PRIORITIZATION IN SERVICE SYSTEMS WITH NONLINEAR DELAY COSTS

Huiyin Ouyang

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research.

Chapel Hill
2016

Approved by:

Nilay T. Argon

Vidyadhar G. Kulkarni

Haipeng Shen

Nur Sunar

Serhan Ziya

ABSTRACT

Huiyin Ouyang: Prioritization in Service Systems with Nonlinear Delay Costs
(Under the direction of Nilay T. Argon and Serhan Ziya)

Prioritization is a common strategy in service systems to improve the overall system performance. In this dissertation, our main objective is to study how priority should be assigned in different systems when the cost of waiting is not linear as is typically assumed in the literature. We study three problems motivated by prioritization decisions in service systems. In the first problem, we consider a single server queueing system with two different types of customers. Each customer incurs a cost depending on its type and waiting time in the queue, and the waiting cost functions are assumed to be nonlinear in the waiting time. We identify the best static policy under different conditions. In the second problem, we consider a stylized, discrete-time model for an Intensive Care Unit (ICU) in which patients' health conditions change over time according to a Markov chain. Our objective is to allocate the ICU beds to minimize the long-run average mortality rate. In the third problem, we consider a multi-server queueing system with impatient customers. Customers are assumed to be in one of two different stages, which can change over time. A reward is obtained at each service completion depending on the stage of the customer. Our objective is to maximize the total discounted reward and the long-run average reward over infinite horizon.

To my daughter Linxi Sun

ACKNOWLEDGEMENTS

I would like to express my special appreciation and thanks to my advisors, Dr. Nilay Argon and Dr. Serhan Ziya, who have been encouraging me and helping me during my study, and providing valuable advice for my research and career. I would also like to thank all my committee members, Dr. Vidyadhar Kulkarni, Dr Nur Sunar, and Dr. Haipeng Shen for serving in my committee and their comments and suggestions on my work.

Finally, I would like to express special thanks to my husband Zhankun Sun and my daughter XiXi, for their support and always being there for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In many service systems customers are heterogeneous in that they have different service requirements and incur different costs. An interesting problem to the service provider is how to control the service process by prioritizing service, and admitting or rejecting customers so that some performance measure is optimized. For example, the director of an emergency department (ED) at a hospital prioritizes the treatment of more severe patients to minimize the overall mortality rate. A call center manager reserves several lines to serve more valued customers since these customers will incur larger cost if they are unsatisfied. Hospital bed managers face the decision of how to allocate beds when the demand exceeds their capacity, e.g., they either discharge an existing patient to admit a new patient, or they reject the new patient. All these service systems provide services to different types of customers, and their decisions are associated with prioritizing some customers over others.

There is a vast literature that investigates models when the cost of holding customers in the system are linear in their waiting time, and the optimal policies for these models are proved to be a priority index rule. However, the costs in reality usually have more complicated structures that are difficult to capture by a simple linear waiting cost approximation. For example, the health state of ED patients does not in general deteriorate at a constant rate. On the other hand in certain cases, it is even not appropriate to capture the cost structure as a function of waiting times. For example in a hospital ward, the cost will depend on the medical outcome of the patients after his/her stay in the ward, which may be indirectly related to waiting times. However, it is difficult to express the cost as a direct function of waiting times in this case.

In this dissertation, we aim to explore the priority decisions in service systems with heterogeneous customers and nonlinear delay costs. In particular, we study three design and control problems motivated by service and health care systems.

In the first part of this dissertation, we consider a single server queueing system with two different types of customers. Each customer incurs a cost depending on its type and its waiting time in the queue. The cost functions are assumed to be non-decreasing in the waiting time. Our objective is to minimize the long-run average cost by controlling the order of service for customers in the queue. The waiting costs for different types of patients are different, and also they are not linear in the waiting time. In general, it is very difficult to keep track of the waiting times of all customers in the system, or these kind of information may be very expensive or even impossible to collect in practice. Hence, we are interested in finding "good" policies that do not need such information under nonlinear costs.

The second part of this dissertation is motivated by the admission and discharge decisions given in the intensive care units (ICUs) in hospitals. When a patient arrives to an ICU that is full, the ICU director needs to decide either rejecting this new patient (by way of transfer to some other ward or hospital) or discharging (transferring) an existing patient to make space for this one. In an ICU setting, we are concerned about the dying probabilities of all patients, which can hardly be approximated by a linear function.

In our model, the heterogeneity of ICU patients is in the sense of patients' severity conditions, and the severity conditions of patients could change during their stay in the hospital. We use different stages to model the different health conditions of the patients, and we use a transition matrix to represent the change of patient stages in the ICU. When patients are not in the ICU either due to being early discharged or being rejected upon arrival, they will stay at some non-ICU care place such as general wards, nursery rooms or even home, which we all refer as "the general ward" throughout the dissertation. The stages of patients in such places will change according to a similar pattern with different parameters from the transition matrix in ICU. The patients could leave with an undesired outcome, e.g., death or readmission to ICU. Our objective is to minimize the expected probability of such an undesired outcome by choosing which patients to keep in the ICU.

Although our work is inspired by the ICU admission and discharge decisions, it can be also extended to other systems. For example, consider a service system that can process jobs by two different types of resources, one general resource with infinite capacity and one more efficient resource with finite capacity. The state of jobs can be described by different stages, and they can

become better or worse during the processing. The quality of completed jobs depends on which stage they are finished with, and our objective is to maximize the total output quality by assigning priorities, where jobs with lower priority will be processed by the general resource when the efficