.

**Reserving slots for urgent patients.** As discussed in Section 2.2.3, one of the reasons for no-shows is long indirect waiting time which happens when a patient reserved an appointment in advance, long time ago. Some researchers suggest switching to advance access appointment system to eliminate delay time between request and actual appointment in order to reduce no-show rate. Literature on this arena is previously discussed in Section 2.2.2.

### 2.3. Scheduling Models and Methodologies

Studies on appointment scheduling can be classified on the basis of the solution approach and methodology into three categories which will be discussed in the following.

#### 2.3.1 Queuing Theory and Markov Chain

Queueing models can provide accurate evaluations of system performance. Moreover, Queueing theory is a useful approach in the health care setting specifically, in appointment systems which designed to reduce patients' waiting times and physician utilization [Fomundam and Herrmann, 2007].

Periodic arrival processes are fundamentally studied within the framework of queueing systems in primary care, where routine patients arrive in random intervals and the service time is random. [Bailey, 1952] introduce the application of queueing theory in health care appointment systems and shows how punctuality of physician

can help a primary care clinic to reduce patients' waiting time. [Taylor and Templeton, 1980] develop a priority queue in steady state with multiple servers and two classes of customers in order to reduce patients' waiting time in an urban ambulance service. [Brahimi and Worthington, 1991] also design a queuing model for a primary care clinic in order to reduce patients' waiting time without increasing the physicians' idle time. In another study [Vasanawala and Desser, 2005] use queuing theory to determine the percentage of open slots reserved for walk-ins or urgent patients. [Balasubramanian et al., 2007] propose a simulation optimization model to determine panel design in a primary care environment in order to maximize continuity of care and reduce patients' waiting time and staffs' overtime. [Hassin and Mendel, 2008] analyzes the scheduling appointment systems with no-shows using a queue with a single server. [Green and Savin, 2008] and [Green, 2010] develop a single-server queueing system and compares the performance metrics with a simulated appointment system. [Qu and Shi, 2011] present a Markov chain model to capture appointment scheduling in open access clinics. [Patrick, 2012] develop a Markov Decision Process model and demonstrate that a short booking window is not better than open access scheduling. [Zacharias and Armony, 2016] study the in-clinic queue in order to characterize the patients' waiting times and physicians' overtime.

### **2.3.2 Simulation Based Studies**

Simulation has the ability to model complex appointment systems and evaluate the performance of those systems with different performance metrics. Therefore, lots of researchers developed simulation models to help primary care providers to assess the efficiency and effectiveness of their appointment systems and find the best justification. In early studies, [Fetter and Thompson, 1966] and [Vissers and Wijngaard, 1979] investigate the impact of several factors on appointment system performance.

In later studies, [Ho and Lau, 1992] compare different scenarios for appointment system using simulation, and demonstrate that none of the rules are optimal in all the environments. [Klassen and Rohleder, 1996] develop a simulation model for a primary care environment and find the best scheduling rule combined with the position of open slots for urgent patients. [Liu and Liu, 1998] design a scheme, based on a simulated model that provides the optimal schedule for a given scheduling environment. [Harper and Gamlin, 2003] examine different appointment schedules to identify critical factors that influence patients' waiting times. [Giachetti et al., 2005] develop a discrete-event simulation model to analyze clinic performance and make recommendations for an open access system. Also, [Kopach et al., 2007] investigate the effects of different parameters on open access scheduling performance and show how open access can lead to significant improvements in clinic throughput. [Cayirli and Gunes, 2014] use simulation to derive solutions for the appointment scheduling problem in order to reduce patients' waiting time, physicians' idle time and overtime.

Application of simulation in capturing the tradeoff between no-shows and system performance is significant. Using simulation, [LaGanga and Lawrence, 2007] show that overbooking provides greater utility when clinics serve larger numbers of patients with high no-show rates, and service variability is low. [Huang and Zuniga, 2012] use simulation optimization to find optimal overbooking scheduling in primary care clinics with high no-show rates. [Berg et al., 2013] develop a discrete event simulation model to determine improved overbooking scheduling policies and examine the effect of no-shows on providers' utilization. If the reader is interested more in this topic, comprehensive reviews on computer simulation modeling in health care delivery can be found in [Fone et al., 2003] and [Günel and Pidd, 2010].

### 2.3.3 Stochastic Optimization

Stochastic programming is a framework for modeling optimization problems that involve uncertainty. In particular, in primary care, managers are dealing with lots of uncertainty, such as random demand, no-shows, walk-ins, random service time, tardiness and etc. Therefore, stochastic optimization is a useful method to capture the effect of these uncertainties and find the optimal solution i.e., appointment scheduling.

The literature on stochastic optimization because of its difficulties is sparse. [Denton and Gupta, 2003] formulate the appointment scheduling problem as a two-stage stochastic linear program to determine optimal appointment times. [Dobson et al., 2011] formulate a stochastic model of appointment scheduling in a primary care practice to minimize the average number of urgent patients that are not handled during normal hours and the average queue of routine patients. [Kim and Giachetti, 2006] develops a stochastic mathematical overbooking model maximize providers' expected profits. [Qu et al., 2007] present a quantitative approach to determine the optimal percentage of appointments held open in a session. [Muthuraman and Lawley, 2008] present a stochastic overbooking model for appointment scheduling in outpatient clinics to reduce patients' waiting time and staffs' overtime.

Recently, [Wang and Gupta, 2011] presents a stochastic model for appointment system using updated patients' preferences, in order to improve providers' revenues, serve more patients, and increase continuity of care. [Samorani and LaGanga, 2011] present a stochastic program to find the optimal schedule with the presence of no-show. [Luo et al., 2012] develop a framework which can be utilized in determining the optimal appointment policy to balance the trade-off between the patients' waiting times and providers' utilization. [Tsai and Teng, 2014] develop a stochastic overbooking model to enhance the service quality and increase the providers' utilization. [Mak et al., 2014] formulate a stochastic appointment scheduling problem to minimize the worst-case expected waiting time and overtime

costs. [Samorani and LaGanga, 2015] use stochastic dynamic programming to develop a dynamic appointment scheduling method which can be used in real time to schedule appointment requests.

## CHAPTER 3

### SINGLE PHYSICIAN APPOINTMENT SCHEDULING

We develop a stochastic model for the appointment planning and scheduling of a single primary care provider to optimize the social welfare, which includes the provider's profit and the cumulative waiting time of the patients. A carve-out appointment system is adopted in order to serve two different types of patients, the routine patients who have reserved their appointments in advance and the urgent ones who need to be served on the day of calling.

Moreover, we find the optimal number of slots reserved for both routine and urgent patients in one session and develop heuristics to estimate them in an accurate manner. We also derive a daily schedule for a physician in which, certain slots, among reserved slots for routine patients, are allowed to be double-booked; while the rest of the available slots are left open for same-day requests. Finally, in a special case where only one slot is left open and just one slot is double-booked, we obtain the conditions under which a certain schedule is optimal. Then, we generalize results of this special case into other cases.

In the next chapter, we extend our model to a joint panel appointment system, which consists two physicians who have separate routine patients but will serve same-day requests jointly.

#### 3.1. Scheduling Modeling

The main goal in this section is to find the optimal appointment schedule that balances the interests of the provider and patients while assisting the provider with capacity planning decisions. To this end, we develop a stochastic analytical model that maximizes a weighted sum of the performance metrics, which include the provider's revenue, the cost of lost sales, overtime cost, and patients' waiting times. The

provider's planning decisions consist the daily number of appointments reserved for urgent patients, for routine patients as well as the maximum number of slots allowed to double book. The provider's schedule specifies where these reserved slots are located. We also characterize the optimal schedules and demonstrate that an efficient heuristic procedure can be used find near optimal schedules.

The demand for the routine patients follows a probability mass function (p.m.f.)  $q_i^r$  for  $i = 0, 1, \dots, U_r$  and the demand for urgent patients follows a p.m.f.  $q_i^s$  for  $i = 0, 1, \dots, U_s$ . Further, the probability functions are independent of each other. Because the majority of the service can be finished within a fixed time interval, we assume that the service times for both types of patient are deterministic and equal. Thus, the daily appointment schedule can be divided into  $N$  intervals of equal length. Let  $i$  ( $i = 1, \dots, N$ ) represents the  $i^{th}$  time slot counted from the beginning of the day. The provider needs to determine the number of slots allocated to the routine patients,  $N_r$ . Further, let  $N_s$  denote the number of slots allocated to the same-day requests,  $N_s = N - N_r$ .

The provider receives revenue of  $p$  for each patient served, regardless of the type. We normalize the operating cost during the regular hours to be zero. If either a routine patient or a same-day request cannot be accommodated, the provider incurs a cost of lost sales ( $c_l$ ). We further assume that with probability  $\gamma$ , a routine patient would not show up for the appointment. All urgent patients, however, show up for sure.

These assumptions are made based on the data we obtained from a clinic, which has an average no-show rate of 8% for routine patients, but less than 0.1% for urgent patients. In order to mitigate the negative impacts associated with no-shows, the provider allows certain time slots to be double booked. However, all booked patients, if shown, must be served on the same day. The provider thus needs to work overtime if there are un-served patients by the end of the day. In this case, the provider incurs

an overtime cost of  $c_v$  for each patient. Moreover, since double-booking all of the slots assigned to routine patients may not be optimal, we introduce another decision variable,  $A$ , to represent the maximum number of slots that can be double booked. Thus, the maximum number of routine patients can be served is  $N_r + A$ . See Table 3.1 for a summary of notation.

**Table 3.1** Single Physician Model's Notation

|             |  |
|-------------|--|
| Parameters: |  |
| $p$ :       | Revenue per patient  |
| $c_l$ :     | Cost of lost sales per patient   |
| $c_w$ :     | Waiting cost per slot per patient  |
| $c_v$ :     | Cost of overtime per slot  |
| $\gamma$ :  | No-show rate for routine patient   |
| $N$ :       | Number of slots available during the regular session   |
| Decisions:  |  |
| $N_r$ :     | Number of slots assigned to routine patients   |
| $A$ :       | Maximum number of double-booked patients   |
| $x_i$ :     | Maximum number of patients allowed to be assigned to slot $i$ ,<br>$i = 1, \dots, N$ and $x_i = 0, 1, 2$ |
| Variables:  |  |
| $D_r$ :     | Demand for routine patients, follows $pmf$ : $q_i^r$ for $i = 0, 1, \dots, U_r$                          |
| $D_s$ :     | Demand for urgent patients, follows $pmf$ : $q_i^s$ for $i = 0, 1, \dots, U_s$                           |
| $N_s$ :     | Number of slots assigned to same-day patients, $N_s = N - N_r$   |
| $k_i$ :     | Number of routine patients that showed up, follows $pmf$ : $B(k; n, \gamma)$                             |
| $w_i$ :     | Number of patients waiting at the end of slot $i$ , $w_i = 0, 1, \dots, i$                               |

**Provider's Appointment Schedule.** Aside from determining the number of slots allocated to the routine patients ( $N_r$ ) and the maximum number of slots that can be double booked ( $A$ ), the provider also needs to specify which time slots are designated for urgent patients, which are for routine patients, and which ones can be double-booked. Let  $x_i$  denote the maximum number of patients allowed to be assigned to time slot  $i$ . If  $x_i = 0$ , time slot  $i$  is reserved for urgent patients and no routine



patients is booked for that slot. If  $x_i = 1$ , time slot  $i$  is reserved for routine patients and no double-bookings is allowed. If  $x_i = 2$ , time slot  $i$  can be double-booked if all other routine slots have been booked. Thus, the number of patients showed up at slot  $i$ ,  $k_i$  can be computed as follows:

$$k_i|\{x_i = 1\} = \begin{cases} 0 & B(0; 1, \bar{\gamma}); \\ 1 & B(1; 1, \bar{\gamma}). \end{cases} \quad \text{and} \quad k_i|\{x_i = 2\} = \begin{cases} 0 & B(0; 2, \bar{\gamma}); \\ 1 & B(1; 2, \bar{\gamma}); \\ 2 & B(2; 2, \bar{\gamma}). \end{cases} \quad (3.1)$$

where  $\bar{\gamma} = 1 - \gamma$  is the probability that a routine patient shows up. The term  $B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$  is the binomial distribution, when  $n = 0$ ,  $B(k; n, p) = 0$ . Given a pre-determined schedule,  $\{x_i\}_{i=1, \dots, N}$ , the provider can expect  $n_r$  routine patients, where

$$n_r = \sum_{i=1}^N \sum_{k_i=0}^{x_i} k_i \cdot B(k_i; x_i, \bar{\gamma}). \quad (3.2)$$

Because the provider also cares about the waiting times of the patients, let  $c_w$  be the waiting cost per time slot for each patient. Let  $w_i$  represent the number of patients waiting at the end of time  $i$ . Then, the probability of having  $j$  patients waiting at the end of slot  $i$ ,  $P(w_i = j)$ , can be derived using the following recursive function:

$$\begin{aligned} P(w_i = j) = & P(w_{i-1} = j + 1) \cdot P(k_i = 0|x_i) + P(w_{i-1} = j) \cdot P(k_i = 1|x_i) \\ & + P(w_{i-1} = j - 1) \cdot P(k_i = 2|x_i). \end{aligned} \quad (3.3)$$

The first term in Eq. (3.3) corresponds to the situation where  $j + 1$  patients are waiting at the end of slot  $i - 1$ , and no patients show up for slot  $i$ . The second term captures the case where  $j$  patients are waiting at the end of slot  $i - 1$  and one patient shows up for slot  $i$ . The last term captures the situation where  $j - 2$  patients are waiting at the end of slot  $i - 1$  and two new patients who are both booked for slot

$i$  show up. In all three cases, only one patient is served in slot  $i$  and all remaining patients would wait.

**Provider's Revenue.** For each patient served, regardless of the type, the provider obtains a revenue  $p$ . Thus, the provider's expected revenue depends on the realized demands  $(D_r, D_s)$  of both routine and urgent patients, and their allocated capacities  $(N_r, N_s, \text{ and } A)$ . Let us first define an operating function,  $\phi(\cdot)$ , as follows:

$$\phi(x, l_r, u_r, l_s, u_s) = \sum_{D_r=l_r}^{u_r} \sum_{D_s=l_s}^{u_s} (x) q_{D_s}^s q_{D_r}^r, \quad (3.4)$$

where  $l_r$  and  $u_r$  are the lower and upper bounds of the demand for routine patients, and  $l_s$  and  $u_s$  are the lower and upper bounds of the demand for same-day requests.

Given a pair of capacity allocations for the routine patients  $(N_r)$  and the urgent patient  $(N_s)$  as well as a maximum number of slots that can be double-booked  $(A)$ , the provider's expected revenue, denoted by  $\Pi(N_r, A)$ , can be computed by

$$\begin{aligned} \Pi(N_r, A) = & p \cdot \phi(D_r \bar{\gamma} + \min\{D_s, N - D_r\}, 0, N_r, 0, U_s) \\ & + p \cdot \phi(n_r + \min\{D_s, N_s\}, N_r, N_r + A, 0, U_s) \\ & + p \cdot \phi(n_r + \min\{D_s, N_s\}, N_r + A, U_r, 0, U_s). \end{aligned} \quad (3.5)$$

The first term in Equation (3.5) refers to the case where the demand for routine patients is less than the allocated capacity. There are  $N - D_r$  un-utilized slots, which can be assigned to the same-day requests. The second term corresponds to the situation where all demand for routine patients can be accommodated with up to  $A$  slots double-booked. Finally, the last term refers to the case where all capacity  $(N_r + A)$  for routine patients is used up and the remaining routine patients are not served.

**Provider's Lost Sales.** Once the capacity planning decisions are made, the provider can only serve up to  $N_r + A$  routine patients and  $N_s$  urgent patients in a day. If any of the demands exceeds the capacity, the provider incurs a lost sale cost. Let

$L(N_r, A)$  denote the provider's expected cost from lost sales, and it can be computed by

$$\begin{aligned}
L(N_r, A) = & c_l \cdot \phi(D_s - (N - D_r), 0, N_r, N - D_r, U_s) \\
& + c_l \cdot \phi(D_s - N_s, N_r, N_r + A, N_s, U_s) \\
& + c_l \cdot \phi(D_r - (N_r + A), N_r + A, U_r, 0, U_s) \\
& + c_l \cdot \phi(D_r + D_s - (N + A), N_r + A, U_r, N_s, U_s).
\end{aligned} \tag{3.6}$$

The first two terms in Equation (3.6) corresponds to the case where there are more urgent patients than available open slots ( $N - D_r$  if the number of routine patients is less than capacity or  $N_s$  otherwise), and thus the urgent patients must be referred to other providers. The third term refers to the scenario where the demand for routine patients is over the maximum capacity ( $N_r + A$ ) with double-booking, and thus some of the routine patients would not be served. The last term refers to the situation where both types of demand exceed their respective capacities.

**Provider's Overtime Cost.** While adopting the double-booking strategy can mitigate the negative effects of no-shows, it may require the provider and other clinic staff to work overtime. Consider the following situation, slot  $i$  is double-booked. When both patients show up, only one can be served during the scheduled time interval. Because clinics often operate in a first come first serve basis, the other patient will be seeing the provider in slot  $i + 1$ . If any patient who is booked for slot  $i + 1$  shows up, that patient will also be delayed by at least one slot. These delays are easily concatenated and may result in some un-served at the end of the day, i.e. slot  $N$ . The provider's expected overtime cost thus depends on the appointment schedule  $\{x_i\}_{i=1, \dots, N}$ . Let  $O(N_r, A, \{x_i\})$  denote the provider's overtime cost, and it can be calculated as follows:

$$O(N_r, A, \{x_i\}) = c_v \cdot \phi\left(\sum_{j=0}^{\min\{D_r - N_r, A\}} j \cdot P(w_N = j), N_r, U_r, 0, U_s\right). \tag{3.7}$$

where  $P(w_N = j)$  in Equation (3.7) is derived using Equation (3.3) recursively. Note that delays at the end of the day only occurs when there are double-booked patients, thus we only need to consider the scenario where the demand for routine patients exceeds its allocated capacity, and thus the number of patients at the end of slot  $N$  will not exceed  $\min\{D_r - N_r, A\}$ .

**Patients' Waiting Cost.** Similar to the provider's overtime cost, the appointment schedule  $\{x_i\}_{i=1,\dots,N}$  with double-booking strategy highly affects the waiting times for patients. Let  $W(N_r, A, \{x_i\})$  denote the aggregated patient overtime cost, and it equals to

$$W(N_r, A, \{x_i\}) = c_w \cdot \phi \left( \sum_{i=1}^N \sum_{j=0}^{\min\{i, D_r - N_r, A\}} j \cdot P(w_i = j), N_r, U_r, 0, U_s \right). \quad (3.8)$$

In Equation (3.8), the number of patients waiting ( $j$ ) at slot  $i$  should exceed neither  $i$  nor the total number of double-booked patients  $\min\{D_r - N_r, A\}$ .

**Provider's Utility.** The provider's utility, denoted by  $U(N_r, A, \{x_i\})$ , can be derived by combining the revenue and cost functions:

$$U(N_r, A, \{x_i\}) = \Pi(N_r, A) - L(N_r, A) - O(N_r, A, \{x_i\}) - W(N_r, A, \{x_i\}). \quad (3.9)$$

Let  $(N_r^*, A^*, \{x_i\}^*)$  be the joint optimal decisions on the number of slots allocated for routine patients, the maximum number of slots allowed to be double-booked and the appointment schedule, respectively that maximize the provider's utility function, that is,

$$(N_r^*, A^*, \{x_i\}^*) = \max_{N_r, A, \{x_i\}} U(N_r, A, \{x_i\}). \quad (3.10)$$

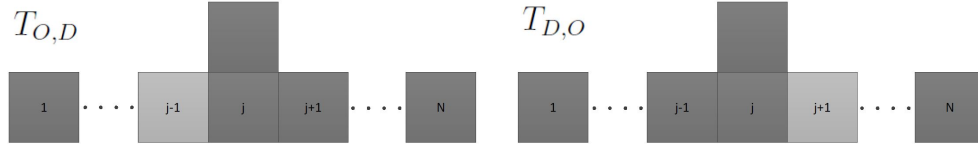
### 3.2. Analytical Solutions

In this section, we present the general form of the optimal solution under different scenarios. We start from the simplest case where it is optimal to double book exactly one patient and leave only one open slot for a same-day request. Let  $T_{O,D}$  be the

schedule in which, the open slot ( $O$ ) is placed before the double-booked slot ( $D$ ). Also,  $T_{D,O}$  be the schedule in which, the open slot is placed after the double-booked slot. Therefore, all possible schedules can be divided into two general scenarios:

- $T_{O,D}$ : the open slot is placed before the double-booked slot.
- $T_{D,O}$ : the open slot is placed after the double-booked slot.

In order to make a better visualization, we can exhibit each schedule as shown in Figure 3.1. Each square represents one slot, which is assumed to have the same length as other slots (see Section 3.1 for further information). The light gray squares present open slots and dark gray ones represent slots which allocated to the routine patients. We also demonstrate double-booked slots with two dark gray squares on top of each other.



**Figure 3.1** General scenarios for optimal schedule.

Both of these schedules are too general to argue. Therefore, first we focus on  $T_{O,D}$  to narrow down all possible forms of it, and then we discuss  $T_{D,O}$ .

**Schedule  $T_{O,D}$ .** In this schedule, the open slot can be anywhere before the double-booked slot since it does not affect the expected waiting and overtime costs. Consequently, The probability of having one patient waiting at the end of slot  $i$ ,  $P(w_i = 1)$  are:

$$\begin{aligned}
 P(w_{i < j} = 1) &= 0 \\
 P(w_j = 1) &= (1 - \gamma)^2 \\
 P(w_{i > j} = 1) &= (1 - \gamma)^{i-j+2}
 \end{aligned}
 \tag{3.11}$$

Note that, the probability of having more than one patient waiting is always zero, since we assume that the schedule has only one double-booked slot. As a result

the expected cost of waiting in this schedule is  $\sum_{i=j}^N (1 - \gamma)^{i-j+2}$  and also the expected overtime cost is  $c_v(1 - \gamma)^{N-j+2}$ . Although, the double-booked slot can be anywhere, in this schedule, we can concisely express a condition in which a certain form of this schedule is optimal.

**Proposition 1.** *Among all schedules which contain an open slot prior to the double-booked slot, the schedule in which the second slot is double-booked is better than the schedule with a double-booked slot at the end if and only if  $c_w \bar{\gamma} \gamma^{-1} \leq c_v$ .*

The proof of this proposition and all subsequent propositions can be found in Appendix A. This proposition drastically reduces the number of schedules, which are derived from  $T_{O,D}$ , into just two schedules. Lets  $T_{O,D}^{1,2}$  be the schedule which has a double-booked slot at the second slot and  $T_{O,D}^{1,N}$  the one with double-booked slot at the end which are demonstrated in Figure 3.2.



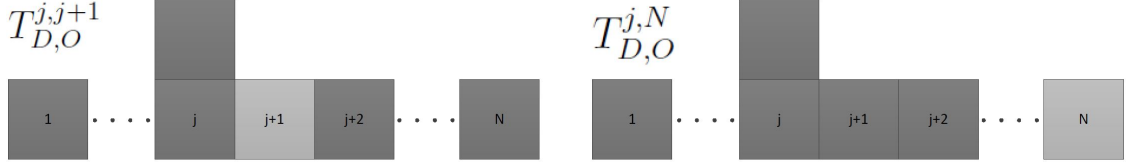
**Figure 3.2** Optimal schedules with  $T_{O,D}$ .

**Schedule  $T_{D,O}$ .** In this schedule the open slot is after the double-booked slot and consequently, its position affect the expected waiting time and overtime. Similarly, we can present a condition in which a particular form of this schedule is optimal.

**Proposition 2.** *It is optimal to place the open slot right after the double-booked slot if and only if  $q_0^s \geq \gamma$ ; otherwise, the open slot should be at the end of schedule (far from the double-booked slot).*

This proposition divides  $T_{D,O}$  into two general forms. Let  $T_{D,O}^{j,j+1}$  denote the schedule in which the open slot is right after the double-booked slot and  $T_{D,O}^{j,N}$  the

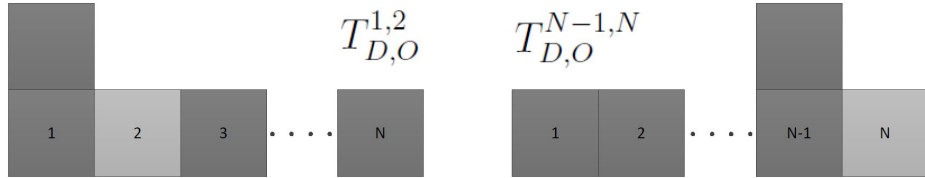
schedule in which the open slot is at the end. However, these two schedules (Shown in Figure 3.3) are still general cases and need more exploration.



**Figure 3.3** Optimal schedules with  $T_{D,O}$ .

**Proposition 3.** *The schedule in which the first slot is double-booked and followed by an open slot is better than the schedule with double-booked at slot  $N - 1$  and an open slot at the end if and only if  $c_w \bar{\gamma} \gamma^{-1} \leq c_v$ .*

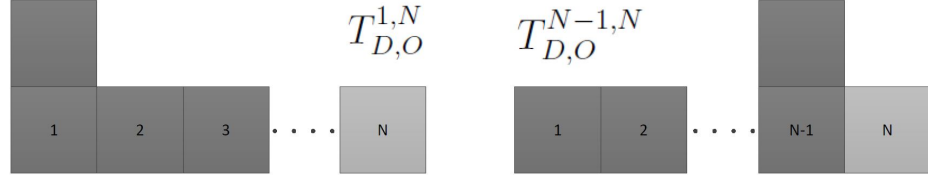
Proposition 3 reduces number of schedules, which are derived from  $T_{D,O}^{j,j+1}$ , into just two schedules. Lets  $T_{D,O}^{1,2}$  be the schedule which has a double-booked slot at the beginning followed by an open slot and  $T_{D,O}^{N-1,N}$  denote the schedule which has a double-booked slot at slot  $N - 1$  and an open slot at the end. Both of these schedules are demonstrated in Figure 3.4.



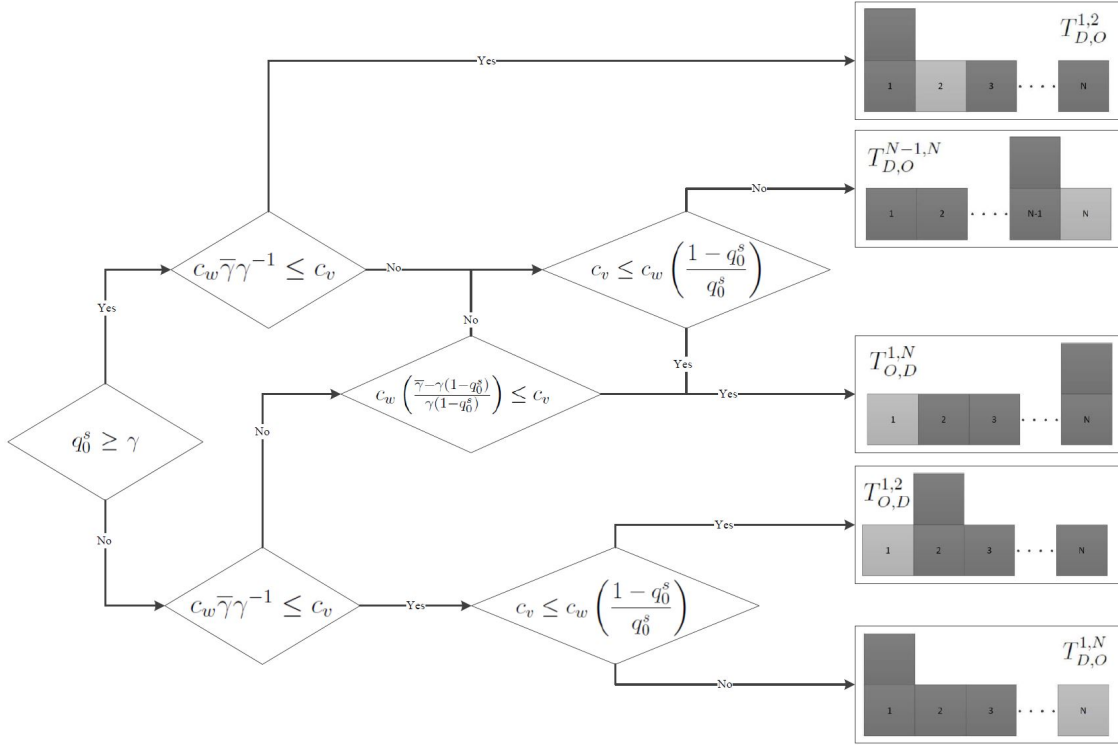
**Figure 3.4** Optimal schedules with  $T_{D,O}^{j,j+1}$ .

**Proposition 4.** *The schedule in which the first slot is double-booked and has an open slot at the end is better than the schedule with double-booked at slot  $N - 1$  and an open slot at the end if and only if  $c_w \left( \frac{\bar{\gamma} - \gamma(1 - q_0^s)}{\gamma(1 - q_0^s)} \right) \leq c_v$ .*

Proposition 4 reduces number of schedules, which are derived from  $T_{D,O}^{j,N}$ , into two distinct schedules. Lets  $T_{D,O}^{1,N}$  be the schedule which has a double-booked slot at



**Figure 3.5** Optimal schedules with  $T_{D,O}^{j,N}$



**Figure 3.6** Flowchart for optimal schedule selection.

the beginning and an open slot at the end of the schedule, which is demonstrated in Figure 3.5.

Figure 3.6 summarizes the propositions stated above. The most dominant factor in determining the position of the double-booked slot in the optimal schedule is the ratio of the overtime cost and the waiting cost. If the waiting cost is relatively small ( $\frac{c_w}{c_v} \leq \gamma \bar{\gamma}^{-1}$ ), then the double-booked slot should be placed at the beginning of the session. The advantage of such as schedule is that any slot opened up due to a no-show after the double-booked slot can be used to serve the patient that is double-booked. As a result, the chance that there is a need for overtime is minimized. Conversely, if



the waiting cost is relatively high ( $\frac{c_w}{c_v} \geq q_0^s \bar{q}_0^{s-1}$ ), then the double-booked slot should be pushed toward the end of the session to shorten the overall waiting time.

The relative position of the open slot, however, is being driven primarily by the relationship of having zero same-day requests ( $q_0^s$ ) and the probability of no-show of the routine patients ( $\gamma$ ). If  $q_0^s \geq \gamma$ , then the open slot should be placed right after the double-booked slot. This allows the service provider to use any of the un-utilized slot due to the lack of same-day requests. On the other hand,  $q_0^s \leq \gamma$ , then the position of the open slot depends on the ratio of the waiting cost and overtime cost. If the ratio is lower than  $q_0^s \bar{q}_0^{s-1}$ , then the open slot should be placed at the end of the session because the overall waiting time is small. Otherwise, it is optimal to place the open slot anywhere prior to the double-booked slot.

### 3.3. Computational Solutions

Due to the combinatorial characterization of the no-shows in each scheduled or double booked slots, the traditional approach of solving newsvendor problems does not apply. We thus investigate two solution methods: complete enumeration (in Section 3.3.1) and heuristic schedules (in Section 3.3.2).

We adopted a two steps procedure to find the optimal schedule through complete enumeration algorithm. In the first step, for each set of capacity planning decisions ( $N_r$ ,  $N_s$  and  $A$ ), we select the best schedule which generates the highest expected utility. Then in the second step, we select the highest expected utility,  $U^*(N_r^*, N_s^*, A^*)$  amongst the best schedules obtained from step one.

At the end of this chapter in Section 3.3.2, we introduce three heuristics, which can instantly estimate the optimal capacity planning decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ) and as a result reduces the time of complete enumeration computations drastically.

### 3.3.1 Characterizations of Optimal Schedules

We first use the complete enumeration algorithm to find the optimal schedules along with the optimal capacity planning decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ). Because the maximum number of patients allowed for each slot  $i$ ,  $x_i$ , can take three integer values: 0, 1 and 2, the runtime of this algorithm is exponential ( $O(3^N)$ ). Therefore, we only use this method for smaller capacity ( $N \leq 15$ ). Results for larger capacity are derived using heuristics and presented afterward.

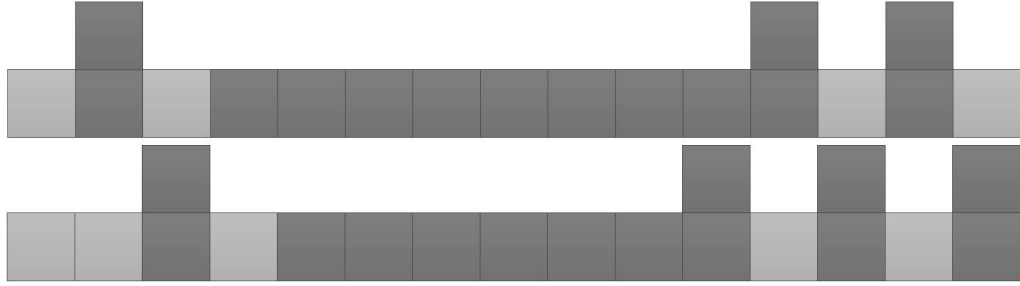
**Case 1: low no-show rate, low waiting cost ( $\gamma \leq q_0^s$  and  $c_w \bar{\gamma} \gamma^{-1} \leq c_v$ ).** The optimal capacity decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ) and the optimal schedules generated using the following parameters  $N = 15$ ,  $\gamma = 8\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 3$  and  $D_r \sim U[0, 14]$  are demonstrated in Table 3.2 and Figure 3.7, respectively. We use two different examples with different distributions ( $D_s \sim U[0, 4]$  and  $D_s \sim U[0, 5]$ ) for same-day requests to examine the general pattern in the optimal schedule.

**Table 3.2** Optimal Capacity Planning Decisions in Case 1

| $D_s$     | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|-----------|---------|---------|-------|-------|
| $U[0, 4]$ | 11      | 4       | 3     | 414.8 |
| $U[0, 5]$ | 10      | 5       | 4     | 434.7 |

Table 3.2 shows that as the demand for the same-day requests increases, the optimal number of open slots ( $N_s^*$ ) and optimal double-booked slots ( $A^*$ ) increase. This indicates that the optimal schedules tends to accommodate urgent patients by assigning as much open slots as needed.

Recall that in the case where only one double booked slot and one open slot are assigned, it is optimal to place a double-booked slot at the beginning of the session followed by an open slot and a series of single slots when the waiting cost is low. The optimal schedule ( $T_{D,O}^{1,2}$ ) allows the physician to serve one of the double-booked



**Figure 3.7** Optimal schedule in Case 1 (Top:  $D_s \sim U[0, 4]$ , Bottom:  $D_s \sim U[0, 5]$ ).

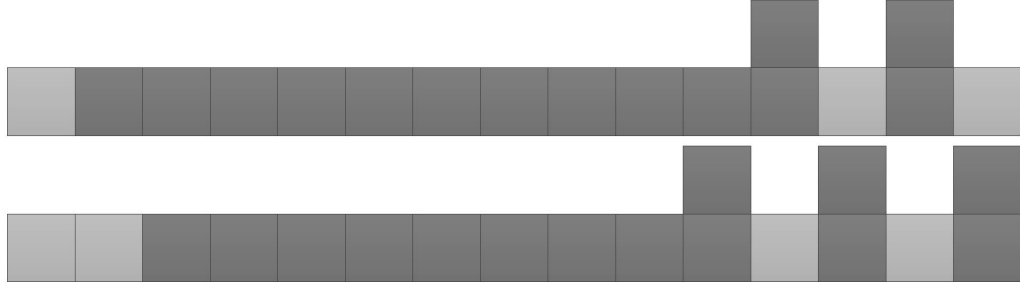
patients if the number of same-day requests is low. If the number of same-day requests is high, however, the physician can still serve the patient in a single booked slot if any of the reserved patients is a no-show. The optimal schedule, in this case, exhibits a similar pattern with a block appended at the beginning and at the end of the session, respectively. Because the probability of one or more same-day requests is relatively high (75% and higher), open slots are placed at the beginning of the session to increase the physician's utilization. Several pairs of a double-booked slot followed by an open slot are allocated at the end of the session. In the case of low same-day demands, the open slots can be used to serve double-booked patients. Otherwise, these patients will be served during overtime.

**Case 2: low no-show rate, moderate waiting cost** ( $\gamma \leq q_0^s$  and  $c_w \bar{\gamma} \gamma^{-1} \geq c_v \geq c_w \bar{q}_0^s q_0^{s-1}$ ). The optimal capacity decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ) and the optimal schedules generated using the following parameters  $N = 15$ ,  $\gamma = 8\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 6$  and  $D_r \sim U[0, 14]$  are demonstrated in in Table 3.3 and Figure 3.8. We use two different examples with different distributions ( $D_s \sim U[0, 3]$  and  $D_s \sim U[0, 4]$ ) for same-day requests to examine a general pattern in the optimal schedule.

Similar to the case 1, the optimal schedule inclines to accommodate urgent patients as much as needed by increasing the number of open slots. As a result, the number of slots assigned to routine patients decreases and the number of double-booked slots increases.

**Table 3.3** Optimal Capacity Planning Decisions in Case 2

| $D_s$     | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|-----------|---------|---------|-------|-------|
| $U[0, 3]$ | 12      | 3       | 2     | 392.6 |
| $U[0, 4]$ | 11      | 4       | 3     | 412.5 |



**Figure 3.8** Optimal schedule in Case 2 (Top:  $D_s \sim U[0, 3]$ , Bottom:  $D_s \sim U[0, 4]$ ).

The optimal schedule, in this Case, is similar to that of Case 1 at first glance. However, as one can see the optimal schedule in Case 1 assigns a double-booked slot at the beginning of the session while all of the double-booked slots are pushed toward the end of the session in the optimal schedule in Case 2. The main differences from the high waiting cost ( $c_w$ ). Therefore, the optimal schedule tries to minimize the overall expected waiting times by placing the double-booked slots at the end of the session.

**Case 3: low no-show rate, high waiting cost** ( $\gamma \leq q_0^s$  and  $c_w \bar{q}_0^s q_0^{s-1} \geq c_v$ ).

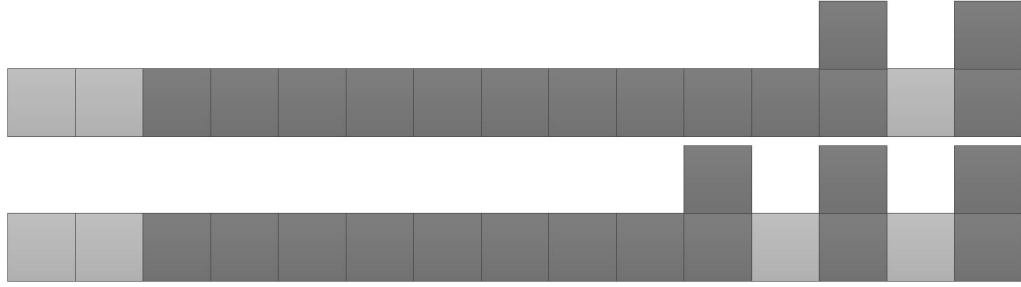
The optimal capacity decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ) and the optimal schedules generated using the following parameters  $N = 15$ ,  $\gamma = 8\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 13$  and  $D_r \sim U[0, 14]$  are demonstrated in Table 3.4 and Figure 3.9. We use two different examples with different distributions ( $D_s \sim U[0, 4]$  and  $D_s \sim U[0, 6]$ ) for same-day requests to examine a general pattern for the optimal schedule.

Comparing to the optimal capacity planning decisions of Case 2, the optimal schedule allocates fewer open slots to urgent patients and reserve more capacity for

**Table 3.4** Optimal Capacity Planning Decisions in Case 3

| $D_s$     | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|-----------|---------|---------|-------|-------|
| $U[0, 4]$ | 12      | 3       | 2     | 410.2 |
| $U[0, 6]$ | 11      | 4       | 3     | 441.1 |

routine patients due to the combination of a high waiting cost ( $c_w \bar{q}_0^s q_0^{s-1} \geq c_v$ ) and a low no-show rate of the routine patients. This allows the provider to minimize the required double-booked slots and thus avoid prolonged waiting times. On the other hand, some same-day requests will be turned away.



**Figure 3.9** Optimal schedule in Case 3 (Top:  $D_s \sim U[0, 4]$ , Bottom:  $D_s \sim U[0, 6]$ ).

We previously show that the optimal schedule ( $T_{O,D}^{1,N}$ ) in the special case, where only one double booked slot and one open slot are assigned, places an open slot at the beginning and assign the double-booked slot at the end of the session. When the waiting cost is high, this arrangement minimizes the overall expected waiting time. The optimal schedule, in this case, exhibits a similar pattern accompanied by several pairs of a double-booked slot followed by an open slot at the end of the session. The rationale of the last block is that, in the case of low same-day demands, the open slots can be used to serve double-booked patients. Otherwise, these patients will be served during overtime.

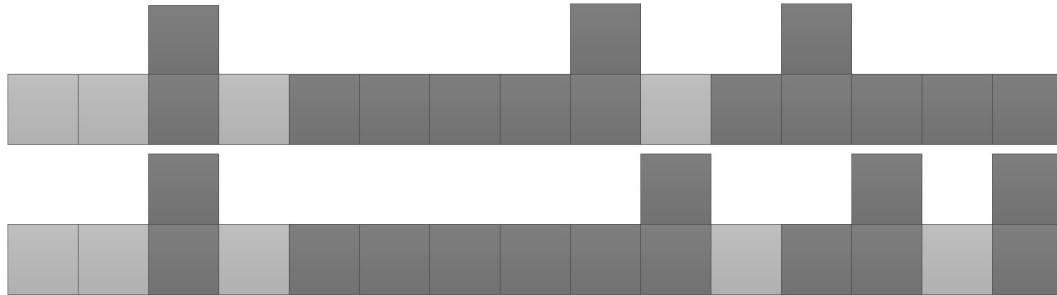
**Case 4: high no-show rate, moderate waiting cost** ( $\gamma \geq q_0^s$  and  $c_w \bar{\gamma} \gamma^{-1} \leq c_v \leq c_w \bar{q}_0^s q_0^{s-1}$ ). The optimal capacity decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ) and the optimal schedules

generated using the following parameters  $N = 15$ ,  $\gamma = 30\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 13$  and  $D_r \sim U[0, 14]$  and are demonstrated in Table 3.5 and Figure 3.10. We use two different examples with different distributions ( $D_s \sim U[0, 4]$  and  $D_s \sim U[0, 5]$ ) for same-day requests to examine a general pattern for the optimal schedule.

**Table 3.5** Optimal Capacity Planning Decisions in Case 4

| $D_s$     | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|-----------|---------|---------|-------|-------|
| $U[0, 4]$ | 11      | 4       | 3     | 338.7 |
| $U[0, 5]$ | 10      | 5       | 4     | 358.7 |

Due to high no-show rate amongst routine patients, the optimal schedule inclines to accommodate urgent patients as much as needed by increasing the number of open slots.



**Figure 3.10** Optimal schedule in Case 4 (Top:  $D_s \sim U[0, 4]$ , Bottom:  $D_s \sim U[0, 5]$ ).

Recall that in the case where only one double booked slot and one open slot are assigned, it is optimal to place an open slot at the beginning followed by a double-booked slot when the waiting cost is moderate and the no-show rate is high. The optimal schedule  $(T_{O,D}^{1,2})$  allows the physician to serve one of the double-booked patients in a single booked slot if any of the reserved patients is a no-show. The optimal schedule in this case, though looks different, follows a similar logic from the special case. Because the probability of having more than one same-day request is high, the schedule allocates a couple of open slots at the beginning of the session

to maximize the physician’s utilization. The second block consists a double booked slot followed by an open slot and a series of single booked slots to serve one of the double-booked patients in the highly likely case of no shows from routine patients. The third block has a similar structure as the second block except the number of single booked slots is much lower, as the overall utility is higher to serve these patients during overtime rather than having double-booked patients wait for a long time.

**Case 5: high no-show rate, low waiting cost** ( $\gamma \geq q_0^s$  and  $c_w \bar{q}_0^s q_0^{s-1} \leq c_v$ ).

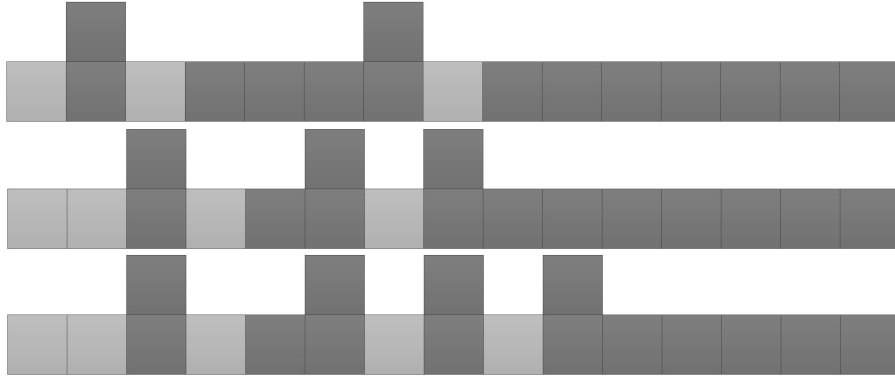
The optimal capacity decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ) and the optimal schedules generated using the following parameters  $N = 15$ ,  $\gamma = 30\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 3$  and  $D_r \sim U[0, 14]$  are demonstrated in Table 3.6 and Figure 3.11. We use three different examples with different distributions ( $D_s \sim U[0, 3]$ ,  $D_s \sim U[0, 4]$  and  $D_s \sim U[0, 5]$ ) for same-day requests to examine a general pattern for the optimal schedule.

**Table 3.6** Optimal Capacity Planning Decisions in Case 5

| $D_s$     | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|-----------|---------|---------|-------|-------|
| $U[0, 3]$ | 12      | 3       | 2     | 319.2 |
| $U[0, 4]$ | 11      | 4       | 3     | 343.2 |
| $U[0, 5]$ | 10      | 5       | 4     | 366.5 |

As one can see, the optimal schedule tends to accommodate urgent patients by assigning as much as open slots needed. Due to the high no-show rate of the routine patients the optimal schedule assigns fewer slots to routine patients and uses double-booking if routine patient’s demand rises.

The optimal schedule, in this case, carries the same structure as the optimal schedule in Case 4. However, due to a low waiting cost, it assigns double-booked slots earlier in the session and reserves a series of single booked slots for routine patients at the end of the session. The single booked slots can be used to serve the double-booked patients in the highly likely case of no-shows from routine patients.



**Figure 3.11** Optimal schedule in Case 5 (Top:  $D_s \sim U[0, 3]$ , Middle:  $D_s \sim U[0, 4]$  Bottom:  $D_s \sim U[0, 5]$ ).

### 3.3.2 Heuristic Schedules

We used a complete enumeration algorithm to obtain the optimal schedule in Section 3.3.1. Since this method is an exhaustive search, it only can be applicable in small scales. Therefore, in this section, we develop three heuristics to reduce the search space for the optimal schedule considerably. To this end, we evaluate the performance of these heuristics in the estimation of optimal capacity decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ). Hence, once the optimal capacity decisions are estimated, we can apply step two of our presented complete enumeration algorithm to find the optimal schedule respectively. Therefore, the required time for the exhaustive search, given optimal capacity decisions, decreases drastically and we are able to obtain the near-optimal schedule in large scales.

The *back loading* schedule (H-SCH 1) allows up to  $N_s$  urgent patients to be taken care of first before any of the routine patients show up. If the realized demand for urgent patients is less than  $N_s$ , however, the physician will be idle during the time allocated for urgent patients. Moreover, if the demand for routine patients is higher than  $N_r$ , double-booking is scheduled toward the end of the day. It is thus more likely that the physicians would work overtime. The optimal schedule using this heuristic for the Case 4 ( $D_s \sim U[0, 4]$ ) is shown in Figure 3.12.





**Figure 3.12** Optimal schedule using back loading heuristic for Case 4.

The *front loading* schedule (H-SCH 2) allocates all appointment slots for routine patients ( $N_r$ ) at the beginning of the day and those for urgent patients ( $N_s^*$ ) toward the end of the day. Double-booking is only exercised if the demand for routine patients is higher than  $N_r$ . In this case, up to  $A$  slots can be double booked at the beginning of the day. This schedule accommodates urgent patients at a later time of the day. Any un-used open slots can then be used to serve overflow routine patients. The optimal schedule using this heuristic for the Case 4 ( $D_s \sim U[0, 4]$ ) is shown in Figure 3.13.



**Figure 3.13** Optimal schedule using front loading heuristic for Case 4.

The *evenly spaced* schedule (H-SCH 3) starts with alternating between an open slot for the urgent patients and a slot for the routine patients. This pattern continuous until all slots for urgent patients ( $N_s$ ) have been allocated. All remaining slots, if applicable, are then assigned to routine patients. Up to  $A$  slots can be double booked. Pairing a double booked slot with an open slot helps to mitigate the no-shows and in the case where same-day requests are low, the open slot can be used to reduce patient waiting time if both patients booked in the same slot show up. The optimal schedule using this heuristic for the Case 4 ( $D_s \sim U[0, 4]$ ) is shown in Figure 3.14.

Tables 3.7 and 3.8 compare the three heuristic schedules against the optimal schedule generated by complete enumeration under three different values of patients' waiting cost ( $c_w$ ) and low no-show rate ( $\gamma = 8\%$ ).



**Figure 3.14** Optimal schedule using evenly spaced heuristic for Case 4.

**Table 3.7** Complete Enumeration ( $\gamma = 8\%$ )

| $c_w$ | Optimal Schedule |         |       |       |          |
|-------|------------------|---------|-------|-------|----------|
|       | $N_r^*$          | $N_s^*$ | $A^*$ | $U^*$ | Schedule |
| 3     | 11               | 4       | 3     | 414.8 |          |
| 6     | 11               | 4       | 3     | 412.5 |          |
| 13    | 12               | 3       | 2     | 410.2 |          |

**Table 3.8** Performance of Heuristics ( $\gamma = 8\%$ )

| $c_w$ | H-SCH1  |         |       |       |                         | H-SCH2  |         |       |       |                         | H-SCH3  |         |       |       |                         |
|-------|---------|---------|-------|-------|-------------------------|---------|---------|-------|-------|-------------------------|---------|---------|-------|-------|-------------------------|
|       | $N_r^*$ | $N_s^*$ | $A^*$ | $U^1$ | $\frac{U^* - U^1}{U^*}$ | $N_r^*$ | $N_s^*$ | $A^*$ | $U^2$ | $\frac{U^* - U^2}{U^*}$ | $N_r^*$ | $N_s^*$ | $A^*$ | $U^3$ | $\frac{U^* - U^3}{U^*}$ |
| 3     | 13      | 2       | 1     | 410.0 | 1.17%                   | 12      | 3       | 2     | 411.3 | 0.85%                   | 11      | 4       | 3     | 414.4 | 0.10%                   |
| 6     | 13      | 2       | 1     | 409.8 | 0.66%                   | 13      | 2       | 1     | 409.1 | 0.82%                   | 12      | 3       | 2     | 410.7 | 0.45%                   |
| 13    | 13      | 2       | 1     | 409.4 | 0.19%                   | 14      | 1       | 0     | 408.7 | 0.37%                   | 14      | 1       | 0     | 408.7 | 0.37%                   |

Tables 3.9 and 3.10 compare the three heuristic schedules against the optimal schedule generated by complete enumeration under three different values of patients' waiting cost ( $c_w$ ) and high no-show rate ( $\gamma = 30\%$ ). As one can see, in the case of high no-show rate, H-SCH 3 outperforms the other two heuristics. Moreover, it allocates the same amount of slots to urgent patients ( $N_s$ ) and routine patients ( $N_r$ ) as well as the maximum number of slots allowed to be double-booked ( $A$ ). As a result, the expected profit is close to optimal (less than 0.5% error).

Seemingly, the presented heuristics estimate the optimal capacity decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ) accurately in most of the cases. However, the results demonstrated that the H-SCH3 outperforms the other two heuristics when waiting cost ( $c_w$ ) is low. On the other hand, H-SCH1 perform best when the waiting cost is high. Results also

**Table 3.9** Complete Enumeration ( $\gamma = 30\%$ )

| $c_w$ | Optimal Schedule |         |       |       |          |
|-------|------------------|---------|-------|-------|----------|
|       | $N_r^*$          | $N_s^*$ | $A^*$ | $U^*$ | Schedule |
| 3     | 11               | 4       | 3     | 343.2 |          |
| 6     | 11               | 4       | 3     | 341.8 |          |
| 13    | 11               | 4       | 3     | 338.7 |          |

**Table 3.10** Performance of Heuristics ( $\gamma = 30\%$ )

| $c_w$ | H-SCH1  |         |       |       |                         | H-SCH2  |         |       |       |                         | H-SCH3  |         |       |       |                         |
|-------|---------|---------|-------|-------|-------------------------|---------|---------|-------|-------|-------------------------|---------|---------|-------|-------|-------------------------|
|       | $N_r^*$ | $N_s^*$ | $A^*$ | $U^1$ | $\frac{U^* - U^1}{U^*}$ | $N_r^*$ | $N_s^*$ | $A^*$ | $U^2$ | $\frac{U^* - U^2}{U^*}$ | $N_r^*$ | $N_s^*$ | $A^*$ | $U^3$ | $\frac{U^* - U^3}{U^*}$ |
| 3     | 11      | 4       | 3     | 336.8 | 1.87%                   | 11      | 4       | 3     | 342.0 | 0.37%                   | 11      | 4       | 3     | 343.0 | 0.07%                   |
| 6     | 12      | 3       | 2     | 336.3 | 1.62%                   | 11      | 4       | 3     | 339.2 | 0.77%                   | 11      | 4       | 3     | 341.2 | 0.17%                   |
| 13    | 12      | 3       | 2     | 335.2 | 1.03%                   | 12      | 3       | 2     | 335.7 | 0.88%                   | 11      | 4       | 3     | 337.1 | 0.48%                   |

indicate that the H-SCH2 only outperforms the other heuristics when the waiting cost is negligible ( $c_w \approx 0$ ).

Next, we evaluate the performance of the three heuristics in two dimensions: the waiting cost and the no-show rate. We use the following common parameters for all cases: the following parameters capacity:  $N = 15$ , revenue:  $p = 50$ , cost of loss sales:  $c_l = 50$ , cost of overtime:  $c_v = 45$  and the demand for routine patients follows  $D_r: U \sim [5, 18]$ . We also uses  $D_s: U \sim [0, 5]$  as the same-day requests distribution. We also consider three different patients' waiting costs:  $c_w^0 = 0$ ,  $c_w^L = 3$ , and  $c_w^H = 6$ , as well as two no show rates:  $\gamma^L = 8\%$  and  $\gamma^H = 30\%$ .

As shown in Table 3.11, when the physician has no regard of the patient waiting time ( $c_w = 0$ ), the *front loading* schedule (H-SCH 2) always generates the same utility as the optimal schedule does. Because this schedule tends to double-book at the beginning of the day, it minimizes the chance of the physician being idle, and as a result, more routine patients are served. Overflow patients can be cared for during open slots later of the day if the same-day requests are low or during overtime. This

schedule may result in really long delays for some of the routine patients that are double-booked, and as a result, it only performs well when patient waiting time is not considered.

The *back loading* schedule (H-SCH 1) performs the best when the patient waiting cost is high, the no-show rate is low. When compared to the optimal schedule, the *back loading* schedule tends to assign more slots to routine patients while allowing fewer slots to be double-booked, and consequently, the physician is able to better accommodate high demand. Moreover, because only the slots toward the end of the day are allowed to be double-booked, patient waiting time is reduced and overflow patients from the double-booked slots can be served during overtime.

The *evenly spaced* schedule (H-SCH 3) takes the lead in under two scenarios: 1) when the patients' waiting cost is low, or 2) when the patient waiting cost is high and the no-show rate is high. Pairing an open slot for urgent patient with a double-booked slot for routine patient is a key advantage of this schedule as it helps the physician mitigate the negative impact of both no-shows and low same-day requests.

**Table 3.11** Heuristics Efficiency Comparison

| Cost of waiting | Low no-show rate | High no-show rate |
|-----------------|------------------|-------------------|
| Zero            | H-SCH2           | H-SCH2            |
| Low             | H-SCH3           | H-SCH3            |
| High            | H-SCH1           | H-SCH3            |

### 3.4. Managerial Insight

The implications of the results from our analysis in Sections 3.2 and 3.3 are summarized in the following.

In the case where it is optimal to double book exactly one patient and leave only one open slot for a same-day request, the optimal schedule highly depends on

the no-show rate ( $\gamma$ ), the probability of having no same-day request ( $q_0^s$ ) and the ratio of the waiting cost to the overtime cost ( $\frac{c_w}{c_v}$ ). Our analysis shows that if the ratio of the waiting cost to the overtime cost is small, then the double-booked slot should be placed at the beginning of the session to increase physician's utilization. On the other hand, if the ratio is large, then the double-booked slot should be placed at the end of the session to shorten the overall waiting time.

Moreover, if receiving zero same-day request is more likely than a no-show, then the open slot should be placed right after the double-booked slot. This allows the physician to use the un-utilized open slot due to the lack of same-day requests. On the other hand, if the no-show rate is higher than having zero same-day request and the ratio of the waiting cost to the overtime cost is small, then the open slot should be placed at the end of the session because the overall waiting time is small. Otherwise, it is optimal to place the open slot anywhere prior to the double-booked slot.

Our numerical results indicate that in the case of low waiting cost, the provider should accommodate urgent patients by assigning as many open slots as needed and allows more slots to be double-booked for routine patients to increase the overall utilization. On the other hand, in the case of high waiting cost combined with low no-show rate, the provider should allocate fewer open slots to urgent patients and reserve more capacity for routine patients to minimize the required double-booked slots and thus avoid prolonged waiting times.

Furthermore, if the probability of having one or more same-day requests is relatively high, then a proportion of open slots should be placed at the beginning of the session to increase the physician's utilization.

We also observe that the combination of a double-booked slot followed by an open slot at the end of the session can be very effective since it allows the physician to serve double-booked patients in the case of low same-day demands. The same

reasoning holds, in the case of high no-show rate where a series of single-booked slots acts like an open slot and can be used to serve double-booked patients.

Finally, our proposed heuristics schedules are fast and efficient in order to estimate optimal capacity planning decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ), thus one can perform the complete enumeration algorithm based on these estimated decisions and largely reduces the time needed to find the optimal schedule. When the patient waiting time is negligible, the *front loading* schedule (H-SCH 2) is the optimal schedule. The *back loading* schedule (H-SCH 1) performs the best when the patients' waiting cost is high, the no-show rate is low and the *evenly spaced* schedule (H-SCH 3) takes the lead in all other cases.

## CHAPTER 4

### JOINT PANEL SCHEDULING

Most literature has focused on single-server systems for appointment scheduling in primary care. However, in practices in most of the clinics, there are several primary care physicians (PCP) who may have their own list of patients [Gupta and Denton, 2007]. Many researchers suggest a single common queue for all physician, which has short waiting time, as the proper appointment system. These kinds of appointment systems, mostly assign the patient to the first available doctor and jeopardize the continuity of care, which is one of the main goals of primary care. These system makes the optimal policy more complicated and also results in lower revenues [OHare and Corlett, 2004]. Thus, in this study, we propose an approach which allows primary care providers to use independent queues for each doctor as been suggested by few other researchers [Rising et al., 1973, Cox et al., 1985].

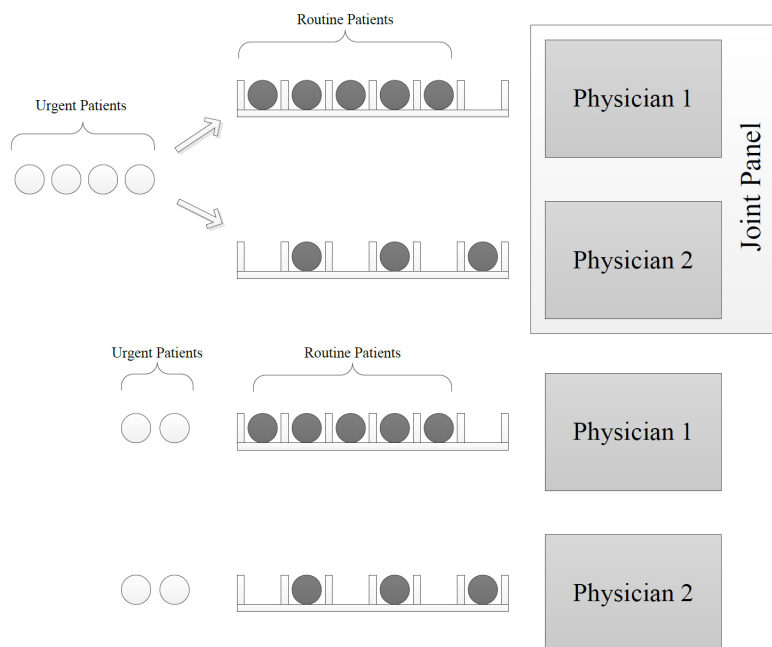
On the other hand, some public clinics do not give appointments to urgent patients, sending them to the first available doctor. This is the case in clinics studied by [Babes and Sarma, 1991] in Algeria and [Liu and Liu, 1998] in Hong Kong. Similarly, in our proposed model the same-day requests will be sent to the first available open slot regardless of the patients' preferred PCP.

#### 4.1. Joint Panel Scheduling Modeling

Similar to the presented model in Section 3.1, the main goal in this section is to find the optimal appointment schedule which minimizes patients' waiting time while maximizing providers' expected utility. Therefore, we develop a stochastic model that maximizes a weighted sum of provider's revenue, cost of lost sales, overtime cost, and patients' waiting cost. The provider's planning decisions consist the daily number of open slots for urgent patients and maximum numbers of slots allowed to be double

booked for all physicians who are working in a joint panel. The provider’s schedule specifies where these reserved slots are located.

Both single physician scheduling and joint panel scheduling systems are depicted in Figure 4.1. The main difference between these two systems is handling same-day requests. In the joint panel scheduling, urgent patients (same-day requests) will be assigned to the first open slot regardless of the physician; while in the single physician scheduling, each doctor receives his/her own same-day request.



**Figure 4.1** Joint panel scheduling vs. single physician scheduling.

Initially, we assume a particular case where we have just two physicians who have their own routine patients but they serve urgent patients based on first in first serve order. Later, we generalize our finding to characterize the general pattern in the optimal schedule in the case of multiple physicians. The demand for the routine patients follows a probability mass function (p.m.f.)  $q_i^{1r}$  for  $i = 0, 1, \dots, U_r^1$  for first physician and  $q_i^{2r}$  for  $i = 0, 1, \dots, U_r^2$  for the second one. The demand for urgent patients follows a p.m.f.  $q_i^s$  for  $i = 0, 1, \dots, U_s$  which is shared between both physician. We also assume the probability functions are independent of each other. Similar to



the previous model we assume that the service times for both types of patient are deterministic and equal. Thus, the daily appointment schedule can be divided into  $N$  intervals of equal length for each physician. Let  $i$  ( $i = 1, \dots, N$ ) represents the  $i^{\text{th}}$  time slot counted from the beginning of the day. The provider needs to determine the number of slots allocated to the routine patients for each physician,  $N_r^1$  and  $N_r^2$ . Further, let  $N_s^1$  denote the number of slots allocated to the same-day requests,  $N_s^1 = N - N_r^1$  and  $N_s^2 = N - N_r^2$ .

The provider receives revenue of  $p$  for each patient served, regardless of the type. If either a routine patient or a same-day request cannot be accommodated, the provider incurs a cost of lost sales ( $c_l$ ). We further assume that with probability  $\gamma$ , a routine patient would not show up for the appointment. All urgent patients, however, show up for sure.

All the aforementioned assumptions are made based on the data we obtained from a clinic, which has an average no-show rate of 8% for routine patients, but less than 0.1% for urgent patients. In order to mitigate the negative impacts associated with no-shows, the provider allows certain time slots to be double booked for each physician. However, all booked patients, if shown, must be served on the same day. The provider thus needs to work overtime if there are un-served patients by the end of the day. In this case, the provider incurs an overtime cost of  $c_v$  for each patient. Moreover, since double-booking all of the slots assigned to routine patients may not be optimal, we introduce another decision variables,  $A^1$  and  $A^2$ , to represent the maximum number of slots that can be double booked for the first and the second physician. See Table 4.1 for a summary of notation.

**Joint Panel Appointment Schedule.** In this appointment schedule the provider not only need to determine the number of slots allocated to the routine patients ( $N_s^1$  and  $N_s^2$ ) and the maximum number of slots that can be double booked ( $A^1$  and  $A^2$ ), But also needs to specify which time slots are designated for urgent patients, which

**Table 4.1** Joint Panel Scheduling Model's Notation

| Parameters: |   |
|-------------|---|
| $p$ :       | Revenue per patient   |
| $c_l$ :     | Cost of lost sales per patient  |
| $c_w$ :     | Waiting cost per slot per patient   |
| $c_v$ :     | Cost of overtime per slot   |
| $\gamma$ :  | No-show rate for routine patient  |
| $N$ :       | Number of slots available during the regular session for both physicians  |
| Decisions:  |   |
| $N_r^h$ :   | Number of slots assigned to routine patients for physician $h = 1, 2$   |
| $A^h$ :     | Maximum number of double-booked patients for physician $h = 1, 2$   |
| $x_i^h$ :   | Maximum number of patients allowed to be assigned to slot $i$ for physician $h = 1, 2$ ,<br>$i = 1, \dots, N$ and $x_i = 0, 1, 2$ |
| Variables:  |   |
| $D_r^h$ :   | Demand for routine patients, follows <i>pmf</i> : $q_i^r$ for $i = 0, 1, \dots, U_r^h$ for physician $h = 1, 2$                   |
| $D_s$ :     | Demand for urgent patients, follows <i>pmf</i> : $q_i^s$ for $i = 0, 1, \dots, U_s$   |
| $N_s^h$ :   | Number of slots assigned to same-day patients, $N_s^h = N - N_r^h$ for physician $h = 1, 2$                                       |
| $k_i^h$ :   | Number of routine patients that showed up, follows <i>pmf</i> : $B(k; n, \gamma)$ for physician $h = 1, 2$                        |
| $w_i^h$ :   | Number of patients waiting at the end of slot $i$ , $w_i = 0, 1, \dots, i$ for physician $h = 1, 2$                               |

are for routine patients, and which ones can be double-booked. Let  $x_i^h$  denote the maximum number of patients allowed to be assigned to time slot  $i$  for physician  $h$ . If  $x_i^h = 0$ , time slot  $i$  is reserved for urgent patients and no routine patients is booked for that slot. If  $x_i^h = 1$ , time slot  $i$  is reserved for routine patients and no double-bookings is allowed. If  $x_i^h = 2$ , time slot  $i$  can be double-booked if all other routine slots have been booked. Thus, the number of patients showed up at slot  $i$ ,  $k_i^h$  can be computed as follows:

$$k_i^h | \{x_i^h = 1\} = \begin{cases} 0 & B(0; 1, \bar{\gamma}); \\ 1 & B(1; 1, \bar{\gamma}). \end{cases} \quad \text{and} \quad k_i^h | \{x_i^h = 2\} = \begin{cases} 0 & B(0; 2, \bar{\gamma}); \\ 1 & B(1; 2, \bar{\gamma}); \\ 2 & B(2; 2, \bar{\gamma}). \end{cases} \quad (4.1)$$

where  $\bar{\gamma} = 1 - \gamma$  is the probability that a routine patient shows up. The term  $B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$  is the binomial distribution, when  $n = 0$ ,  $B(k; n, p) = 0$ . Given a pre-determined schedule,  $\{x_i^1\}_{i=1, \dots, N}$  and  $\{x_i^2\}_{i=1, \dots, N}$ , the provider can expect  $n_r^1$  routine patients for the first physician and  $n_r^2$  routine patients for the second one, where

$$\begin{aligned} n_r^1 &= \sum_{i=1}^N \sum_{k_i^1=0}^{x_i^1} k_i^1 \cdot B(k_i^1; x_i^1, \bar{\gamma}), \\ n_r^2 &= \sum_{i=1}^N \sum_{k_i^2=0}^{x_i^2} k_i^2 \cdot B(k_i^2; x_i^2, \bar{\gamma}). \end{aligned} \quad (4.2)$$

Because the provider also cares about the waiting times of the patients, let  $c_w$  be the waiting cost per time slot for each patient. Let  $w_i^h$  represent the number of patients waiting at the end of time  $i$  for physician  $h$ . Then, the probability of having  $j$  patients waiting at the end of slot  $i$ ,  $P^h(w_i^h = j)$  for physician  $h$ , can be derived using the following recursive function:

$$\begin{aligned} P^h(w_i^h = j) &= P^h(w_{i-1}^h = j + 1) \cdot P(k_i^h = 0 | x_i^h) + P^h(w_{i-1}^h = j) \cdot P(k_i^h = 1 | x_i^h) \\ &\quad + P^h(w_{i-1}^h = j - 1) \cdot P(k_i^h = 2 | x_i^h). \end{aligned} \quad (4.3)$$

The first term in Equation (4.3) corresponds to the situation where  $j+1$  patients are waiting at the end of slot  $i-1$  for physician  $h = 1, 2$ , and no patients show up for slot  $i$ . The second term captures the case where  $j$  patients are waiting at the end of slot  $i-1$  and one patient shows up for slot  $i$ . The last term captures the situation where  $j-2$  patients are waiting at the end of slot  $i-1$  and two new patients who are both booked for slot  $i$  show up. In all three cases, only one patient is served in slot  $i$  and all remaining patients would wait.

**Provider's Revenue.** For each patient served, regardless of the type, the provider obtains a revenue  $p$ . Thus, the provider's expected revenue depends on the realized demands  $(D_r^1, D_r^2, D_s^1, D_s^2)$  of both routine and urgent patients for both physician, and their allocated capacities  $(N_r^1, N_r^2, N_s^1, N_s^2, A^1$  and  $A^2)$ . Let us first

define an operating function,  $\psi(\cdot)$ , as follows:

$$\psi(x, l_r^1, u_r^1, l_r^2, u_r^2, l_s, u_s) = \sum_{D_r^1=l_r^1}^{u_r^1} \sum_{D_r^2=l_r^2}^{u_r^2} \sum_{D_s=l_s}^{u_s} (x) q_{D_s}^s q_{D_r^2}^r q_{D_r^1}^r, \quad (4.4)$$

where  $l_r^1$  and  $u_r^1$  are the lower and upper bounds of the demand for routine patients for the first physician,  $l_r^2$  and  $u_r^2$  are the lower and upper bounds of the demand for routine patients for the second physician and  $l_s$  and  $u_s$  are the lower and upper bounds of the demand for same-day requests. Given a pair of capacity allocations for the routine patients ( $N_r^1$  and  $N_r^2$ ) and the urgent patient ( $N_s^1$  and  $N_s^2$ ) as well as a maximum number of slots that can be double-booked ( $A^1$  and  $A^2$ ), the provider's expected revenue, denoted by  $\Pi(N_r^1, N_r^2, A^1, A^2)$ , can be computed by:

$$\begin{aligned} \Pi(N_r^1, N_r^2, A^1, A^2) &= p \cdot \psi((D_r^1 + D_r^2)\bar{\gamma} + \min\{D_s, 2N - D_r^1 - D_r^2\}, 0, N_r^1, 0, N_r^2, 0, U_s) \\ &+ p \cdot \psi(D_r^1\bar{\gamma} + n_r^2 + \min\{D_s, N - D_r^1 + N_s^2\}, 0, N_r^1, N_r^2, N_r^2 + A^2, 0, U_s) \\ &+ p \cdot \psi(D_r^1\bar{\gamma} + n_r^2 + \min\{D_s, N - D_r^1 + N_s^2\}, 0, N_r^1, N_r^2 + A^2, U_r^2, 0, U_s) \\ &+ p \cdot \psi(n_r^1 + D_r^2\bar{\gamma} + \min\{D_s, N_s^1 + N_s^2\}, N_r^1, N_r^1 + A^1, 0, N_r^2, 0, U_s) \\ &+ p \cdot \psi(n_r^1 + n_r^2 + \min\{D_s, N_s^1 + N_s^2\}, N_r^1, N_r^1 + A^1, N_r^2, N_r^2 + A^2, 0, U_s) \\ &+ p \cdot \psi(n_r^1 + n_r^2 + \min\{D_s, N_s^1 + N_s^2\}, N_r^1, N_r^1 + A^1, N_r^2 + A^2, U_r^2, 0, U_s) \\ &+ p \cdot \psi(n_r^1 + D_r^2\bar{\gamma} + \min\{D_s, N_s^1 + N_s^2\}, N_r^1 + A^1, U_r^1, 0, N_r^2, 0, U_s) \\ &+ p \cdot \psi(n_r^1 + n_r^2 + \min\{D_s, N_s^1 + N_s^2\}, N_r^1 + A^1, U_r^1, N_r^2, N_r^2 + A^2, 0, U_s) \\ &+ p \cdot \psi(n_r^1 + n_r^2 + \min\{D_s, N_s^1 + N_s^2\}, N_r^1 + A^1, U_r^1, N_r^2 + A^2, U_r^2, 0, U_s). \end{aligned} \quad (4.5)$$

The first term in Equation (4.5) refers to the case where the demand for routine patients for both physicians is less than the allocated capacity. Consequently, there are  $2N - D_r^1 - D_r^2$  un-utilized slots, which can be assigned to the same-day requests. The second term corresponds to the situation where all demand for routine patients of the second physician can be accommodated with up to  $A^2$  slots double-booked. The third term refers to the case where all capacity of the second physician ( $N_r^2 + A^2$ )

for routine patients is used up and the remaining routine patients are not served. Similarly, the next three terms represent the situation where the demand of routine patients for the first physician exceeds  $N_r^1$ . Finally, the last three terms refer to the situation where all the available capacity for routine patients, for both physicians, is used up.

**Provider's Lost Sales.** Once the capacity planning decisions are made, the provider can only serve up to  $N_r^1 + N_r^2 + A^1 + A^2$  routine patients and  $N_s^1 + N_s^2$  urgent patients in a day. If any of the demands exceeds the capacity, the provider incurs a lost sale cost. Let  $L(N_r^1, N_r^2, A^1, A^2)$  denote the provider's expected cost from lost sales, and it can be computed by

$$\begin{aligned}
L(N_r^1, N_r^2, A^1, A^2) &= c_l \cdot \psi(D_s - (2N - D_r^1 - D_r^2), 0, N_r^1, 0, N_r^2, 2N - D_r^1 - D_r^2, U_s) \\
&+ c_l \cdot \psi(D_s - (N - D_r^1 - N_s^2), 0, N_r^1, N_r^2, N_r^2 + A^2, N - D_r^1 + N_s^2, U_s) \\
&+ c_l \cdot \psi(D_r^2 - (N_r^2 + A^2), 0, N_r^1, N_r^2 + A^2, U_r^2, 0, N - D_r^1 + N_s^2) \\
&+ c_l \cdot \psi(D_r^2 + D_s - (2N - 2N_s^2 + A^2 - D_r^1), 0, N_r^1, N_r^2 + A^2, U_r^2, N - D_r^1 + N_s^2, U_s) \\
&+ c_l \cdot \psi(D_s - (N_s^1 + N - D_r^2), N_r^1, N_r^1 + A^1, 0, N_r^2, N - D_r^2 + N_s^1, U_s) \\
&+ c_l \cdot \psi(D_s - (N_s^1 - N_s^2), N_r^1, N_r^1 + A^1, N_r^2, N_r^2 + A^2, N_s^1 + N_s^2, U_s) \\
&+ c_l \cdot \psi(D_r^2 - (N_r^2 + A^2), N_r^1, N_r^1 + A^1, N_r^2 + A^2, U_r^2, 0, N_s^1 + N_s^2) \\
&+ c_l \cdot \psi(D_r^2 + D_s - (N_r^2 + A^2 + N_s^1 - N_s^2), N_r^1, N_r^1 + A^1, N_r^2 + A^2, U_r^2, N_s^1 + N_s^2, U_s) \\
&+ c_l \cdot \psi(D_r^1 - (N_r^1 + A^1), N_r^1 + A^1, U_r^1, 0, N_r^2, 0, N - D_r^2 + N_s^1) \\
&+ c_l \cdot \psi(D_s + D_r^1 - (2N - 2N_s^1 + A^1 - D_r^2), N_r^1 + A^1, U_r^1, 0, N_r^2, N - D_r^2 + N_s^1, U_s) \\
&+ c_l \cdot \psi(D_r^1 - (N_r^1 + A^1), N_r^1 + A^1, U_r^1, N_r^2, N_r^2 + A^2, 0, N_s^1 + N_s^2) \\
&+ c_l \cdot \psi(D_s + D_r^1 - (N_r^1 + A^1 + N_s^1 + N_s^2), N_r^1 + A^1, U_r^1, N_r^2, N_r^2 + A^2, N_s^1 + N_s^2, U_s) \\
&+ c_l \cdot \psi(D_r^1 + D_r^2 - (N_r^1 + A^1 + N_r^2 + A^2), N_r^1 + A^1, U_r^1, N_r^2 + A^2, U_r^2, 0, N_s^1 + N_s^2) \\
&+ c_l \cdot \psi(D_s + D_r^2 + D_r^2 - (2N + A^1 + A^2), N_r^1 + A^1, U_r^1, N_r^2 + A^2, U_r^2, N_s^1 + N_s^2, U_s).
\end{aligned} \tag{4.6}$$

The first two terms in Equation (4.6) correspond to the case where there are more urgent patients than available open slots, and thus the urgent patients must be referred to other providers. The third term refers to the scenario where the demand for routine patients, from the second physician's list, is over the maximum capacity ( $N_r^1 + A^1$ ) with double-booking, and thus some of the routine patients would not be served. Similarly, all the other terms represent the situation where demand (either routine patients or same-day request) exceed the provider's capacities.

**Provider's Overtime Cost.** While adopting the double-booking strategy can mitigate the negative effects of no-shows, it may require the provider and other clinic staff to work overtime. Consider the following situation, slot  $i$  is double-booked for one of the physicians. When both patients show up, only one can be served during the scheduled time interval. Because clinics regularly operate based on "first come, first serve", the other patient will be seeing the provider in slot  $i + 1$ . If any patient who is booked for slot  $i + 1$  shows up, that patient will also be delayed by at least one slot. These delays are easily concatenated and may result in some un-served at the end of the day, i.e. slot  $N$ . The provider's expected overtime cost thus depends on the appointment schedules  $\{x_i^h\}_{i=1,\dots,N}$  and  $h = 1, 2$ . Let  $O(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\})$  denote the provider's overtime cost, and it can be calculated as follows:

$$\begin{aligned}
O(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\}) &= c_v \cdot \psi\left(\sum_{j=0}^{\min\{D_r^2 - N_r^2, A^2\}} j \cdot P(w_N^2 = j), 0, N_r^1, N_r^2, U_r^2, 0, U_s\right) \\
&+ c_v \cdot \psi\left(\sum_{j=0}^{\min\{D_r^1 - N_r^1, A^1\}} j \cdot P(w_N^1 = j), N_r^1, U_r^1, 0, N_r^2, 0, U_s\right) \\
&+ c_v \cdot \psi\left(\sum_{j=0}^{\min\{D_r^1 - N_r^1, A^1\}} j \cdot P(w_N^1 = j) + \sum_{j=0}^{\min\{D_r^2 - N_r^2, A^2\}} j \cdot P(w_N^2 = j), N_r^1, U_r^1, N_r^2, U_r^2, 0, U_s\right). \tag{4.7}
\end{aligned}$$

where  $P(w_N^1 = j)$  and  $P(w_N^2 = j)$  in Equation (4.7) is derived using Equation (4.3) recursively. Note that delays at the end of the day only occurs when there are double-booked patients, thus we only need to consider the scenarios where the demand

for routine patients exceeds its allocated capacity, and thus the number of patients at the end of slot  $N$  will not exceed  $\min\{D_r^1 - N_r^1, A^1\}$  for the first physician and  $\min\{D_r^2 - N_r^2, A^2\}$  for the second one.

**Patients' Waiting Cost.** Similar to the provider's overtime cost, the appointment schedules  $\{x_i^1\}_{i=1,\dots,N}$  and  $\{x_i^2\}_{i=1,\dots,N}$  with double-booking strategy highly affects the waiting times for patients. Let  $W(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\})$  denote the aggregated patients' waiting cost, and it equals to

$$\begin{aligned}
W(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\}) &= c_w \cdot \psi\left(\sum_{i=1}^N \min\{i, D_r^2 - N_r^2, A^2\} \sum_{j=0}^{\min\{i, D_r^2 - N_r^2, A^2\}} j P(w_i^2 = j), 0, N_r^1, N_r^2, U_r^2, 0, U_s\right) \\
&+ c_w \cdot \psi\left(\sum_{i=1}^N \min\{i, D_r^1 - N_r^1, A^1\} \sum_{j=0}^{\min\{i, D_r^1 - N_r^1, A^1\}} j \cdot P(w_i^1 = j), N_r^1, U_r^1, 0, N_r^2, 0, U_s\right) \\
&+ c_w \cdot \psi\left(\sum_{i=1}^N \min\{i, D_r^1 - N_r^1, A^1\} \sum_{j=0}^{\min\{i, D_r^1 - N_r^1, A^1\}} j \cdot P(w_i^1 = j) + \sum_{j=0}^{\min\{i, D_r^2 - N_r^2, A^2\}} j \cdot P(w_i^2 = j), N_r^1, U_r^1, N_r^2, U_r^2, 0, U_s\right).
\end{aligned} \tag{4.8}$$

In Equation (4.8), the number of patients waiting ( $j$ ) at slot  $i$  should exceed neither  $i$  nor the total number of double-booked patients  $\min\{D_r^h - N_r^h, A^h\}$  where  $h = 1, 2$ .

**Provider's Utility.** The provider's utility can be derived by combining the revenue and cost functions. Let  $U(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\})$  denote the provider's utility:

$$\begin{aligned}
U(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\}) &= \Pi(N_r^1, N_r^2, A^1, A^2) - L(N_r^1, N_r^2, A^1, A^2) \\
&\quad - O(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\}) \\
&\quad - W(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\}).
\end{aligned} \tag{4.9}$$

Let  $(N_r^{1*}, N_r^{2*}, A^{1*}, A^{2*}, \{x_i^{1*}\}, \{x_i^{2*}\})$  be the joint optimal decisions on the number of slots allocated for routine patients, the maximum number of slots allowed to be double-booked and the appointment schedule respectively for both physicians, that

maximize the provider's utility function, that is,

$$(N_r^{1*}, N_r^{2*}, A^{1*}, A^{2*}, \{x_i^{1*}\}, \{x_i^{2*}\}) = \max_{N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\}} U(N_r^1, N_r^2, A^1, A^2, \{x_i^1\}, \{x_i^2\}). \quad (4.10)$$

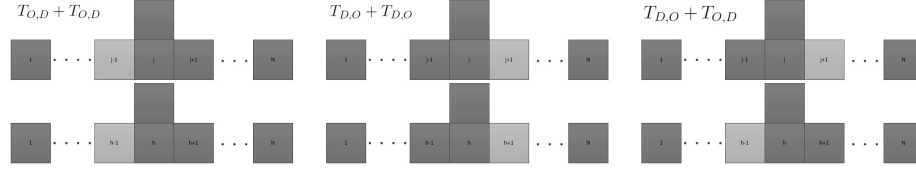
## 4.2. Analytical Solutions

In this section, we start with the simplest case where it is optimal to double book exactly one routine patient and leave only one slot open for each physician. Let  $T_{O,D}$  be the schedule in which, the open slot ( $O$ ) is placed before the double-booked slot ( $D$ ). Also,  $T_{D,O}$  be the schedule in which, the open slot is placed after the double-booked slot. Therefore, all possible schedules can be divided into three general scenarios:

- $T_{O,D} + T_{O,D}$ : the open slots are placed before the double-booked slot for both physicians.
- $T_{D,O} + T_{D,O}$ : the open slots are placed after the double-booked slot for both physicians.
- $T_{D,O} + T_{O,D}$ : one open slot is placed after the double-booked slot for first physicians and the other one is placed before the double-booked slot for the second physician.

In order to make a better visualization, we can exhibit each schedule as shown in Figure 4.2. Each square represents one slot, which has the same length as the other slots (see Section 3.1 for further information). The light gray squares present open slots and dark gray ones represent slots which allocated to the routine patients. We also demonstrate double-booked slots with two dark gray squares on top of each other. Moreover, the top schedule demonstrates the schedule of first physician and the bottom one shows the second physician's schedule.





**Figure 4.2** General scenarios for joint panel scheduling.

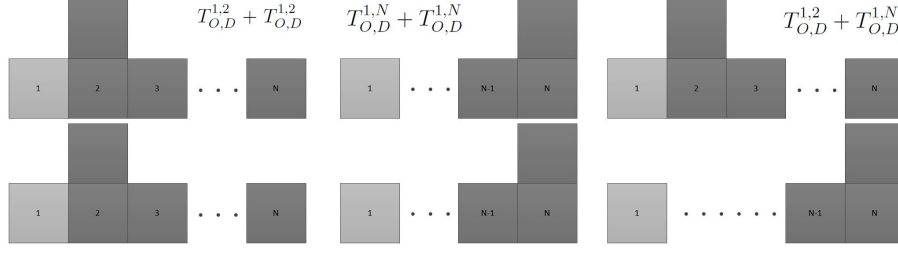
These schedules are too general to argue and be compared. Therefore, first, we focus on  $T_{O,D} + T_{O,D}$  to narrow down all possible forms of it.

**Schedule  $T_{O,D} + T_{O,D}$ .** In this schedule, the open slot can be anywhere before the double-booked slot and its placement does not affect the total expected waiting time and overtime costs. Although, in this schedule, the double-booked slot can be anywhere, but we can concisely express a condition in which a certain form of this schedule is optimal.

**Proposition 5.** *Among all possible schedule  $T_{O,D} + T_{O,D}$ , if  $c_w \left( \frac{1-\gamma}{\gamma} \right) \leq c_v$  then the schedule which has an open slot at the beginning followed by a double-booked slot right after that for both physicians is the best. on the other hand, if  $c_w \left( \frac{1-\gamma}{\gamma} \right) \geq c_v$  the schedule which has an open slot at the beginning and a double-booked slot at the end of the session for both physician is the best.*

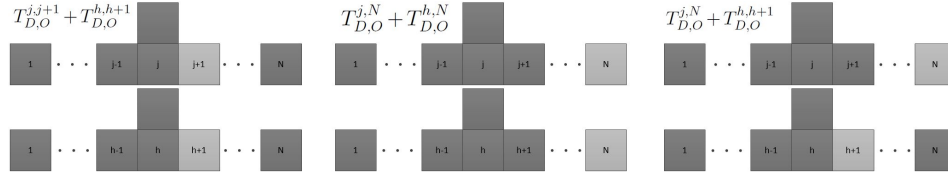
The proof of this proposition and all subsequent propositions can be found in Appendix B. Proposition 5 drastically reduces the number of schedules, which are derived from  $T_{O,D} + T_{O,D}$ , into three schedules. These three schedules are demonstrated in Figure 4.3.

**Schedule  $T_{D,O} + T_{D,O}$ .** In this schedule, the placement of the open slot affects the expected waiting time and overtime costs since it is after the double-booked slot. Similarly, we can present a condition in which a particular form of this schedule is optimal.



**Figure 4.3** Optimal schedule with  $T_{O,D} + T_{O,D}$ .

**Proposition 6.** Among all  $T_{D,O} + T_{D,O}$  schedules (the open slot can be anywhere after the double-booked slot), if  $\gamma \leq q_0^s + q_1^s$  then schedule  $T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1}$  is the best amongst all the other schedules. Otherwise, schedule  $T_{D,O}^{j,N} + T_{D,O}^{h,N}$  is the best. The schedule  $T_{D,O}^{j,N} + T_{D,O}^{h,h+1}$  is never optimal.

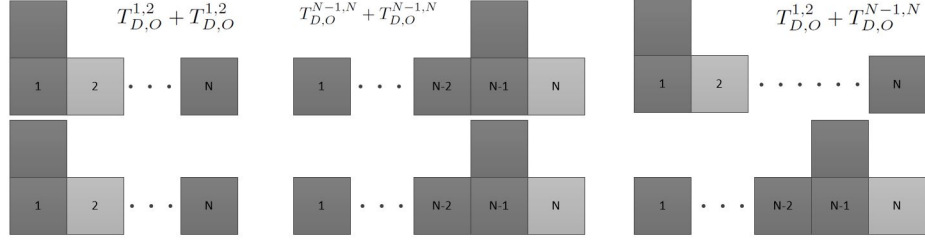


**Figure 4.4** Optimal schedule with  $T_{D,O} + T_{D,O}$ .

This proposition divides  $T_{D,O} + T_{D,O}$  into three general forms and states that the third one ( $T_{D,O}^{j,N} + T_{D,O}^{h,h+1}$ ) is always worst than the other two. However, these two schedules (Shown in Figure 4.4) are still general cases and need more exploration.

**Proposition 7.** Among all possible schedule  $T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1}$  (the open slot is right after the double-booked slot for both physician), if  $c_w \left( \frac{1-\gamma}{\gamma} \right) \leq c_v$  then schedule  $T_{D,O}^{1,2} + T_{D,O}^{1,2}$  is the best. otherwise, schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$  is the best. The schedule  $T_{D,O}^{1,2} + T_{D,O}^{N-1,N}$  is never optimal.

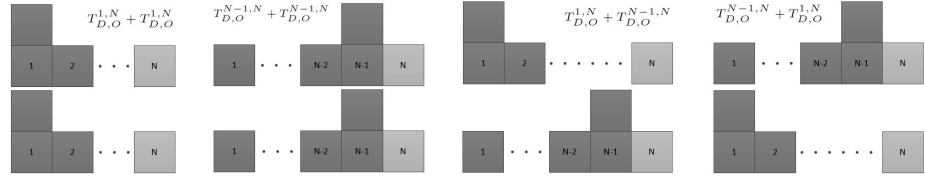
This Proposition reduces number of schedules, which are derived from  $T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1}$ , into three schedules. All of these schedules are demonstrated in Figure 4.5. This proposition also states that the third schedule ( $T_{D,O}^{1,2} + T_{D,O}^{N-1,N}$ ) is never optimal.



**Figure 4.5** Optimal schedule with  $T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1}$ .

**Proposition 8.** Among all possible schedule  $T_{D,O}^{j,N} + T_{D,O}^{h,N}$  (the open slot is after the double-booked slot and at the end of the session for both physician), if  $c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s)} - 1 \right) \leq c_v$  then the schedule  $T_{D,O}^{1,N} + T_{D,O}^{1,N}$  is the best. Moreover, if  $c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s)} - 1 \right) \geq c_v$  then the schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$  is the best. Otherwise, the schedule  $T_{D,O}^{1,N} + T_{D,O}^{N-1,N}$  is the best amongst the others. The schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{1,N}$  is never optimal.

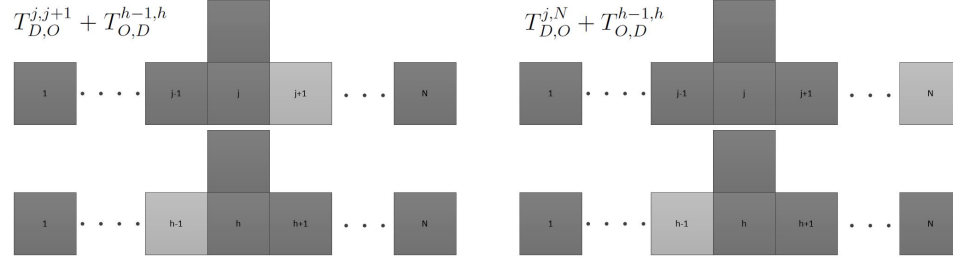
Proposition 8 reduces number of schedules, which are derived from  $T_{D,O}^{j,N} + T_{D,O}^{h,N}$  into four schedules. All of these schedules are demonstrated in Figure 4.6. This proposition also states that the fourth schedule ( $T_{D,O}^{N-1,N} + T_{D,O}^{1,N}$ ) is never optimal.



**Figure 4.6** Optimal schedule with  $T_{D,O}^{j,N} + T_{D,O}^{h,N}$ .

**Schedule  $T_{D,O} + T_{O,D}$ .** In this schedule the placement of the open slot affects the expected waiting time and overtime costs since it is after the double-booked slot. Similarly, we can present a condition in which a particular form of this schedule is optimal.

**Proposition 9.** Among all  $T_{D,O} + T_{O,D}$  schedules, if  $\gamma \leq q_0^s + q_1^s$ , then schedule  $T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h}$  is the best amongst all the other schedules. Otherwise, schedule  $T_{D,O}^{j,N} + T_{O,D}^{h-1,h}$  is the best.



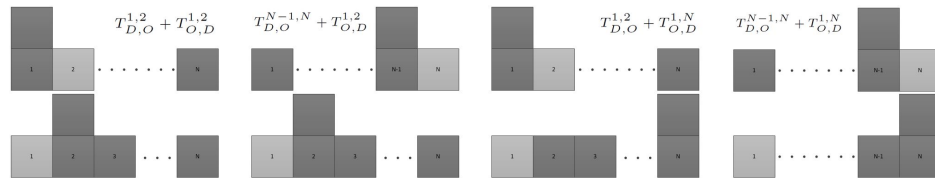
**Figure 4.7** Optimal schedule with  $T_{D,O} + T_{O,D}$  .

This proposition divide  $T_{D,O} + T_{O,D}$  into two general forms. However, these two schedules (shown in Figure 4.7) are still general cases and need more exploration.

**Proposition 10.** *Among all possible schedule  $T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h}$ , it is optimal to place first open slot, the one which is before the double-booked slot, prior to the open slot which is after the double-booked slot.*

This proposition states that in  $T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h}$  it is optimal to have  $j \geq h - 2$ . Therefore, we can divide  $T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h}$  into four schedules. all these four schedules are shown in Figure 4.8.

**Proposition 11.** *Among all  $T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h}$  schedules, if  $c_w \left( \frac{1-\gamma}{\gamma} \right) \leq c_v$ , then schedule  $T_{D,O}^{1,2} + T_{O,D}^{1,2}$  is the best. Otherwise, the schedule  $T_{D,O}^{N-1,N} + T_{O,D}^{1,N}$  is the best.*

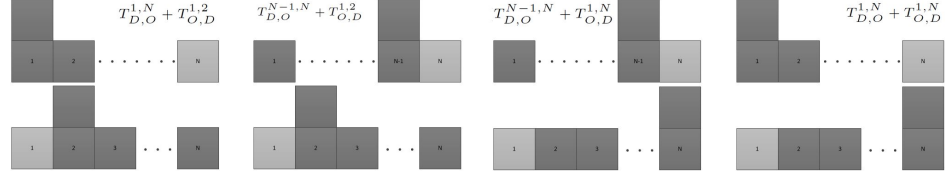


**Figure 4.8** Optimal schedule with  $T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h}$  .

Proposition 11 states that the other two schedules ( $T_{D,O}^{N-1,N} + T_{O,D}^{1,2}$  and  $T_{D,O}^{1,2} + T_{O,D}^{1,N}$ ), which are derived from Proposition 10, are never optimal.

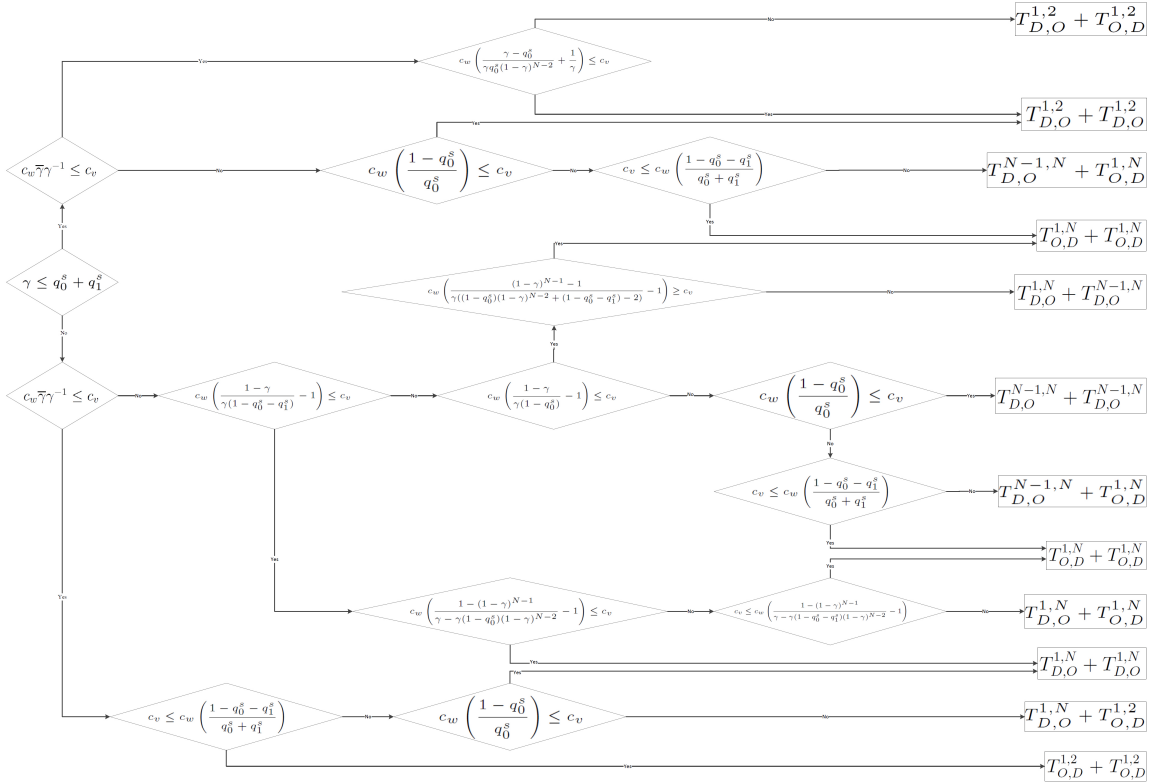
**Proposition 12.** *Among all possible schedule  $T_{D,O}^{j,N} + T_{O,D}^{h-1,h}$ , if  $c_w \left( \frac{1-\gamma}{\gamma} \right) \leq c_v$  then the schedule  $T_{D,O}^{1,N} + T_{O,D}^{1,2}$  is the best. Moreover, if  $c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s - q_1^s)} - 1 \right) \geq c_v$  then the*

schedule  $T_{D,O}^{N-1,N} + T_{O,D}^{1,N}$  is the best. Otherwise, the schedule  $T_{D,O}^{1,N} + T_{O,D}^{1,N}$  is the best amongst the others. The schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{1,N}$  is never optimal.



**Figure 4.9** Optimal schedule with  $T_{D,O}^{j,N} + T_{D,O}^{h-1,h}$ .

Proposition 12 divide all forms of schedule  $T_{D,O}^{j,N} + T_{D,O}^{h-1,h}$  into four schedules which are shown in Figure 4.9. It also states that the schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{1,N}$  is never optimal.



**Figure 4.10** Flowchart for optimal schedule selection.

Figure 4.10 summarizes all the propositions stated above which conceptually, provides similar guidelines as the ones presented in Figure 3.6 for the single physician.

As one can see, the ratio of the overtime cost and the waiting cost highly influences the position of the double-booked slots in the optimal schedule. If the ratio is relatively small, then the double-booked slots should be placed at the beginning of the session for both physicians in order to take advantage of any unutilized slot due to a no-show. In contrast, if the ratio is +high, then the double-booked slots should be pushed toward the end of the session to minimize the overall expected waiting time. These findings are aligned with our previous claims in the single physician scheduling (Section 3.2.)

Furthermore, the relative position of open slots to the double-booked slots are determined by two major factors: (i) The relationship between the probability of no-show of routine patients and the probability of having one or fewer same-day request and (ii) the ratio of the waiting cost to the overtime cost. If  $\gamma \leq q_0^s + q_1^s$ , which implies that chance of not having enough same-day requests is higher than that of the no-show of routine patients, and the ratio of waiting cost to the overtime cost is low, then the open slot should be placed right after the double-booked slot for both physicians. This allows both physicians to serve the double-booked patient in the highly likely case of no same-day request. In addition, if the ratio is moderate, then the open slot should be placed right after the double-booked slot for one physician and prior to the double-booked slot for the other physicians. This allows the joint panel to justifies the usage of unutilized open slot due to lack of requests and the reduction of overall waiting time. On the other hand, if  $\gamma \geq q_0^s + q_1^s$  and the ratio is less than  $q_0^s \bar{q}_0^{s-1}$ , then the open slot should be placed after the double-booked slot and at the end of the session for both physicians. Since the possibility of a no-show is higher than having one or zero same-day request and waiting cost is relatively small, both physicians are willing to use a block of single-booked slots instead of an open slot to serve one of the double-booked patients. Also, if the ratio is moderate,  $(q_0^s \bar{q}_0^{s-1} \leq \frac{c_w}{c_v} \leq (q_0^s + q_1^s)(1 - q_0^s - q_1^s)^{-1})$ , it is optimal to place the open slot prior to the double-booked slot for one of the physicians and at the end of the session for

the other one. This allows the joint panel to balance the utilization of any opened-up single-booked slot, due to high no-shows, and the reduction of overall waiting time from pushing double-booked slots toward the end of the session. Finally, if the ratio of the waiting cost and overtime cost is higher than  $(q_0^s + q_1^s)(1 - q_0^s - q_1^s)^{-1}$ , regardless of no-show rate and same-day requests, it is optimal to place the open slot anywhere prior to the double-booked slot for both physicians. The high waiting cost forces both physicians to not rely on unutilized open slot, due to lack of sam-day requests, and therefore, the open slots are placed prior to the double-booked slots for both physicians.

### 4.3. Computational Solutions

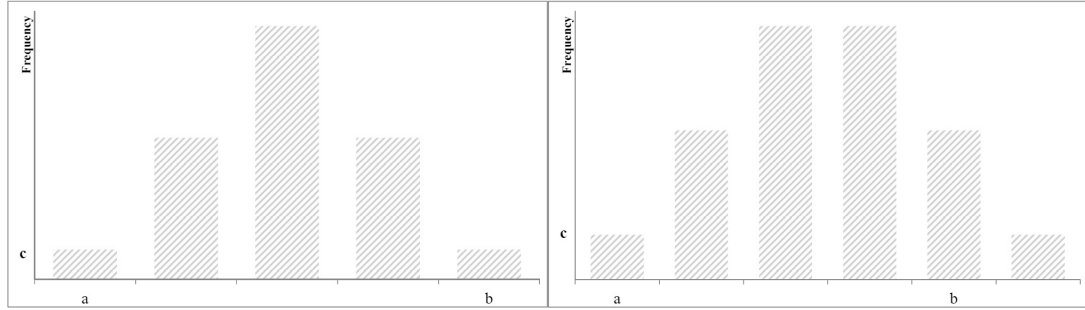
Due to the combinatorial characterization of the no-shows in each scheduled or double booked slots, the traditional approach of solving newsvendor problems does not apply and we use complete enumeration algorithm to solve this problem.

#### 4.3.1 Characterizations of Joint Panel Optimization

We use the complete enumeration algorithm to find the optimal schedules along with the optimal capacity planning decisions ( $N_r^*$ ,  $N_s^*$  and  $A^*$ ). Because the maximum number of patients allowed for each slot  $i$ ,  $x_i^1$  and  $x_i^2$ , can take three integer values: 0, 1 and 2, the runtime of this algorithm is exponential ( $O(3^{2N})$ ). Therefore, we only use this method for smaller capacity ( $N \leq 7$ ).

Additionally, Since the same-day requests will be assigned to both physicians, in order to maintain similarity of demand with the single physician schedules, we use a symmetric discrete triangular distribution  $T[a, b, c]$  for same-day requests which is shown in Figure 4.11.

**Case 1: low no-show rate, low waiting cost** ( $\gamma \leq q_0^s + q_1^s$  and  $c_w \bar{\gamma} \gamma^{-1} \leq c_v$ ). The optimal capacity planning decisions ( $N_r^{1*}$ ,  $N_r^{2*}$ ,  $N_s^{1*}$ ,  $N_s^{2*}$ ,  $A^{1*}$  and  $A^{2*}$ ) along



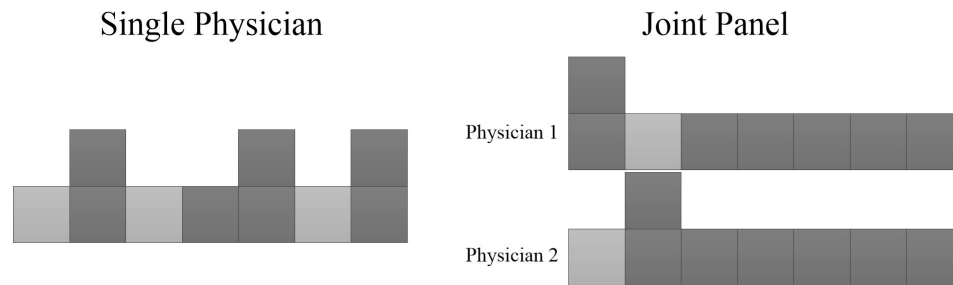
**Figure 4.11** Different types of symmetric discrete triangular distributions.

with optimal expected utilities ( $U^*$ ) and the optimal joint panel generated using the following parameters  $N = 7$ ,  $\gamma = 8\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 3$ ,  $D_r \sim U[0, 7]$  for both physicians, and  $D_s \sim T[0, 6]$  are provided in Table 4.2 and Figure 4.12.

**Table 4.2** Optimal Capacity Planning Decisions - Case 1

|                  | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|------------------|---------|---------|-------|-------|
| Single Physician | 4       | 3       | 3     | 219.9 |
| Joint Panel      | 6       | 1       | 1     | 448.8 |
|                  | 6       | 1       | 1     |       |

The joint panel schedule provides same-day access with less waiting time for urgent patients. As a result, as demonstrated in Table 4.2, the joint panel schedule allocates fewer slots for urgent patients compared to the single physician schedule and consequently fewer number of slots are allowed to be double-booked.



**Figure 4.12** Optimal schedules: single physician vs. joint panel - case 1.



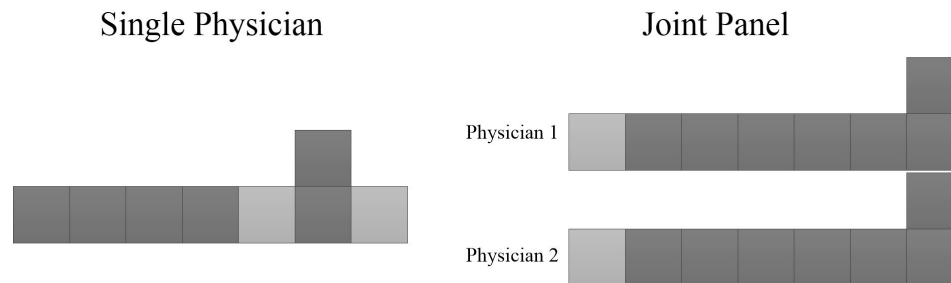
The optimal joint panel schedule assigns only one double booked slot and one open slot for each physician, which resembles the same assumption we made in Section 4.2. Therefore, it is optimal to place a double-booked slot at the beginning of the session followed by an open slot for the first physician and leave one slot open at the beginning of the session followed by a double-booked slot for the second physician.

**Case 2: low no-show rate, high waiting cost** ( $\gamma \leq q_0^s + q_1^s$  and  $c_w \left( \frac{1-q_0^s-q_1^s}{q_0^s+q_1^s} \right) \geq c_v$ ). The optimal capacity planning decisions ( $N_r^{1*}, N_r^{2*}, N_s^{1*}, N_s^{2*}, A^{1*}$  and  $A^{2*}$ ) along with optimal expected utilities ( $U^*$ ) and the optimal joint panel generated using the following parameters  $N = 7, \gamma = 8\%, p = 50, c_l = 50, c_v = 45, c_w = 13, D_r \sim U[0, 7]$  for both physicians, and  $D_s \sim T[0, 6]$  are provided in Table 4.3 and Figure 4.13.

**Table 4.3** Optimal Capacity Planning Decisions - Case 2

|                  | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|------------------|---------|---------|-------|-------|
| Single Physician | 5       | 2       | 1     | 200.8 |
| Joint Panel      | 6       | 1       | 1     | 445.5 |
|                  | 6       | 1       | 1     |       |

In this case, both schedules (single physician and joint panel) leave two open slots in the session. However, the joint panel schedule can serve more routine patients than single physician schedule due to urgent patient sharing among both physicians.



**Figure 4.13** Optimal schedules: single physician vs. joint panel - case 2.

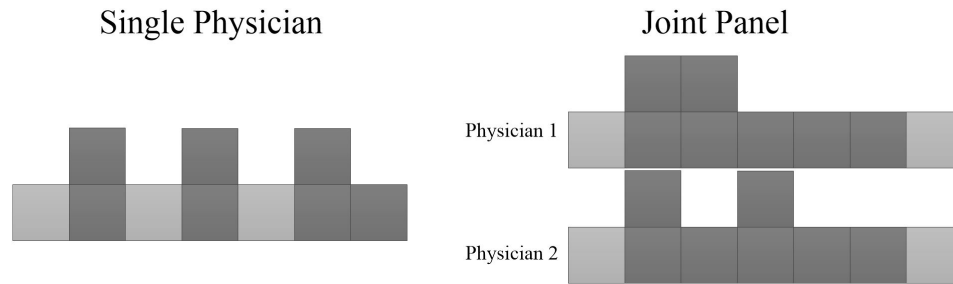
Similar to the previous case, the optimal joint panel schedule assigns only one double booked slot and one open slot for each physician. Therefore, as shown in Figure 4.10, The optimal schedule assigns one open slot at the beginning of the session and places the double-booked slot at the end of the session for both physicians in order to minimize the waiting time.

**Case 3: high no-show rate, low waiting cost** ( $\gamma \geq q_0^s + q_1^s$  and  $c_w \left(\frac{1-q_0^s}{q_0^s}\right) \leq c_v$ ). The optimal capacity planning decisions ( $N_r^{1*}$ ,  $N_r^{2*}$ ,  $N_s^{1*}$ ,  $N_s^{2*}$ ,  $A^{1*}$  and  $A^{2*}$ ) along with optimal expected utilities ( $U^*$ ) and the optimal joint panel generated using the following parameters  $N = 7$ ,  $\gamma = 30\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 3$ ,  $D_r \sim U[0, 7]$  for both physicians, and  $D_s \sim T[0, 6]$  are provided in Table 4.4 and Figure 4.14.

**Table 4.4** Optimal Capacity Planning Decisions - Case 3

|                  | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|------------------|---------|---------|-------|-------|
| Single Physician | 4       | 3       | 3     | 177.3 |
| Joint Panel      | 5       | 2       | 2     | 382.6 |
|                  | 5       | 2       | 2     |       |

Due to the high no-show rate, in this case, the joint panel schedule allocates more open slots compared to the single physician schedule to provide more access for urgent patients and uses even fewer number of double-booked slots.



**Figure 4.14** Optimal schedules: single physician vs. joint panel - case 3.

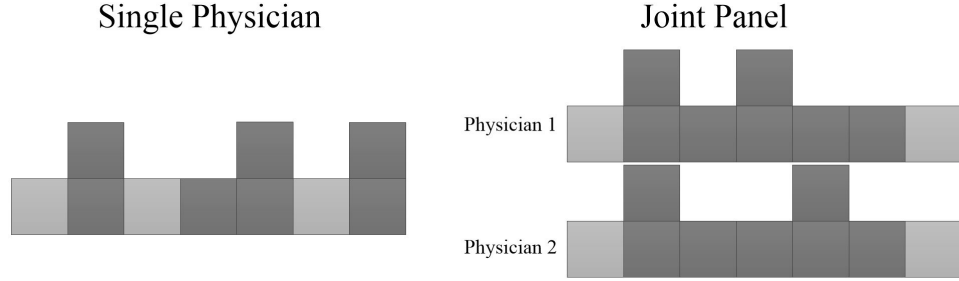
Recall that in the case where only one double booked slot and one open slot are assigned, it is optimal to place a double-booked slot at the beginning of the session and leave an open slot at the end of the session for both physicians when the waiting cost is low and the no-show rate is high. The optimal schedule  $(T_{D,O}^{1,N} + T_{D,O}^{1,N})$  allows the physician to serve one of the double-booked patients in a single booked slot if any of the reserved patients is a no-show. The optimal joint panel schedule in this case, though looks different, follows a similar logic from the special case. Because the probability of having more than one same-day request is high, the schedule allocates one open slot at the beginning of the session to maximize the physician's utilization. The rest of the optimal schedule is similar to the special case with one extra double-booked slot. The extra double-booked slot is placed closely after the first double-booked slot for both physicians to take advantage of high no-show rate and low waiting cost at the same time.

**Case 4: high no-show rate, moderate waiting cost** ( $\gamma \geq q_0^s + q_1^s$  and  $c_w \left( \frac{1-q_0^s-q_1^s}{q_0^s+q_1^s} \right) \leq c_v \leq c_w \left( \frac{1-q_0^s}{q_0^s} \right)$ ). The optimal capacity planning decisions ( $N_r^{1*}, N_r^{2*}, N_s^{1*}, N_s^{2*}, A^{1*}$  and  $A^{2*}$ ) along with optimal expected utilities ( $U^*$ ) and the optimal joint panel generated using the following parameters  $N = 7, \gamma = 30\%, p = 50, c_l = 50, c_v = 45, c_w = 6, D_r \sim U[0, 7]$  for both physicians, and  $D_s \sim T[0, 6]$  are provided in Table 4.5 and Figure 4.15.

**Table 4.5** Optimal Capacity Planning Decisions - Case 4

|                  | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|------------------|---------|---------|-------|-------|
| Single Physician | 4       | 3       | 3     | 175.9 |
| Joint Panel      | 5       | 2       | 2     | 379.2 |
|                  | 5       | 2       | 2     |       |

Similar to the Case 3, the joint panel schedule allocates more open slots than the single physician schedule due to the high no-show rate.



**Figure 4.15** Optimal schedules: single physician vs. joint panel - case 4.

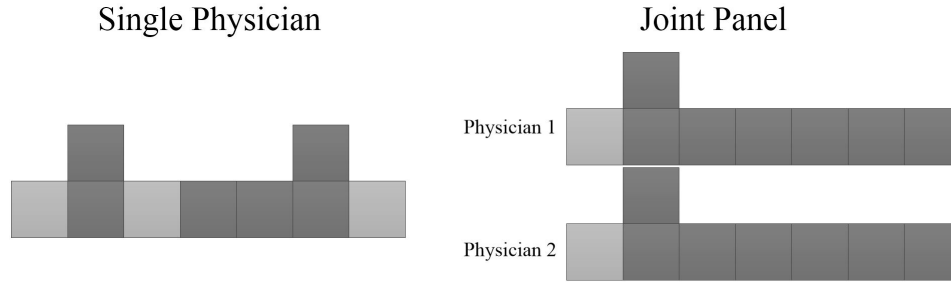
The optimal joint panel schedule, in this case, carries the same structure as the one in the Case 3. However, due to higher waiting cost, the optimal joint schedule pushes the extra double-booked slots toward middle of the session to justify the moderate overall expected waiting time.

**Case 5: high no-show rate, high waiting cost** ( $\gamma \leq q_0^s + q_1^s$ , and  $c_w \left( \frac{1-q_0^s-q_1^s}{q_0^s+q_1^s} \right) \geq c_v$ ). The optimal capacity planning decisions ( $N_r^{1*}$ ,  $N_r^{2*}$ ,  $N_s^{1*}$ ,  $N_s^{2*}$ ,  $A^{1*}$  and  $A^{2*}$ ) along with optimal expected utilities ( $U^*$ ) and the optimal joint panel generated using the following parameters  $N = 7$ ,  $\gamma = 30\%$ ,  $p = 50$ ,  $c_l = 50$ ,  $c_v = 45$ ,  $c_w = 13$ ,  $D_r \sim U[0, 7]$  for both physicians, and  $D_s \sim T[0, 6]$  are provided in Table 4.6 and Figure 4.16.

**Table 4.6** Optimal Capacity Planning Decisions - Case 5

|                  | $N_r^*$ | $N_s^*$ | $A^*$ | $U^*$ |
|------------------|---------|---------|-------|-------|
| Single Physician | 4       | 3       | 3     | 172.9 |
| Joint Panel      | 6       | 1       | 1     | 375.1 |
|                  | 6       | 1       | 1     |       |

Similar to the special case which we discussed in Section 4.2, the optimal joint panel schedule, in this case, assigns only one open slot at the beginning of the session followed by a double-booked slot for both physicians. The rest of the schedule is filled with single-booked routine patients which allows the physician to serve one of the double-booked patients if any of the reserved patients is a no-show.



**Figure 4.16** Optimal schedules: single physician vs. joint panel - case 5.

#### 4.4. Efficiency of Joint Panel Scheduling

In this section, we compare the joint panel scheduling with single physician scheduling to investigate the efficiency of this approach.

The efficiency can be measured in terms of improvement in expected provider's utility ( $U^*$ ). In order to do a fair comparison, we compared the joint panel's expected utility with the sum of expected utilities of two individual physicians. The results show that the joint panel scheduling can improve provider's expected utility up to 7% which is a significant improvement in efficiency because it causes less waiting time for patients as well as higher utilization for provider.

Figure 4.1 shows the difference between both schedules. As one can see, in the joint panel scheduling urgent patients (same-day requests) will be assigned to the first open slot; while in the single physician scheduling each doctor receives his/her own same-day request and it causes long waiting time for the whole system. On the other hand, single physician scheduling may result in lower physician utilization since the physician may stay idle in an open slot because of lack of same-day requests since it should leave more open slots compared to the joint panel schedule. Moreover, in the single physician scheduling some urgent patient may be turned away due to appointment unavailability; however, in the joint panel scheduling, urgent patients have higher chance to get the same-day appointment.

Let  $U_{Joint}^*$  be the optimal expected utility generated using joint panel scheduling, and  $U_{Single}^*$  be the optimal expected profit for single physician scheduling. Therefore,

efficiency improvement ( $\Lambda$ ) can be calculated as follow:

$$\Lambda = \frac{U_{Joint}^* - 2U_{Single}^*}{U_{Joint}^*} \times 100 \quad (4.11)$$

Remarkably, we can claim that the joint panel scheduling not only reduces patients' waiting time and increase physicians' utilization but also provides more accessibility to patients. Results of numerical cases in Section 4.3 demonstrate the efficiency improvement of joint panel scheduling. The summary of results is shown in Table 4.7.

**Table 4.7** Joint Panel Scheduling Efficiency Comparison

|                      | Low No-show Rate ( $\gamma = 8\%$ ) |               |           | High No-show Rate ( $\gamma = 30\%$ ) |               |           |
|----------------------|-------------------------------------|---------------|-----------|---------------------------------------|---------------|-----------|
|                      | $U_{Single}^*$                      | $U_{Joint}^*$ | $\Lambda$ | $U_{Single}^*$                        | $U_{Joint}^*$ | $\Lambda$ |
| Low<br>Waiting Cost  | 236.4                               | 488.0         | 3.1%      | 207.8                                 | 425.5         | 2.4%      |
| High<br>Waiting Cost | 214.4                               | 462.5         | 7.3%      | 193.0                                 | 404.8         | 4.6%      |

According to Table 4.7 joint panel scheduling has higher expected utility in all of the scenarios. However, it is evident that the joint panel scheduling is more effective when the provider expects low no-show rate among routine patients.

#### 4.5. Managerial Insight

The implications of the results from our analysis in Sections 3.2 and 3.3 are summarized in the following.

In the case where it is optimal to double book exactly one patient and leave only one open slot for both physicians, the optimal schedule depends on the no-show rate ( $\gamma$ ), the probability of having one or less same-day request ( $q_0^s$  and  $q_1^s$ ) and the ratio of the waiting cost to the overtime cost ( $\frac{c_w}{c_v}$ ). Our analysis confirm our previous

findings in single physician scheduling (Section 3.4) that if the ratio of the waiting cost to the overtime cost is small, then the double-booked slot should be placed at the beginning of the session for both physicians and if the ratio is large, then the double-booked slot should be placed at the end of the session for both physicians.

Furthermore, the relative position of open slots to the double-booked slots are determined by two major factors: (i) The relationship between the probability of no-show of routine patients and the probability of having one or fewer same-day request and (ii) the ratio of the waiting cost to the overtime cost. If receiving one or less same-day request is more likely than a no-show and the ratio of waiting cost to the overtime cost is low, then the open slot should be placed right after the double-booked slot for both physicians to increase the utilization. In addition, if the ratio is moderate, then the open slot should be placed right after the double-booked slot for one physician and prior to the double-booked slot for the other physicians to enable provider in reduction of overall waiting time and utilization of un-occupied open slot simultaneously. On the other hand, if receiving one or less same-day request is less likely than a no-show and the ratio is small, then the open slot should be placed after the double-booked slot and at the end of the session for both physicians. Also, if the ratio is moderate then the open slot should be placed after the double-booked slot and at the end of the session for both physicians. Finally, if the ratio of the waiting cost to the overtime cost is high, regardless of no-show rate and same-day requests, it is optimal to place the open slot anywhere prior to the double-booked slots.

Comparing joint panel scheduling with two physicians who are using individual schedules, we can point out two major differentiations. (i) The joint panel scheduling assigns fewer open slots for urgent patients. This allows both physicians to serve more routine patients and utilize fewer double-booked slots which results in lower expected waiting time. (ii) The joint panel scheduling results in higher efficiency and

higher expected utility because it not only causes less waiting time for patients but also reduces physician's idle time.



## CHAPTER 5

### CONCLUSION AND FUTURE EXTENSIONS

In this dissertation, we develop an innovative stochastic model combined with a Newsvendor framework that calculates the probability of the number of patients waiting at the end of each appointment slot in order to find the provider's joint optimal capacity allocation and scheduling decisions that maximize the social welfare. The optimal curve-out schedule stipulates the optimal capacity allocation among urgent patients and routine patients, the maximum number of appointment slots that can be double-booked, as well as where are the different types of slots located in the schedule.

We also develop heuristic schedules for single physician scheduling and compare their performances against the one optimal schedule generated from the complete enumeration algorithm. Furthermore, we extend our model to a joint panel to examine the efficiency of such systems and extract guidelines for primary care providers.

We analyze both the single physician scheduling and the joint panel scheduling in a case where it is optimal to double book exactly one patient and leave only one open slot for a same-day request for each physician and characterize the optimal schedule under specific circumstances. Our analysis indicates that the optimal schedule is determined by two factors (i) the ratio of waiting cost to the overtime cost and (ii) the relation between no-show rate and distribution of same-day requests.

In addition, through several numerical illustrations, we demonstrate that in the case of low waiting cost, the provider should accommodate urgent patients by assigning as many open slots as needed. On the other hand, in the case of high waiting cost combined with low no-show rate, the provider should allocate fewer open slots to urgent patients and reserve more capacity for routine patients.

Furthermore, we investigate both single physician scheduling and the joint panel schedule and show that sharing the same-day requests can increase the efficiency of the clinic. Our analysis shows that the joint panel is more effective in the cases where the provider expects low no-show rate and consider a high cost for patients' waiting time.

This dissertation has laid the groundwork for several future research opportunities. We assume that the service time is deterministic and has an equal length for all patients. However, the effect of random service time on optimal schedule can be one direction to expand the following model. Furthermore, we assume that all physicians and patients are punctual while this assumption is not practical and may need to be relaxed. In addition, we did not consider a constraint on overtime hours which can be taken into account in a future extension of the model. Moreover, we assume that the no-show rate is homogenous among routine patients. However, the no-show probability can be different among patients.

## APPENDIX A

### PROOF OF PROPOSITIONS (SINGLE PHYSICIAN)

Since we assumed the optimal schedule has only one open slot and one double-booked slot, then the expected profit and lost sales are all the same in all schedules. Therefore, we compare only the expected cost of schedules.

**Proof of proposition 1.** Let  $TC_{SCH_1^1, j}$  be the expected cost of schedule  $SCH_1^1$  which has the double-booked slot at slot  $j$ .

$$\begin{aligned}
 TC_{SCH_1^1, j} &\leq TC_{SCH_1^1, j+1} \\
 \Rightarrow c_w \sum_{i=j}^N (1-\gamma)^{i-j+2} + c_v(1-\gamma)^{N-j+2} &\leq c_w \sum_{i=j+1}^N (1-\gamma)^{i-j+1} + c_v(1-\gamma)^{N-j+1} \\
 \Rightarrow c_w(1-\gamma)^{N-j+2} \leq \gamma c_v(1-\gamma)^{N-j+1} &\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
 \end{aligned}$$

**Proof of proposition 2.** Let  $TC_{SCH_2^1, j}$  be the expected cost of schedule  $SCH_2^1$  which has the double-booked slot at slot  $j$ .

$$\begin{aligned}
 TC_{SCH_2^1, j} &\leq TC_{SCH_2^1, j+1} \\
 \Rightarrow c_w \left( \sum_{i=1}^{j-1} (1-\gamma)^{i+1} + (1-q_0^s) \sum_{i=j}^N (1-\gamma)^i \right) &+ c_v(1-q_0^s)(1-\gamma)^N \\
 \leq c_w \left( \sum_{i=1}^j (1-\gamma)^{i+1} + (1-q_0^s) \sum_{i=j+1}^N (1-\gamma)^i \right) &+ c_v(1-q_0^s)(1-\gamma)^N \\
 \Rightarrow \sum_{i=1}^{j-1} (1-\gamma)^{i+1} + (1-q_0^s) \sum_{i=j}^N (1-\gamma)^i &\leq \sum_{i=1}^j (1-\gamma)^{i+1} + (1-q_0^s) \sum_{i=j+1}^N (1-\gamma)^i \\
 \Rightarrow \sum_{i=1}^{j-1} (1-\gamma)^{i+1} + (1-q_0^s) \sum_{i=j}^N (1-\gamma)^i &\leq \sum_{i=1}^j (1-\gamma)^{i+1} + (1-q_0^s) \sum_{i=j+1}^N (1-\gamma)^i \\
 \Rightarrow (1-q_0^s)(1-\gamma)^j \leq (1-\gamma)^{j+1} &\Rightarrow q_0^s \geq \gamma
 \end{aligned}$$

**Proof of proposition 3.** Let  $TC_{SCH_{2-1,j}^1}$  be the expected cost of schedule  $SCH_{2-1}^1$  which has the double-booked slot at slot  $j$ .

$$\begin{aligned}
TC_{SCH_{2-1,j}^1} &\leq TC_{SCH_{2-1,j}^1} \\
&\Rightarrow c_w((1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^{N-j+1} (1-\gamma)^i) + c_v(1-q_0^s)(1-\gamma)^{N-j+1} \\
&\leq c_w((1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^{N-j} (1-\gamma)^i) + c_v(1-q_0^s)(1-\gamma)^{N-j} \\
&\Rightarrow c_w(1-q_0^s)(1-\gamma)^{N-j+1} \leq \gamma c_v(1-q_0^s)(1-\gamma)^{N-j} \Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

**Proof of proposition 4.** Let  $TC_{SCH_{2-2,j}^1}$  be the expected cost of schedule  $SCH_{2-2}^1$  which has the double-booked slot at slot  $j$ .

$$\begin{aligned}
TC_{SCH_{2-2,j}^1} &\leq TC_{SCH_{2-2,j}^1} \\
&\Rightarrow c_w \left( \sum_{i=2}^{N-j+1} (1-\gamma)^i + (1-q_0^s)(1-\gamma)^{N-j+1} \right) + c_v(1-q_0^s)(1-\gamma)^{N-j+1} \\
&\leq c_w \left( \sum_{i=2}^{N-j} (1-\gamma)^i + (1-q_0^s)(1-\gamma)^{N-j} \right) + c_v(1-q_0^s)(1-\gamma)^{N-j} \\
&\Rightarrow c_w \leq c_v \frac{\gamma(1-q_0^s)}{(1-\gamma) - \gamma(1-q_0^s)}
\end{aligned}$$

In order to create the presented flowchart, we need to compare the identical schedules which are introduced after each proposition. Lets assume  $q_0^s \geq \gamma$  and

$$c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right).$$

$$\begin{aligned} TC_{SCH_{1-1}^1} &\leq TC_{SCH_{2-1-1}^1} \\ &\Rightarrow c_w \sum_{i=2}^N (1-\gamma)^i + c_v(1-\gamma)^N \\ &\leq c_w \left( (1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^N (1-\gamma)^i \right) + c_v(1-q_0^s)(1-\gamma)^N \\ &\Rightarrow c_v q_0^s (1-\gamma)^N \leq c_w \left( (1-\gamma)^2 - q_0^s \sum_{i=2}^N (1-\gamma)^i \right) \\ &\Rightarrow c_v q_0^s (1-\gamma)^{N-2} \leq c_w \left( 1 - q_0^s \sum_{i=0}^{N-2} (1-\gamma)^i \right) \\ &\Rightarrow c_v \leq c_w \left( \frac{\gamma + q_0^s (1-\gamma)^{N-1}}{\gamma q_0^s (1-\gamma)^{N-2}} - \frac{1}{\gamma (1-\gamma)^{N-2}} \right) \end{aligned}$$

However, we can prove that this inequality is never valid:

$$\begin{aligned} q_0^s \geq \gamma &\Rightarrow \frac{\gamma}{q_0^s} \leq 1 \\ &\Rightarrow \frac{\gamma}{q_0^s} + (1-\gamma)^{N-1} - 1 \leq (1-\gamma)^{N-1} \\ &\Rightarrow \frac{\gamma + q_0^s (1-\gamma)^{N-1}}{\gamma q_0^s (1-\gamma)^{N-2}} - \frac{1}{\gamma (1-\gamma)^{N-2}} \leq \frac{(1-\gamma)^{N-1}}{\gamma (1-\gamma)^{N-2}} \\ &\Rightarrow \frac{\gamma + q_0^s (1-\gamma)^{N-1}}{\gamma q_0^s (1-\gamma)^{N-2}} - \frac{1}{\gamma (1-\gamma)^{N-2}} \leq \frac{1-\gamma}{\gamma} \end{aligned}$$

Therefore, given  $c_w \left( \frac{1-\gamma}{\gamma} \right) \leq c_v$ , the schedule  $SCH_{1-1}^1$  is always worst than  $SCH_{2-1-1}^1$ .

Now lets move the next comparison. Lets assume  $q_0^s \geq \gamma$  and  $c_w \geq c_v \left( \frac{\gamma}{1-\gamma} \right)$ .

$$\begin{aligned} TC_{SCH_{1-2}^1} &\leq TC_{SCH_{2-1-2}^1} \\ &\Rightarrow c_w(1-\gamma)^2 + c_v(1-\gamma)^2 \leq c_w \left( (1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 \right) + c_v(1-q_0^s)(1-\gamma)^2 \\ &\Rightarrow c_v q_0^s (1-\gamma)^2 \leq c_w \left( (1-q_0^s)(1-\gamma)^2 \right) \\ &\Rightarrow c_v \leq c_w \left( \frac{1-q_0^s}{q_0^s} \right) \end{aligned}$$

The next comparison is between schedule  $SCH_{1-1}^1$  and  $SCH_{2-2-1}^1$ . Lets assume  $q_0^s \leq \gamma$  and  $c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)$  also  $c_w \leq c_v \frac{\gamma(1-q_0^s)}{(1-\gamma)-\gamma(1-q_0^s)}$ .

$$\begin{aligned}
TC_{SCH_{1-1}^1} &\leq TC_{SCH_{2-2-1}^1} \\
&\Rightarrow c_w \sum_{i=2}^N (1-\gamma)^i + c_v(1-\gamma)^N \\
&\leq c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N \right) + c_v(1-q_0^s)(1-\gamma)^N \\
&\Rightarrow c_w q_0^s (1-\gamma)^N \leq c_w ((1-q_0^s)(1-\gamma)^N) \\
&\Rightarrow c_v \left( \frac{q_0^s}{1-q_0^s} \right) \leq c_w
\end{aligned}$$

Finally, lets assume  $q_0^s \leq \gamma$  and  $c_w \geq c_v \left( \frac{\gamma}{1-\gamma} \right)$  also  $c_w \leq c_v \frac{\gamma(1-q_0^s)}{(1-\gamma)-\gamma(1-q_0^s)}$ .

$$\begin{aligned}
TC_{SCH_{1-2}^1} &\leq TC_{SCH_{2-2-1}^1} \\
&\Rightarrow c_w(1-\gamma)^2 + c_v(1-\gamma)^2 \\
&\leq c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N \right) + c_v(1-q_0^s)(1-\gamma)^N \\
&\Rightarrow c_v ((1-\gamma)^2 - (1-q_0^s)(1-\gamma)^N) \leq c_w \left( \sum_{i=3}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N \right) \\
&\Rightarrow c_v \left( \frac{(1-\gamma)^2 - (1-q_0^s)(1-\gamma)^N}{\sum_{i=3}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N} \right) \leq c_w
\end{aligned}$$

However, we can prove that this inequality is always satisfied and consequently, schedule  $SCH_{1-2}^1$  is always better than  $SCH_{2-2-1}^1$ .

$$\begin{aligned}
q_0^s \leq \gamma &\Rightarrow \frac{q_0^s}{1 - q_0^s} \leq \frac{\gamma}{1 - \gamma} \\
&\Rightarrow \gamma(1 - q_0^s)(1 - \gamma)^{N-3} \leq q_0^s(1 - \gamma)^{N-2} \\
&\Rightarrow -(1 - \gamma)^{N-2} + \gamma(1 - q_0^s)(1 - \gamma)^{N-3} \leq -(1 - q_0^s)(1 - \gamma)^{N-2} \\
&\Rightarrow \gamma \left( \frac{1 - (1 - \gamma)^{N-2}}{\gamma} \right) + \gamma(1 - q_0^s)(1 - \gamma)^{N-3} \leq 1 - (1 - q_0^s)(1 - \gamma)^{N-2} \\
&\Rightarrow \gamma \sum_{i=0}^{N-3} (1 - \gamma)^i + \gamma(1 - q_0^s)(1 - \gamma)^{N-3} \leq 1 - (1 - q_0^s)(1 - \gamma)^{N-2} \\
&\Rightarrow \gamma \sum_{i=3}^N (1 - \gamma)^i + \gamma(1 - q_0^s)(1 - \gamma)^N \leq 1 - (1 - q_0^s)(1 - \gamma)^{N+1} \\
&\Rightarrow \gamma \sum_{i=3}^N (1 - \gamma)^i + \gamma(1 - q_0^s)(1 - \gamma)^N \leq (1 - \gamma)^3 - (1 - q_0^s)(1 - \gamma)^{N+1} \\
&\Rightarrow \frac{\gamma}{1 - \gamma} \leq \frac{(1 - \gamma)^2 - (1 - q_0^s)(1 - \gamma)^N}{\sum_{i=3}^N (1 - \gamma)^i + (1 - q_0^s)(1 - \gamma)^N}
\end{aligned}$$

## APPENDIX B

### PROOF OF PROPOSITIONS (JOINT PANEL)

**Proof of proposition 5.** Let  $U(T_{O,D}^{i,j} + T_{O,D}^{i,h})$  be the expected cost of  $T_{O,D} + T_{O,D}$  which has the double-booked slot at slot  $j$  for first physician and slot  $h$  for the second physician and  $i \leq j$ .

$$\begin{aligned}
U(T_{O,D}^{i,j} + T_{O,D}^{i,h}) &\leq U(T_{O,D}^{i,j+1} + T_{O,D}^{i,h+1}) \\
&\Rightarrow c_w \left( \sum_{i=2}^{N-j+2} (1-\gamma)^i + \sum_{i=2}^{N-h+2} (1-\gamma)^i \right) + c_v \left( (1-\gamma)^{N-j+2} + (1-\gamma)^{N-h+2} \right) \\
&\leq c_w \left( \sum_{i=2}^{N-j+1} (1-\gamma)^i + \sum_{i=2}^{N-h+1} (1-\gamma)^i \right) + c_v \left( (1-\gamma)^{N-j+1} + (1-\gamma)^{N-h+1} \right) \\
&\Rightarrow c_w \left( (1-\gamma)^{N-j+2} + (1-\gamma)^{N-h+2} \right) \\
&\leq c_v \left( (1-\gamma)^{N-j+1} + (1-\gamma)^{N-h+1} - (1-\gamma)^{N-j+2} - (1-\gamma)^{N-h+2} \right) \\
&\Rightarrow c_w \leq c_v \left( \frac{(1-\gamma)^{N-j+1} + (1-\gamma)^{N-h+1}}{(1-\gamma)^{N-j+2} + (1-\gamma)^{N-h+2}} - 1 \right) \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{O,D}^{i,j} + T_{O,D}^{i,h}) &\leq U(T_{O,D}^{i,j+1} + T_{O,D}^{i,h-1}) \\
&\Rightarrow c_w \left( \sum_{i=2}^{N-j+2} (1-\gamma)^i + \sum_{i=2}^{N-h+2} (1-\gamma)^i \right) + c_v \left( (1-\gamma)^{N-j+2} + (1-\gamma)^{N-h+2} \right) \\
&\leq c_w \left( \sum_{i=2}^{N-j+1} (1-\gamma)^i + \sum_{i=2}^{N-h+3} (1-\gamma)^i \right) + c_v \left( (1-\gamma)^{N-j+1} + (1-\gamma)^{N-h+3} \right) \\
&\Rightarrow c_w \left( (1-\gamma)^{N-j+2} - (1-\gamma)^{N-h+3} \right) \leq c_v \gamma \left( (1-\gamma)^{N-j+1} - (1-\gamma)^{N-h+2} \right) \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$



$$\begin{aligned}
U(T_{O,D}^{i,N} + T_{O,D}^{i,N}) &\leq U(T_{O,D}^{i,2} + T_{O,D}^{i,N}) \\
&\Rightarrow 2c_w(1-\gamma)^2 + 2c_v(1-\gamma)^2 \\
&\leq c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-\gamma)^2 \right) + c_v((1-\gamma)^N + (1-\gamma)^2) \\
&\Rightarrow c_w \left( (1-\gamma)^2 - \sum_{i=2}^N (1-\gamma)^i \right) \leq c_v \left( (1-\gamma)^N - (1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( 1 - \sum_{i=0}^{N-2} (1-\gamma)^i \right) \leq c_v \left( (1-\gamma)^{N-2} - 1 \right) \Rightarrow c_w \geq c_v \left( \frac{1 - (1-\gamma)^{N-2}}{\sum_{i=1}^{N-2} (1-\gamma)^i} \right) \\
&\Rightarrow c_w \geq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

**Proof of proposition 6.** Let  $U(T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1})$  be the expected cost of the schedule which has the double-booked slot at slot  $j$  and an open slot at slot  $j+1$  for first physician and a double-booked slot at  $h$  and open slot at  $h+1$  for the second physician.

$$\begin{aligned}
U(T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1}) &\leq U(T_{D,O}^{j,j+1} + T_{D,O}^{h,h+2}) \\
&\Rightarrow \left( 2(1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^{N-j+1} (1-\gamma)^i + (1-q_0^s - q_1^s) \sum_{i=2}^{N-h+1} (1-\gamma)^i \right) \\
&\leq \left( 2(1-\gamma)^2 + (1-\gamma)^3 + (1-q_0^s) \sum_{i=2}^{N-j+1} (1-\gamma)^i + (1-q_0^s - q_1^s) \sum_{i=3}^{N-h+1} (1-\gamma)^i \right) \\
&\Rightarrow \left( (1-q_0^s - q_1^s) \sum_{i=2}^{N-h+1} (1-\gamma)^i \right) \leq \left( (1-\gamma)^3 + (1-q_0^s - q_1^s) \sum_{i=3}^{N-h+1} (1-\gamma)^i \right) \\
&\Rightarrow (1-q_0^s - q_1^s) \leq (1-\gamma) \\
&\Rightarrow \gamma \leq q_0^s + q_1^s
\end{aligned}$$

**Proof of proposition 7.** Let  $U(T_{D,O}^{j,j+1} + T_{D,O}^{h,h+1})$  be the expected cost of the schedule which has the double-booked slot at slot  $j$  for first physician and slot  $h$  for

the second physician.

$$\begin{aligned}
U(T_{D,O}^{1,2} + T_{D,O}^{1,2}) &\leq U(T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}) \\
&\Rightarrow c_w \left( 2(1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( (1-q_0^s) \sum_{i=1}^{N-2} (1-\gamma)^i + (1-q_0^s - q_1^s) \sum_{i=1}^{N-2} (1-\gamma)^i \right) \\
&+ c_v \left( (2-2q_0^s - q_1^s)(1-\gamma)^{N-2} \right) \leq c_v(2-2q_0^s - q_1^s) \\
&\Rightarrow c_w \left( (1-q_0^s)(1 - (1-\gamma)^{N-1}) + (1-q_0^s - q_1^s)(1 - (1-\gamma)^{N-1}) - \gamma(2-2q_0^s - q_1^s) \right) \\
&\leq c_v \gamma \left( (2-2q_0^s - q_1^s)(1 - (1-\gamma)^{N-2}) \right) \\
&\Rightarrow c_w(1-\gamma - (1-\gamma)^{N-1}) \leq c_v \gamma(1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w(1-\gamma) - \gamma c_v \leq (1-\gamma)^{N-2} (c_w(1-\gamma) - \gamma c_v)
\end{aligned}$$

Therefore, if  $c_w \leq (\frac{\gamma}{1-\gamma})c_v$  schedule  $T_{D,O}^{1,2} + T_{D,O}^{1,2}$  is better. Otherwise, schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$  is better.

$$\begin{aligned}
U(T_{D,O}^{1,2} + T_{D,O}^{1,2}) &\leq U(T_{D,O}^{1,2} + T_{D,O}^{N-1,N}) \\
&\Rightarrow c_w \left( 2(1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\quad + c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\leq c_w \left( (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( \sum_{i=1}^{N-2} (1-\gamma)^i \right) + c_v \left( (1-\gamma)^{N-2} \right) \leq c_v \\
&\Rightarrow c_w(1 - (1-\gamma)^{N-1} - \gamma) \leq \gamma c_v(1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w(1-\gamma) - \gamma c_v \leq (c_w(1-\gamma) - \gamma c_v)(1-\gamma)^{N-2}
\end{aligned}$$

Therefore, if  $c_w \leq \frac{\gamma}{1-\gamma}c_v$  schedule  $T_{D,O}^{1,2} + T_{D,O}^{1,2}$  is better. Otherwise, schedule  $T_{D,O}^{1,2} + T_{D,O}^{N-1,N}$  is better.

$$\begin{aligned}
U(T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}) &\leq U(T_{D,O}^{1,2} + T_{D,O}^{N-1,N}) \\
&\Rightarrow c_w (2(1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
&\quad + c_v ((1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\quad + c_v ((1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
&\Rightarrow c_v ((1-q_0^s)(1-\gamma)^2) \leq c_w \left( (1-q_0^s) \sum_{i=3}^N (1-\gamma)^i \right) \\
&\quad + c_v ((1-q_0^s)(1-\gamma)^N) \\
&\Rightarrow c_v \leq c_w \sum_{i=1}^{N-2} (1-\gamma)^i + c_v(1-\gamma)^{N-2} \\
&\Rightarrow \gamma c_v - (1-\gamma)c_w \leq (\gamma c_v - (1-\gamma)c_w)(1-\gamma)^{N-2}
\end{aligned}$$

Therefore, if  $c_w \geq (\frac{\gamma}{1-\gamma})c_v$  schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$  is better. Otherwise, schedule  $T_{D,O}^{1,2} + T_{D,O}^{N-1,N}$  is better. Consequently,  $T_{D,O}^{1,2} + T_{D,O}^{N-1,N}$  is never outperform  $T_{D,O}^{1,2} + T_{D,O}^{1,2}$  and  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$ .

**Proof of proposition 8.** Let  $U(T_{D,O}^{j,N} + T_{D,O}^{h,N})$  be the expected cost of the schedule which has the double-booked slot at slot  $j$  for first physician and slot  $h$  for

the second physician.

$$\begin{aligned}
U(T_{D,O}^{1,N} + T_{D,O}^{1,N}) &\leq U(T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}) \\
&\Rightarrow c_w \left( 2 \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( 2 \sum_{i=0}^{N-2} (1-\gamma)^i + (1-q_0^s)(1-\gamma)^{N-2} + (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^{N-2} + (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \right) \\
&\leq c_w (2 + (2 - 2q_0^s - q_1^s)) + c_v ((2 - 2q_0^s - q_1^s)) \\
&\Rightarrow c_w (2(1 - (1-\gamma)^{N-1}) + \gamma(2 - 2q_0^s - q_1^s)(1-\gamma)^{N-2} - 2\gamma - \gamma(2 - 2q_0^s - q_1^s)) \\
&\leq c_w \gamma (2 - 2q_0^s - q_1^s) (1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w (2(1-\gamma)(1 - (1-\gamma)^{N-2}) - \gamma(2 - 2q_0^s - q_1^s)(1 - (1-\gamma)^{N-2})) \\
&\leq c_w \gamma (2 - 2q_0^s - q_1^s) (1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w (2(1-\gamma) - \gamma(2 - 2q_0^s - q_1^s)) \leq c_w \gamma (2 - 2q_0^s - q_1^s) \\
&\Rightarrow c_w \left( \frac{2(1-\gamma)}{\gamma(2 - 2q_0^s - q_1^s)} - 1 \right) \leq c_v
\end{aligned}$$

Therefore, if  $c_v \left( \frac{\gamma(1-q_0^s-0.5q_1^s)}{(1-\gamma)-\gamma(1-1q_0^s-0.5q_1^s)} \right) \geq c_w$  then schedule  $T_{D,O}^{1,N} + T_{D,O}^{1,N}$  is better. Otherwise, schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$  is better.

$$\begin{aligned}
U(T_{D,O}^{1,N} + T_{D,O}^{1,N}) &\leq U(T_{D,O}^{1,N} + T_{D,O}^{N-1,N}) \\
&\Rightarrow c_w \left( 2 \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N + (1-q_0^s-q_1^s)(1-\gamma)^N \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s-q_1^s)(1-\gamma)^N \right) \\
&\leq c_w \left( (1-\gamma)^2 + (1-q_0^s-q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s-q_1^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( \sum_{i=0}^{N-2} (1-\gamma)^i + (1-q_0^s-q_1^s)(1-\gamma)^{N-2} \right) \\
&+ c_v \left( (1-q_0^s-q_1^s)(1-\gamma)^{N-2} \right) \\
&\leq c_w(2-q_0^s-q_1^s) + c_v \left( (1-q_0^s-q_1^s) \right) \\
&\Rightarrow c_w \left( (1-(1-\gamma)^{N-1} + \gamma(1-q_0^s-q_1^s)(1-\gamma)^{N-2} - \gamma(2-q_0^s-q_1^s)) \right) \\
&\leq c_v \gamma \left( (1-q_0^s-q_1^s)(1-(1-\gamma)^{N-2}) \right) \\
&\Rightarrow c_w \left( (1-\gamma)(1-(1-\gamma)^{N-2} - \gamma(1-q_0^s-q_1^s)(1-(1-\gamma)^{N-2})) \right) \\
&\leq c_v \gamma \left( (1-q_0^s-q_1^s)(1-(1-\gamma)^{N-2}) \right) \\
&\Rightarrow c_w \left( (1-\gamma) - \gamma(1-q_0^s-q_1^s) \right) \leq c_v \gamma \left( (1-q_0^s-q_1^s) \right) \\
&\Rightarrow c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s-q_1^s)} - 1 \right) \leq c_v
\end{aligned}$$

Therefore, if  $c_v \left( \frac{\gamma(1-q_0^s-q_1^s)}{(1-\gamma)-\gamma(1-q_0^s-q_1^s)} \right) \geq c_w$  then schedule  $T_{D,O}^{1,N}+T_{D,O}^{1,N}$  is better. Otherwise, schedule  $T_{D,O}^{1,N} + T_{D,O}^{N-1,N}$  is better.

$$\begin{aligned}
U(T_{D,O}^{1,N} + T_{D,O}^{1,N}) &\leq U(T_{D,O}^{N-1,N} + T_{D,O}^{1,N}) \\
&\Rightarrow c_w \left( 2 \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N + (1-q_0^s-q_1^s)(1-\gamma)^N \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s-q_1^s)(1-\gamma)^N \right) \\
&\leq c_w \left( (1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s-q_1^s)(1-\gamma)^N \right) \\
&+ c_v \left( (1-q_0^s-q_1^s)(1-\gamma)^N + (1-q_0^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( \sum_{i=0}^{N-2} (1-\gamma)^i + (1-q_0^s)(1-\gamma)^{N-2} \right) \\
&+ c_v \left( (1-q_0^s)(1-\gamma)^{N-2} \right) \\
&\leq c_w (1 + (1-q_0^s)) + c_v ((1-q_0^s)) \\
&\Rightarrow c_w \left( (1 - (1-\gamma)^{N-1}) + \gamma(1-q_0^s)(1-\gamma)^{N-2} - \gamma - \gamma(1-q_0^s) \right) \\
&\leq c_v \gamma (1-q_0^s) (1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w \left( (1-\gamma)(1 - (1-\gamma)^{N-2}) - (\gamma(1-q_0^s))(1 - (1-\gamma)^{N-2}) \right) \\
&\leq c_v \gamma (1-q_0^s) (1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w \left( (1-\gamma) - (\gamma(1-q_0^s)) \right) \leq c_v \gamma (1-q_0^s) \\
&\Rightarrow c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s)} - 1 \right) \leq c_v
\end{aligned}$$

Therefore, if  $c_v \left( \frac{\gamma(1-q_0^s)}{(1-\gamma)-\gamma(1-q_0^s)} \right) \geq c_w$  then schedule  $T_{D,O}^{1,N} + T_{D,O}^{1,N}$  is better. Otherwise, schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{1,N}$  is better.

$$\begin{aligned}
& U(T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}) \leq U(T_{D,O}^{1,N} + T_{D,O}^{N-1,N}) \\
& \Rightarrow c_w (2(1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
& + c_v ((1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
& \leq c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N \right) \\
& + c_v ((1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
& \Rightarrow c_w ((1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2) + c_v ((1-q_0^s)(1-\gamma)^2) \\
& \leq c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N \right) + c_v ((1-q_0^s)(1-\gamma)^N) \\
& \Rightarrow c_w \left( 1 + (1-q_0^s) - \sum_{i=0}^{N-2} (1-\gamma)^i - (1-q_0^s)(1-\gamma)^{N-2} \right) \\
& \leq c_v(1-q_0^s)((1-\gamma)^{N-2} - 1) \\
& \Rightarrow c_w ((1-\gamma)((1-\gamma)^{N-2} - 1) - \gamma(1-q_0^s)((1-\gamma)^{N-2} - 1)) \\
& \leq c_v \gamma(1-q_0^s)((1-\gamma)^{N-2} - 1) \\
& \Rightarrow c_w ((1-\gamma) - \gamma(1-q_0^s)) \geq c_v \gamma(1-q_0^s) \\
& \Rightarrow c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s)} - 1 \right) \geq c_v
\end{aligned}$$



Therefore, if  $c_v \left( \frac{\gamma(1-q_0^s)}{(1-\gamma)-\gamma(1-q_0^s)} \right) \geq c_w$  then schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$  is better. Otherwise, schedule  $T_{D,O}^{1,N} + T_{D,O}^{N-1,N}$  is better.

$$\begin{aligned}
& U(T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}) \leq U(T_{D,O}^{N-1,N} + T_{D,O}^{1,N}) \\
& \Rightarrow c_w (2(1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
& + c_v ((1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) \\
& \leq c_w \left( (1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
& + c_v ((1-q_0^s - q_1^s)(1-\gamma)^N + (1-q_0^s)(1-\gamma)^2) \\
& \Rightarrow c_w ((1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) + c_v ((1-q_0^s - q_1^s)(1-\gamma)^2) \\
& \leq c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N \right) + c_v ((1-q_0^s - q_1^s)(1-\gamma)^N) \\
& \Rightarrow c_w \left( 1 + (1-q_0^s - q_1^s) - \sum_{i=0}^{N-2} (1-\gamma)^i - (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \right) \\
& \leq c_v (1 - q_0^s - q_1^s) ((1-\gamma)^{N-2} - 1) \\
& \Rightarrow c_w (\gamma + \gamma(1 - q_0^s - q_1^s) - (1 - (1-\gamma)^{N-1}) - \gamma(1 - q_0^s - q_1^s)(1-\gamma)^{N-2}) \\
& \leq c_v \gamma (1 - q_0^s - q_1^s) ((1-\gamma)^{N-2} - 1) \\
& \Rightarrow c_w ((1-\gamma)((1-\gamma)^{N-1} - 1) - \gamma(1 - q_0^s - q_1^s)((1-\gamma)^{N-2} - 1)) \\
& \leq c_v \gamma (1 - q_0^s - q_1^s) ((1-\gamma)^{N-2} - 1) \\
& \Rightarrow c_w ((1-\gamma) - \gamma(1 - q_0^s - q_1^s)) \geq c_v \gamma (1 - q_0^s - q_1^s) \\
& \Rightarrow c_w \left( \frac{1-\gamma}{\gamma(1 - q_0^s - q_1^s)} - 1 \right) \geq c_v
\end{aligned}$$

Therefore, if  $c_v \left( \frac{\gamma(1-q_0^s - q_1^s)}{(1-\gamma) - \gamma(1-q_0^s - q_1^s)} \right) \leq c_w$  then schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{N-1,N}$  is better. Otherwise, schedule  $T_{D,O}^{N-1,N} + T_{D,O}^{1,N}$  is better.

$$\begin{aligned}
U(T_{D,O}^{1,N} + T_{D,O}^{N-1,N}) &\leq U(T_{D,O}^{N-1,N} + T_{D,O}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s)(1-\gamma)^N \right) \\
&\quad + c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\leq c_w \left( (1-\gamma)^2 + (1-q_0^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-q_0^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-q_0^s)(1-\gamma)^N \right) \\
&\quad + c_v \left( (1-q_0^s)(1-\gamma)^N + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\leq c_w \left( (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-q_0^s)(1-\gamma)^2 \right) \\
&\Rightarrow (c_w + c_v) \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-q_0^s)(1-\gamma)^N \right) \\
&\leq (c_w + c_v) \left( (1-q_0^s)(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\Rightarrow (1-q_0^s - q_1^s) + (1-q_0^s)(1-\gamma)^{N-2} \leq (1-q_0^s) + (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \\
&\Rightarrow (1-q_0^s - q_1^s)(1 - (1-\gamma)^{N-2}) \leq (1-q_0^s)(1 - (1-\gamma)^{N-2}) \\
&\Rightarrow q_1^s \geq 0
\end{aligned}$$

Therefore, all the time schedule  $T_{D,O}^{1,N} + T_{D,O}^{N-1,N}$  is better than  $T_{D,O}^{N-1,N} + T_{D,O}^{1,N}$ .

**Proof of proposition 10.** Let  $U(T_{D,O}^{j,j+1} + T_{O,D}^{h-1,h})$  be the expected cost of the schedule which has the double-booked slot at slot  $j$  for first physician and slot  $h$  for

the second physician.

$$\begin{aligned}
& T_{D,O}^{j \geq h-2, j+1} + T_{O,D}^{h-1, h} \leq T_{D,O}^{j \leq h-2, j+1} + T_{O,D}^{h-1, h} \\
& \Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s) \sum_{i=2}^{N-j+1} (1-\gamma)^i + \sum_{i=2}^{N-h+2} (1-\gamma)^i \right) \\
& + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^{N-j+1} + (1-\gamma)^{N-h+2} \right) \\
& \leq c_w \left( (1-\gamma)^2 + (1-q_0^s) \sum_{i=2}^{N-j+1} (1-\gamma)^i + \sum_{i=2}^{N-h+2} (1-\gamma)^i \right) \\
& + c_v \left( (1-q_0^s)(1-\gamma)^{N-j+1} + (1-\gamma)^{N-h+2} \right) \\
& \Rightarrow -c_w \sum_{i=2}^{N-j+1} (1-\gamma)^i \leq c_v (1-\gamma)^{N-j+1}
\end{aligned}$$

**Proof of proposition 11.** Let  $U(T_{D,O}^{j, j+1} + T_{O,D}^{h-1, h})$  be the expected cost of the schedule which has the double-booked slot at slot  $j$  for first physician and slot  $h$  for

the second physician.

$$\begin{aligned}
U(T_{D,O}^{1,2} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{N-1,N} + T_{O,D}^{1,2}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^N \right) \\
&\leq c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^N \right) \\
&\Rightarrow c_w \left( (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N \right) \leq c_w \left( (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \sum_{i=2}^N (1-\gamma)^i + c_v (1-\gamma)^N \leq c_w (1-\gamma)^2 + c_v (1-\gamma)^2 \\
&\Rightarrow c_w \sum_{i=0}^{N-2} (1-\gamma)^i + c_v (1-\gamma)^{N-2} \leq (c_w + c_v) \\
&\Rightarrow c_w (1 - (1-\gamma)^{N-1}) + c_v \gamma (1-\gamma)^{N-2} \leq \gamma (c_w + c_v) \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{1,2} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{1,2} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^N \right) \\
&\leq c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i + (1-\gamma)^2 \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( \sum_{i=2}^N (1-\gamma)^i \right) + c_v \left( (1-\gamma)^N \right) \\
&\leq c_w (1-\gamma)^2 + c_v (1-\gamma)^2 \\
&\Rightarrow c_w \left( \sum_{i=0}^{N-2} (1-\gamma)^i \right) + c_v \left( (1-\gamma)^{N-2} \right) \leq c_w + c_v \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{1,2} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{N-1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^N \right) \leq \\
&c_w \left( 2(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \sum_{i=2}^N (1-\gamma)^i + c_v (1-\gamma)^N \leq c_w (1-\gamma)^2 + c_v (1-\gamma)^2 \\
&\Rightarrow c_w \sum_{i=0}^{N-2} (1-\gamma)^i + c_v (1-\gamma)^{N-2} \leq c_w + c_v \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{N-1,N} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{1,2} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^N \right) \\
&\leq c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i + (1-\gamma)^2 \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( (1-q_0^s - q_1^s) + \sum_{i=0}^{N-2} (1-\gamma)^i \right) + c_v \left( (1-q_0^s - q_1^s) + (1-\gamma)^{N-2} \right) \\
&\leq c_w \left( (1-q_0^s - q_1^s) \sum_{i=0}^{N-2} (1-\gamma)^i + 1 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^{N-2} + 1 \right) \\
&\Rightarrow c_w \left( \gamma(1-q_0^s - q_1^s) + (1-(1-\gamma)^{N-1}) \right) + c_v \gamma \left( (1-q_0^s - q_1^s) + (1-\gamma)^{N-2} \right) \\
&\leq c_w \left( (1-q_0^s - q_1^s)(1-(1-\gamma)^{N-1}) + \gamma \right) + c_v \gamma \left( (1-q_0^s - q_1^s)(1-\gamma)^{N-2} + 1 \right) \\
&\Rightarrow c_w \left( \gamma(1-q_0^s - q_1^s) + (1-(1-\gamma)^{N-1}) - (1-q_0^s - q_1^s)(1-(1-\gamma)^{N-1}) - \gamma \right) \\
&\leq c_v \gamma \left( (1-q_0^s - q_1^s)(1-\gamma)^{N-2} + 1 - (1-q_0^s - q_1^s) - (1-\gamma)^{N-2} \right) \\
&\Rightarrow c_w (q_0^s + q_1^s) (1-\gamma - (1-\gamma)^{N-1}) \leq c_v \gamma (q_0^s + q_1^s) (1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{N-1,N} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{N-1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^N \right) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \sum_{i=0}^{N-2} (1-\gamma)^i + c_v (1-\gamma)^{N-2} \leq c_w + c_v \\
&\Rightarrow c_w (1 - (1-\gamma)^{N-1}) + c_v \gamma (1-\gamma)^{N-2} \leq (c_w + c_v) \gamma \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{1,2} + T_{O,D}^{1,N}) &\leq U(T_{D,O}^{N-1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s) \sum_{i=2}^N (1-\gamma)^i + (1-\gamma)^2 \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \sum_{i=2}^N (1-\gamma)^i + c_v (1-\gamma)^N \leq c_w (1-\gamma)^2 + c_v (1-\gamma)^2 \\
&\Rightarrow c_w \sum_{i=0}^{N-2} (1-\gamma)^i + c_v (1-\gamma)^{N-2} \leq c_w + c_v \\
&\Rightarrow c_w (1 - (1-\gamma)^{N-1}) + c_v \gamma (1-\gamma)^{N-2} \leq (c_w + c_v) \gamma \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

**Proof of proposition 12.** Let  $U(T_{D,O}^{j,N} + T_{O,D}^{h-1,h})$  be the expected cost of the schedule which has the double-booked slot at slot  $j$  for first physician and slot  $h$  for the second

physician.

$$\begin{aligned}
U(T_{D,O}^{1,N} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{N-1,N} + T_{O,D}^{1,2}) \\
&\Rightarrow c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N + \sum_{i=2}^N (1-\gamma)^i \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^N \right) \leq \\
c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^N \right) \\
&\Rightarrow c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&+ c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N \right) \\
&\leq c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( \sum_{i=1}^{N-2} (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \right) \\
&\leq (c_w + c_v)(1-q_0^s - q_1^s) \\
&\Rightarrow c_w(1 - (1-\gamma)^{N-1} - \gamma + (1-q_0^s - q_1^s)\gamma(1-\gamma)^{N-2}) + c_v\gamma(1-q_0^s - q_1^s)(1-\gamma)^{N-2} \\
&\leq (c_w + c_v)\gamma(1-q_0^s - q_1^s) \\
&\Rightarrow c_w(1 - (1-\gamma)^{N-1} - \gamma + (1-q_0^s - q_1^s)\gamma(1-\gamma)^{N-2} - \gamma(1-q_0^s - q_1^s)) \\
&\leq c_v\gamma(1-q_0^s - q_1^s)(1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w(1 - \gamma - \gamma(1-q_0^s - q_1^s)) \leq c_v\gamma(1-q_0^s - q_1^s) \\
&\Rightarrow c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s - q_1^s)} - 1 \right) \leq c_v
\end{aligned}$$



$$\begin{aligned}
U(T_{D,O}^{1,N} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{N-1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^N \right) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( 2 \sum_{i=0}^{N-2} (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \right) + c_v (1-\gamma)^{N-2} ((1-q_0^s - q_1^s) + 1) \\
&\leq c_w (1-q_0^s - q_1^s + 2) + c_v ((1-q_0^s - q_1^s) + 1) \\
&\Rightarrow c_w \left( 2(1 - (1-\gamma)^{N-1}) - 2\gamma + (1-q_0^s - q_1^s)\gamma(1-\gamma)^{N-2} - \gamma(1-q_0^s - q_1^s) \right) \\
&\leq c_w \gamma ((1-q_0^s - q_1^s) + 1)(1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w \left( \frac{2-\gamma}{\gamma(2-q_0^s - q_1^s)} - 1 \right) \leq c_v
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{1,N} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^N \right) \\
&\leq c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \sum_{i=2}^N (1-\gamma)^i + c_v (1-\gamma)^N \leq c_w (1-\gamma)^2 + c_v (1-\gamma)^2 \\
&\Rightarrow c_w \sum_{i=0}^{N-2} (1-\gamma)^i + c_v (1-\gamma)^{N-2} \leq c_w + c_v \\
&\Rightarrow c_w (1 - (1-\gamma)^{N-1}) + c_v \gamma (1-\gamma)^{N-2} \leq (c_w + c_v) \gamma \\
&\Rightarrow c_w (1 - \gamma - (1-\gamma)^{N-1}) \leq c_v \gamma (1 - (1-\gamma)^{N-2}) \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma}{1-\gamma} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{N-1,N} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{N-1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^N \right) \\
&\leq c_w \left( 2(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \sum_{i=0}^{N-2} (1-\gamma)^i + c_v (1-\gamma)^{N-2} \leq c_w + c_v \\
&\Rightarrow c_w (1 - (1-\gamma)^{N-1}) + c_v \gamma (1-\gamma)^{N-2} \leq (c_w + c_v) \gamma \\
&\Rightarrow c_w \leq c_v \left( \frac{\gamma (1 - (1-\gamma)^{N-2})}{(1-\gamma - (1-\gamma)^{N-1})} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{N-1,N} + T_{O,D}^{1,2}) &\leq U(T_{D,O}^{1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w \left( (1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2 + \sum_{i=2}^N (1-\gamma)^i \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^N \right) \leq \\
&c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\quad + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\Rightarrow c_w \left( (1-q_0^s - q_1^s)(1-\gamma)^2 \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^N \right) \\
&\leq c_w \left( (1-q_0^s - q_1^s)(1-\gamma)^N \right) + c_v \left( (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\Rightarrow c_w (1 - q_0^s - q_1^s) (1 - (1-\gamma)^{N-2}) \leq c_v (1 - (1-\gamma)^{N-2}) (q_0^s + q_1^s) \\
&\Rightarrow c_w \leq c_v \left( \frac{q_0^s + q_1^s}{1 - q_0^s - q_1^s} \right)
\end{aligned}$$

$$\begin{aligned}
U(T_{D,O}^{N-1,N} + T_{O,D}^{1,N}) &\leq U(T_{D,O}^{1,N} + T_{O,D}^{1,N}) \\
&\Rightarrow c_w (2(1-\gamma)^2 + (1-q_0^s - q_1^s)(1-\gamma)^2) + c_v ((1-q_0^s - q_1^s)(1-\gamma)^2 + (1-\gamma)^2) \\
&\leq c_w \left( \sum_{i=2}^N (1-\gamma)^i + (1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2 \right) \\
&\quad + c_v ((1-q_0^s - q_1^s)(1-\gamma)^N + (1-\gamma)^2) \\
&\Rightarrow c_w \left( 1 + (1-q_0^s - q_1^s) - \sum_{i=0}^{N-2} (1-\gamma)^i - (1-q_0^s - q_1^s)(1-\gamma)^{N-2} \right) \\
&\leq c_w (1-q_0^s - q_1^s)((1-\gamma)^{N-2} - 1) \\
&\Rightarrow c_w (\gamma + \gamma(1-q_0^s - q_1^s) - (1-(1-\gamma)^{N-1}) - (1-q_0^s - q_1^s)\gamma(1-\gamma)^{N-2}) \\
&\leq c_w \gamma (1-q_0^s - q_1^s)((1-\gamma)^{N-2} - 1) \\
&\Rightarrow c_w (1-\gamma - \gamma(1-q_0^s - q_1^s)) \leq c_v \gamma (1-q_0^s - q_1^s) \\
&\Rightarrow c_w \left( \frac{1-\gamma}{\gamma(1-q_0^s - q_1^s)} - 1 \right) \geq c_v
\end{aligned}$$

## BIBLIOGRAPHY

- [Adan and Vissers, 2002] Adan, I. and Vissers, J. (2002). Patient mix optimisation in hospital admission planning: a case study. *International Journal of Operations and Production Management*, 22(4):445–461.
- [Aharonson-Daniel et al., 1996] Aharonson-Daniel, L., Paul, R. J., and Hedley, A. J. (1996). Management of queues in out-patient departments: the use of computer simulation. *Journal of Management in Medicine*, 10(6):50–58.
- [Almog et al., 2003] Almog, D. M., Devries, J. A., Borrelli, J. A., and Kopycka-Kedzierawski, D. T. (2003). The reduction of broken appointment rates through an automated appointment confirmation system. *Journal of Dental Education*, 67(9):1016–1022.
- [Armstrong et al., 2004] Armstrong, B., Levesque, O., Perlin, J. B., Rick, C., and Schechtman, G. (2004). Reinventing veterans health administration: focus on primary care. *Journal of Healthcare Management/American College of Healthcare Executives*, 50(6):399–408.
- [Babes and Sarma, 1991] Babes, M. and Sarma, G. (1991). Out-patient queues at the ibn-rochd health centre. *Journal of the Operational Research Society*, 42(10):845–855.
- [Bailey, 1952] Bailey, N. T. (1952). A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1:185–199.
- [Balasubramanian et al., 2007] Balasubramanian, H., Banerjee, R., Gregg, M., and Denton, B. T. (2007). Improving primary care access using simulation optimization. In *2007 Winter Simulation Conference*, pages 1494–1500.
- [Ballard and Kuhl, 2006] Ballard, S. M. and Kuhl, M. E. (2006). The use of simulation to determine maximum capacity in the surgical suite operating room. In *Proceedings of the 2006 Winter Simulation Conference*, pages 433–438. IEEE.
- [Barnes et al., 1995] Barnes, D., Eribes, C., Juarbe, T., Nelson, M., Proctor, S., Sawyer, L., Shaul, M., and Meleis, A. I. (1995). Primary health care and primary care: A confusion of philosophies. *Nursing Outlook*, 43(1):7–16.
- [Bean and Talaga, 1994] Bean, A. and Talaga, J. (1994). Predicting appointment breaking. *Journal of Health Care Marketing*, 15(1):29–34.
- [Bean and Talaga, 1992] Bean, A. G. and Talaga, J. (1992). Appointment breaking: causes and solutions. *Journal of Health Care Marketing*, 12(4):14–25.

- [Beliën et al., 2006] Beliën, J., Demeulemeester, E., and Cardoen, B. (2006). Visualizing the demand for various resources as a function of the master surgery schedule: A case study. *Journal of Medical Systems*, 30(5):343–350.
- [Benjamin-Bauman et al., 1984] Benjamin-Bauman, J., Reiss, M. L., and Bailey, J. S. (1984). Increasing appointment keeping by reducing the call-appointment interval. *Journal of Applied Behavior Analysis*, 17(3):295–301.
- [Berg et al., 2013] Berg, B. P., Murr, M., Chermak, D., Woodall, J., Pignone, M., Sandler, R. S., and Denton, B. T. (2013). Estimating the cost of no-shows and evaluating the effects of mitigation strategies. *Medical Decision Making*, 33(8):976–985.
- [Birchall et al., 1983] Birchall, J., Cox, T., and Wong, H. (1983). Computer modelling of ent out-patients. *Clinical Otolaryngology and Allied Sciences*, 8(6):411–415.
- [Bloom, 2002] Bloom, B. S. (2002). Crossing the quality chasm: a new health system for the 21st century. *The Journal of the American Medical Association*, 287(5):646–647.
- [Bosch and Dietz, 2000] Bosch, P. M. V. and Dietz, D. C. (2000). Minimizing expected waiting in a medical appointment system. *IIE Transactions*, 32(9):841–848.
- [Bourdais et al., 2003] Bourdais, S., Galinier, P., and Pesant, G. (2003). Hibiscus: A constraint programming application to staff scheduling in health care. In *International Conference on Principles and Practice of Constraint Programming*, pages 153–167. Springer.
- [Brahimi and Worthington, 1991] Brahimi, M. and Worthington, D. (1991). Queueing models for out-patient appointment systems a case study. *Journal of the Operational Research Society*, 42(9):733–746.
- [Bundy et al., 2005] Bundy, D. G., Randolph, G. D., Murray, M., Anderson, J., and Margolis, P. A. (2005). Open access in primary care: results of a north carolina pilot project. *Pediatrics*, 116(1):82–87.
- [Cardoen and Demeulemeester, 2008] Cardoen, B. and Demeulemeester, E. (2008). Capacity of clinical pathways a strategic multi-level evaluation tool. *Journal of Medical Systems*, 32(6):443–452.
- [Cardoen et al., 2010] Cardoen, B., Demeulemeester, E., and Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201(3):921–932.
- [Cayirli and Gunes, 2014] Cayirli, T. and Gunes, E. D. (2014). Outpatient appointment scheduling in presence of seasonal walk-ins. *Journal of the Operational Research Society*, 65(4):512–531.

- [Cayirli et al., 2012] Cayirli, T., Yang, K. K., and Quek, S. A. (2012). A universal appointment rule in the presence of no-shows and walk-ins. *Production and Operations Management*, 21(4):682–697.
- [Cayirli, 2003] Cayirli, T., V. E. (2003). Outpatient scheduling in health care: a review of literature. *Production and Operations Management*, 12(4):519–549.
- [Centeno et al., 2003] Centeno, M. A., Giachetti, R., Linn, R., and Ismail, A. M. (2003). Emergency departments ii: A simulation-ilp based tool for scheduling er staff. In *Proceedings of the 35th Conference on Winter Simulation: Driving Innovation*, WSC '03, pages 1930–1938.
- [Chakraborty et al., 2013] Chakraborty, S., Muthuraman, K., and Lawley, M. (2013). Sequential clinical scheduling with patient no-show: The impact of pre-defined slot structures. *Socio-Economic Planning Sciences*, 47(3):205–219.
- [Chen et al., 2010] Chen, B.-l., Li, E.-d., Yamawuchi, K., Kato, K., Naganawa, S., and Miao, W.-j. (2010). Impact of adjustment measures on reducing outpatient waiting time in a community hospital: application of a computer simulation. *Chinese Medical Journal (English Edition)*, 123(5):574.
- [Cox et al., 1985] Cox, T. F., Birchall, J. P., and Wong, H. (1985). Optimising the queuing system for an ear, nose and throat outpatient clinic. *Journal of Applied Statistics*, 12(2):113–126.
- [Davis et al., 2014] Davis, K., Stremikis, K., Schoen, C., and Squires, D. (2014). Mirror, mirror on the wall, 2014 update: How the performance of the us health care system compares internationally. *CommonWealth Fund, New York*, 16.
- [Davis et al., 2010] Davis, K., Stremikis, K., Squires, D., and Schoen, C. (2010). Mirror, mirror on the wall, how the performance of the us health care system compares internationally. *CommonWealth Fund, New York*.
- [Denton and Gupta, 2003] Denton, B. and Gupta, D. (2003). A sequential bounding approach for optimal appointment scheduling. *IIE Transactions*, 35(11):1003–1016.
- [Dexter and Macario, 2002] Dexter, F. and Macario, A. (2002). Changing allocations of operating room time from a system based on historical utilization to one where the aim is to schedule as many surgical cases as possible. *Anesthesia and Analgesia*, 94(5):1272–1279.
- [Dobson et al., 2011] Dobson, G., Hasija, S., and Pinker, E. J. (2011). Reserving capacity for urgent patients in primary care. *Production and Operations Management*, 20(3):456–473.
- [Dove and Schneider, 1981] Dove, H. G. and Schneider, K. C. (1981). The usefulness of patients' individual characteristics in predicting no-shows in outpatient clinics. *Medical Care*, 19(7):734–740.

- [Downer et al., 2005] Downer, S. R., Meara, J. G., and Da Costa, A. C. (2005). Use of sms text messaging to improve outpatient attendance. *Medical Journal of Australia*, 183(7):366.
- [Dowland, 1998] Dowland, K. A. (1998). Nurse scheduling with tabu search and strategic oscillation. *European Journal of Operational Research*, 106(2):393–407.
- [Dreier et al., 2007] Dreier, J., Froimovici, M., Bibi, Y., Vardy, D. A., Cicurel, A., and Cohen, A. D. (2007). Nonattendance in obstetrics and gynecology patients. *Gynecologic and Obstetric Investigation*, 66(1):40–43.
- [DuMontier et al., 2013] DuMontier, C., Rindfleisch, K., Pruszynski, J., and Frey, J. (2013). A multi-method intervention to reduce no shows in an urban residency clinic. *Family Medicine*, 45(9):634–641.
- [Elmendorf, 2014] Elmendorf, D. W. (2014). The 2014 long-term budget outlook. Congressional Budget Office (US Congress), Washington, DC.
- [Erdogan et al., 2015] Erdogan, S. A., Gose, A., and Denton, B. T. (2015). Online appointment sequencing and scheduling. *IIE Transactions*, 47(11):1267–1286.
- [Exadaktylos et al., 2008] Exadaktylos, A. K., Evangelopoulos, D. S., Wullschleger, M., Bürki, L., and Zimmermann, H. (2008). Strategic emergency department design: An approach to capacity planning in healthcare provision in overcrowded emergency rooms. *Journal of Trauma Management and Outcomes*, 2(1):1.
- [Fetter and Thompson, 1966] Fetter, R. B. and Thompson, J. D. (1966). Patients’ waiting time and doctors’ idle time in the outpatient setting. *Health Services Research*, 1(1):66.
- [Fomundam and Herrmann, 2007] Fomundam, S. and Herrmann, J. W. (2007). A survey of queuing theory applications in healthcare. Technical Report 24, ISR Technical Report, University of Maryland, College Park, MD.
- [Fone et al., 2003] Fone, D., Hollinghurst, S., Temple, M., Round, A., Lester, N., Weightman, A., Roberts, K., Coyle, E., Bevan, G., and Palmer, S. (2003). Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health*, 25(4):325–335.
- [Fries and Marathe, 1981] Fries, B. E. and Marathe, V. P. (1981). Determination of optimal variable-sized multiple-block appointment systems. *Operations Research*, 29(2):324–345.
- [Gallucci et al., 2005] Gallucci, G., Swartz, W., and Hackerman, F. (2005). Brief reports: Impact of the wait for an initial appointment on the rate of kept appointments at a mental health center. *Psychiatric Services*, 56(3):344–346.

- [Geraghty et al., 2008] Geraghty, M., Glynn, F., Amin, M., and Kinsella, J. (2008). Patient mobile telephone textreminder: a novel way to reduce non-attendance at the ent out-patient clinic. *The Journal of Laryngology and Otolology*, 122(03):296–298.
- [Giachetti et al., 2005] Giachetti, R. E., Centeno, E. A., Centeno, M. A., and Sundaram, R. (2005). Assessing the viability of an open access policy in an outpatient clinic: A discrete-event and continuous simulation modeling approach. In *Proceedings of the 37th Conference on Winter Simulation, WSC '05*, pages 2246–2255.
- [Goitein, 1990] Goitein, M. (1990). Waiting patiently. *The New England Journal of Medicine*, 323(9):604–608.
- [Goldman et al., 1982] Goldman, L., Freidin, R., Cook, E. F., Eigner, J., and Grich, P. (1982). A multivariate approach to the prediction of no-show behavior in a primary care center. *Archives of Internal Medicine*, 142(3):563–567.
- [Green, 2004] Green, L. V. (2004). *Capacity planning and management in hospitals*, pages 15–41. Springer, Boston, MA.
- [Green, 2010] Green, L. V. (2010). *Using Queueing Theory to Alleviate Emergency Department Overcrowding*. Wiley.
- [Green and Savin, 2008] Green, L. V. and Savin, S. (2008). Reducing delays for medical appointments: A queueing approach. *Operations Research*, 56(6):1526–1538.
- [Green et al., 2007] Green, L. V., Savin, S., and Murray, M. (2007). Providing timely access to care: what is the right patient panel size? *Joint Commission Journal on Quality and Patient Safety*, 33(4):211–218.
- [Güenal and Pidd, 2010] Güenal, M. M. and Pidd, M. (2010). Discrete event simulation for performance modelling in health care: a review of the literature. *Journal of Simulation*, 4(1):42–51.
- [Gupta and Denton, 2007] Gupta, D. and Denton, B. (2007). Appointment scheduling in health care: Challenges and opportunities. *IIE transactions*, 40(9):800–819.
- [Gupta and Wang, 2008] Gupta, D. and Wang, L. (2008). Revenue management for a primary-care clinic in the presence of patient choice. *Operations Research*, 56(3):576–592.
- [Hahn-Goldberg et al., 2012] Hahn-Goldberg, S., Carter, M. W., and Beck, J. C. (2012). Dynamic template scheduling to address uncertainty in complex scheduling problems: a case study on chemotherapy outpatient scheduling. In *Society for Health Systems Conference, Las Vegas, NV*.



- [Harper and Gamlin, 2003] Harper, P. R. and Gamlin, H. M. (2003). Reduced outpatient waiting times with improved appointment scheduling: a simulation modelling approach. *Operation Research Spectrum*, 25(2):207–222.
- [Hassin and Mendel, 2008] Hassin, R. and Mendel, S. (2008). Scheduling arrivals to queues: A single-server model with no-shows. *Management Science*, 54(3):565–572.
- [Herriott, 1999] Herriott, S. (1999). Reducing delays and waiting times with open-office scheduling. *Family Practice Management*, 6:38–43.
- [Hick et al., 2004] Hick, J. L., Hanfling, D., Burstein, J. L., DeAtley, C., Barbisch, D., Bogdan, G. M., and Cantrill, S. (2004). Health care facility and community strategies for patient care surge capacity. *Annals of Emergency Medicine*, 44(3):253–261.
- [Hixon et al., 1999] Hixon, A. L., Chapman, R. W., and Nuovo, J. (1999). Failure to keep clinic appointments: implications for residency education and productivity. *Family Medicine*, 31:627–630.
- [Ho and Lau, 1992] Ho, C.-J. and Lau, H.-S. (1992). Minimizing total cost in scheduling outpatient appointments. *Management Science*, 38(12):1750–1764.
- [Huang and Zuniga, 2012] Huang, Y. and Zuniga, P. (2012). Dynamic overbooking scheduling system to improve patient access. *Journal of the Operational Research Society*, 63(6):810–820.
- [Izard, 2005] Izard, T. (2005). Managing the habitual no-show patient. *Family Practice Management*, 12(2):65–66.
- [Jansson, 1966] Jansson, B. (1966). Choosing a good appointment system: a study of queues of the type (d, m, 1). *Operations Research*, 14(2):292–312.
- [Jaumard et al., 1998] Jaumard, B., Semet, F., and Vovor, T. (1998). A generalized linear programming model for nurse scheduling. *European Journal of Operational Research*, 107(1):1–18.
- [Kaandorp and Koole, 2007] Kaandorp, G. C. and Koole, G. (2007). Optimal outpatient appointment scheduling. *Health Care Management Science*, 10(3):217–229.
- [Kim and Giachetti, 2006] Kim, S. and Giachetti, R. E. (2006). A stochastic mathematical appointment overbooking model for healthcare providers to improve profits. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 36(6):1211–1219.
- [Klassen and Rohleder, 1996] Klassen, K. J. and Rohleder, T. R. (1996). Scheduling outpatient appointments in a dynamic environment. *Journal of Operations Management*, 14(2):83–101.

- [Klassen and Rohleder, 2004] Klassen, K. J. and Rohleder, T. R. (2004). Outpatient appointment scheduling with urgent clients in a dynamic, multi-period environment. *International Journal of Service Industry Management*, 15(2):167–186.
- [Koeleman and Koole, 2012] Koeleman, P. M. and Koole, G. M. (2012). Optimal outpatient appointment scheduling with emergency arrivals and general service times. *IIE Transactions on Healthcare Systems Engineering*, 2(1):14–30.
- [Kopach et al., 2007] Kopach, R., DeLaurentis, P.-C., Lawley, M., Muthuraman, K., Ozsen, L., Rardin, R., Wan, H., Intrevado, P., Qu, X., and Willis, D. (2007). Effects of clinical characteristics on successful open access scheduling. *Health Care Management Science*, 10(2):111–124.
- [Lacy et al., 2004] Lacy, N. L., Paulman, A., Reuter, M. D., and Lovejoy, B. (2004). Why we dont come: patient perceptions on no-shows. *The Annals of Family Medicine*, 2(6):541–545.
- [LaGanga and Lawrence, 2007] LaGanga, L. R. and Lawrence, S. R. (2007). Clinic overbooking to improve patient access and increase provider productivity\*. *Decision Sciences*, 38(2):251–276.
- [LaGanga and Lawrence, 2012] LaGanga, L. R. and Lawrence, S. R. (2012). Appointment overbooking in health care clinics to improve patient service and clinic performance. *Production and Operations Management*, 21(5):874–888.
- [Lau and Lau, 2000] Lau, H.-S. and Lau, A. H.-L. (2000). A fast procedure for computing the total system cost of an appointment schedule for medical and kindred facilities. *IIE Transactions*, 32(9):833–839.
- [Lee et al., 2005] Lee, V. J., Earnest, A., Chen, M. I., and Krishnan, B. (2005). Predictors of failed attendances in a multi-specialty outpatient centre using electronic databases. *BMC Health Services Research*, 5(1):51.
- [Lehaney et al., 1999] Lehaney, B., Clarke, S., and Paul, R. (1999). A case of an intervention in an outpatients department. *Journal of the Operational Research Society*, 50(9):877–891.
- [Lesaca, 1995] Lesaca, T. (1995). Assessing the influence of a no-show fee on patient compliance at a cmhc. *Administration and Policy in Mental Health and Mental Health Services Research*, 22(6):629–631.
- [Lindley, 1952] Lindley, D. V. (1952). *The theory of queues with a single server*, volume 48, pages 277–289. Cambridge University Press, New York, NY.
- [Liu and Liu, 1998] Liu, L. and Liu, X. (1998). Block appointment systems for outpatient clinics with multiple doctors. *Journal of the Operational Research Society*, 49(12):1254–1259.

- [Liu et al., 2010] Liu, N., Ziya, S., and Kulkarni, V. G. (2010). Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. *Manufacturing and Service Operations Management*, 12(2):347–364.
- [Lowes, 2005] Lowes, R. (2005). Practice pointers: How to handle no-shows. *Medical Economics*, 82(8):62–65.
- [Luo et al., 2012] Luo, J., Kulkarni, V. G., and Ziya, S. (2012). Appointment scheduling under patient no-shows and service interruptions. *Manufacturing and Service Operations Management*, 14(4):670–684.
- [Macharia et al., 1992] Macharia, W. M., Leon, G., Rowe, B. H., Stephenson, B. J., and Haynes, R. B. (1992). An overview of interventions to improve compliance with appointment keeping for medical services. *Journal of American Medical Association*, 267(13):1813–1817.
- [Mak et al., 2014] Mak, H.-Y., Rong, Y., and Zhang, J. (2014). Appointment scheduling with limited distributional information. *Management Science*, 61(2):316–334.
- [Mäntyjärvi, 1994] Mäntyjärvi, M. (1994). No-show patients in an ophthalmological out-patient department. *Acta ophthalmologica*, 72(3):284–289.
- [Mardiah and Basri, 2013] Mardiah, F. P. and Basri, M. H. (2013). The analysis of appointment system to reduce outpatient waiting time at indonesia's public hospital. *Human Resource Management Research*, 3(1):27–33.
- [Mehrotra et al., 2008] Mehrotra, A., Keehl-Markowitz, L., and Ayanian, J. Z. (2008). Implementing open-access scheduling of visits in primary care practices: a cautionary tale. *Annals of Internal Medicine*, 148(12):915–922.
- [Millen and Shonick, 1978] Millen, R. and Shonick, W. (1978). Comment on "an approach to reducing the adverse effects of broken appointments in primary care systems". *Medical Care*, 16(6):528–530.
- [Mitchell et al., 2007] Mitchell, A. J., Psych, M., and Selmes, T. (2007). A comparative survey of missed initial and follow-up appointments to psychiatric specialties in the united kingdom. *Psychiatric Services*, 58(6):868–871.
- [Murdock et al., 2002] Murdock, A., Rodgers, C., Lindsay, H., and Tham, T. (2002). Why do patients not keep their appointments? prospective study in a gastroenterology outpatient clinic. *Journal of the Royal Society of Medicine*, 95(6):284–286.
- [Murray and Berwick, 2003] Murray, M. and Berwick, D. M. (2003). Advanced access: reducing waiting and delays in primary care. *Jama*, 289(8):1035–1040.
- [Murray et al., 2003] Murray, M., Bodenheimer, T., Rittenhouse, D., and Grumbach, K. (2003). Improving timely access to primary care: case studies of the advanced access model. *Jama*, 289(8):1042–1046.

- [Murray and Tantau, 1999] Murray, M. and Tantau, C. (1999). Redefining open access to primary care. *Managed Care Quarterly*, 7:45–55.
- [Murray and Tantau, 2000] Murray, M. and Tantau, C. (2000). Same-day appointments: exploding the access paradigm. *Family Practice Management*, 7(8):45–45.
- [Muthuraman and Lawley, 2008] Muthuraman, K. and Lawley, M. (2008). A stochastic overbooking model for outpatient clinical scheduling with no-shows. *IIE Transactions*, 40(9):820–837.
- [Neal et al., 2001] Neal, R. D., Lawlor, D. A., Allgar, V., Colledge, M., Ali, S., Hassey, A., Portz, C., and Wilson, A. (2001). Missed appointments in general practice: retrospective data analysis from four practices. *British Journal of General Practice*, 51(471):830–832.
- [Neeleman and Mikhail, 1997] Neeleman, W. and Mikhail, J. (1997). A case-control study of gp and patient-related variables associated with non-attendance at new psychiatric out-patient appointments. *Journal of Mental Health*, 6(3):301–306.
- [Nolte and McKee, 2008] Nolte, E. and McKee, C. M. (2008). Measuring the health of nations: updating an earlier analysis. *Health Affairs*, 27(1):58–71.
- [Norman and Conner, 1996] Norman, P. and Conner, M. (1996). Predicting health-check attendance among prior attenders and nonattenders: The role of prior behavior in the theory of planned behavior. *Journal of Applied Social Psychology*, 26(11):1010–1026.
- [Norris et al., 2014] Norris, J. B., Kumar, C., Chand, S., Moskowitz, H., Shade, S. A., and Willis, D. R. (2014). An empirical investigation into factors affecting patient cancellations and no-shows at outpatient clinics. *Decision Support Systems*, 57:428–443.
- [OConnor et al., 2006] OConnor, M. E., Matthews, B. S., and Gao, D. (2006). Effect of open access scheduling on missed appointments, immunizations, and continuity of care for infant well-child care visits. *Archives of pediatrics and adolescent medicine*, 160(9):889–893.
- [OHare and Corlett, 2004] OHare, C. D. and Corlett, J. (2004). The outcomes of open-access scheduling. *Family Practice Management*, 11(2):35–8.
- [Parikh et al., 2010] Parikh, A., Gupta, K., Wilson, A. C., Fields, K., Cosgrove, N. M., and Kostis, J. B. (2010). The effectiveness of outpatient appointment reminder systems in reducing no-show rates. *The American Journal of Medicine*, 123(6):542–548.

- [Parker et al., 2012] Parker, M. M., Moffet, H. H., Schillinger, D., Adler, N., Fernandez, A., Ciechanowski, P., and Karter, A. J. (2012). Ethnic differences in appointment-keeping and implications for the patient-centered medical home findings from the diabetes study of northern california (distance). *Health Services Research*, 47(2):572–593.
- [Patrick, 2012] Patrick, J. (2012). A markov decision model for determining optimal outpatient scheduling. *Health Care Management Science*, 15(2):91–102.
- [Patrick et al., 2008] Patrick, J., Puterman, M. L., and Queyranne, M. (2008). Dynamic multipriority patient scheduling for a diagnostic resource. *Operations Research*, 56(6):1507–1525.
- [Paul, 1995] Paul, R. (1995). Outpatient clinics. *Operation Research Insight*, 8(2):24–27.
- [Perron et al., 2010] Perron, N. J., Dao, M. D., Kossovsky, M. P., Miserez, V., Chuard, C., Calmy, A., and Gaspoz, J.-M. (2010). Reduction of missed appointments at an urban primary care clinic: a randomised controlled study. *British Medical Journal; Family Practice*, 11(1):1.
- [Petterson et al., 2015] Petterson, S. M., Liaw, W. R., Tran, C., and Bazemore, A. W. (2015). Estimating the residency expansion required to avoid projected primary care physician shortages by 2035. *The Annals of Family Medicine*, 13(2):107–114.
- [Phan and Brown, 2009] Phan, K. and Brown, S. R. (2009). Decreased continuity in a residency clinic: a consequence of open access scheduling. *Family Medicine*, 41(1):46–50.
- [Pierdon et al., 2004] Pierdon, S., Charles, T., McKinley, K., and Myers, L. (2004). Implementing advanced access in a group practice network. *Family Practice Management*, 11(5):35–39.
- [PWC, 2013] PWC (2013). Medical cost trend: behind the numbers. Technical Report 8, Health Research Institute, PricewaterhouseCoopers LLP, Dallas, Texas.
- [Qu et al., 2011] Qu, X., Rardin, R. L., and Williams, J. A. S. (2011). Single versus hybrid time horizons for open access scheduling. *Computers and Industrial Engineering*, 60(1):56–65.
- [Qu et al., 2012] Qu, X., Rardin, R. L., and Williams, J. A. S. (2012). A mean–variance model to optimize the fixed versus open appointment percentages in open access scheduling systems. *Decision Support Systems*, 53(3):554–564.
- [Qu et al., 2007] Qu, X., Rardin, R. L., Williams, J. A. S., and Willis, D. R. (2007). Matching daily healthcare provider capacity to demand in advanced access scheduling systems. *European Journal of Operational Research*, 183(2):812–826.

- [Qu and Shi, 2011] Qu, X. and Shi, J. (2011). Modeling the effect of patient choice on the performance of open access scheduling. *International Journal of Production Economics*, 129(2):314–327.
- [Rising et al., 1973] Rising, E. J., Baron, R., and Averill, B. (1973). A systems analysis of a university-health-service outpatient clinic. *Operations Research*, 21(5):1030–1047.
- [Robinson and Chen, 2003] Robinson, L. W. and Chen, R. R. (2003). Scheduling doctors’ appointments: optimal and empirically-based heuristic policies. *IIE Transactions*, 35(3):295–307.
- [Robinson and Chen, 2010] Robinson, L. W. and Chen, R. R. (2010). A comparison of traditional and open-access policies for appointment scheduling. *Manufacturing and Service Operations Management*, 12(2):330–346.
- [Rohrer et al., 2007] Rohrer, J. E., Bernard, M., Naessens, J., Furst, J., Kircher, K., and Adamson, S. (2007). Impact of open-access scheduling on realized access. *Health Services Management Research*, 20(2):134–139.
- [Rosenthal, 2013] Rosenthal, E. (2013). Health cares road to ruin. Technical Report SR1, The New York Times, New York, NY.
- [Rust et al., 1995] Rust, C. T., Gallups, N. H., Clark, W. S., Jones, D. S., and Wilcox, W. D. (1995). Patient appointment failures in pediatric resident continuity clinics. *Archives of Pediatrics and Adolescent Medicine*, 149(6):693–695.
- [Samorani and LaGanga, 2011] Samorani, M. and LaGanga, L. (2011). A stochastic programming approach to improve overbooking in clinic appointment scheduling. In *POMS Annual Meeting Proceedings*, Reno, Nevada.
- [Samorani and LaGanga, 2015] Samorani, M. and LaGanga, L. R. (2015). Outpatient appointment scheduling given individual day-dependent no-show predictions. *European Journal of Operational Research*, 240(1):245–257.
- [Sandy et al., 2009] Sandy, L. G., Bodenheimer, T., Pawlson, L. G., and Starfield, B. (2009). The political economy of us primary care. *Health Affairs*, 28(4):1136–1145.
- [Schectman et al., 2008] Schectman, J. M., Schorling, J. B., and Voss, J. D. (2008). Appointment adherence and disparities in outcomes among patients with diabetes. *Journal of General Internal Medicine*, 23(10):1685–1687.
- [Schoen et al., 2006] Schoen, C., Davis, K., How, S. K., and Schoenbaum, S. C. (2006). Us health system performance: a national scorecard. *Health Affairs*, 25(6):w457–w475.

- [Schoen et al., 2009] Schoen, C., Osborn, R., Doty, M. M., Squires, D., Peugh, J., and Applebaum, S. (2009). A survey of primary care physicians in eleven countries, 2009: perspectives on care, costs, and experiences. *Health Affairs*, 28(6):w1171–w1183.
- [Sepulveda et al., 2008] Sepulveda, M.-J., Bodenheimer, T., and Grundy, P. (2008). Primary care: Can it solve employers health care dilemma? *Health Affairs*, 27(1):151–158.
- [Shonick and Klein, 1977] Shonick, W. and Klein, B. (1977). An approach to reducing the adverse effects of broken appointments in primary care systems: development of a decision rule based on estimated conditional probabilities. *Medical Care*, 15(5):419–429.
- [Siferd and Benton, 1992] Siferd, S. P. and Benton, W. (1992). Workforce staffing and scheduling: Hospital nursing specific models. *European Journal of Operational Research*, 60(3):233–246.
- [Smith and Yawn, 1994] Smith, C. M. and Yawn, B. P. (1994). Factors associated with appointment keeping in a family practice residency clinic. *The Journal of Family Practice*, 38(1):25–29.
- [Smith et al., 2013] Smith, M., Saunders, R., Stuckhardt, L., and McGinnis, J. M. (2013). *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. The National Academies Press, Washington, DC.
- [Soriano, 1966] Soriano, A. (1966). Comparison of two scheduling systems. *Operations Research*, 14(3):388–397.
- [Starfield, 1998] Starfield, B. (1998). *Primary care: balancing health needs, services, and technology*. Oxford University Press, Cary, NC, USA.
- [Starfield and Shi, 2002] Starfield, B. and Shi, L. (2002). Policy relevant determinants of health: an international perspective. *Health Policy*, 60(3):201–218.
- [Swisher et al., 2001] Swisher, J. R., Jacobson, S. H., Jun, J. B., and Balci, O. (2001). Modeling and analyzing a physician clinic environment using discrete-event (visual) simulation. *Computers and Operations Research*, 28(2):105–125.
- [Tai and Williams, 2012] Tai, G. and Williams, P. (2012). Optimization of scheduling patient appointments in clinics using a novel modelling technique of patient arrival. *Computer Methods and Programs in Biomedicine*, 108(2):467–476.
- [Taylor, 1984] Taylor, B. (1984). Patient use of a mixed appointment system in an urban practice. *British Medical Journal*, 289(6454):1277–1278.
- [Taylor and Templeton, 1980] Taylor, I. and Templeton, J. (1980). Waiting time in a multi-server cutoff-priority queue, and its application to an urban ambulance service. *Operations Research*, 28(5):1168–1188.

- [Tsai and Teng, 2014] Tsai, P.-F. J. and Teng, G.-Y. (2014). A stochastic appointment scheduling system on multiple resources with dynamic call-in sequence and patient no-shows for an outpatient clinic. *European Journal of Operational Research*, 239(2):427–436.
- [Van Houdenhoven et al., 2007] Van Houdenhoven, M., Hans, E. W., Klein, J., Wullink, G., and Kazemier, G. (2007). A norm utilisation for scarce hospital resources: evidence from operating rooms in a dutch university hospital. *Journal of Medical Systems*, 31(4):231–236.
- [VanBerkel and Blake, 2007] VanBerkel, P. T. and Blake, J. T. (2007). A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health Care Management Science*, 10(4):373–385.
- [Vasanawala and Desser, 2005] Vasanawala, S. S. and Desser, T. S. (2005). Accommodation of requests for emergency us and ct: Applications of queueing theory to scheduling of urgent studies 1. *Radiology*, 235(1):244–249.
- [Virji, 1990] Virji, A. (1990). A study of patients attending without appointments in an urban general practice. *British Medical Journal*, 301(6742):22–26.
- [Vissers, 1979] Vissers, J. (1979). Selecting a suitable appointment system in an outpatient setting. *Medical Care*, 1:1207–1220.
- [Vissers and Wijngaard, 1979] Vissers, J. and Wijngaard, J. (1979). The outpatient appointment system: Design of a simulation study. *European Journal of Operational Research*, 3(6):459–463.
- [Wang and Gupta, 2011] Wang, W.-Y. and Gupta, D. (2011). Adaptive appointment systems with patient preferences. *Manufacturing and Service Operations Management*, 13(3):373–389.
- [Weisner et al., 2001] Weisner, C., Mertens, J., Tam, T., and Moore, C. (2001). Factors affecting the initiation of substance abuse treatment in managed care. *Addiction*, 96(5):705–716.
- [Welch and Bailey, 1952] Welch, J. and Bailey, N. J. (1952). Appointment systems in hospital outpatient departments. *The Lancet*, 259(6718):1105–1108.
- [White and Pike, 1964] White, M. B. and Pike, M. (1964). Appointment systems in out-patients’ clinics and the effect of patients’ unpunctuality. *Medical Care*, 1:133–145.
- [Williams et al., 2014] Williams, K. A., Chambers, C. G., Dada, M., McLeod, J. C., and Ulatowski, J. A. (2014). Patient punctuality and clinic performance: observations from an academic-based private practice pain centre: a prospective quality improvement study. *British Medical Journal*, 4(5):e004679.



- [Zacharias and Armony, 2016] Zacharias, C. and Armony, M. (2016). Joint panel sizing and appointment scheduling in outpatient care. *Management Science*.
- [Zeng et al., 2010] Zeng, B., Turkcan, A., Lin, J., and Lawley, M. (2010). Clinic scheduling models with overbooking for patients with heterogeneous no-show probabilities. *Annals of Operations Research*, 178(1):121–144.
- [Zhu et al., 2012] Zhu, Z., Heng, B. H., and Teow, K. L. (2012). Analysis of factors causing long patient waiting time and clinic overtime in outpatient clinics. *Journal of Medical Systems*, 36(2):707–713.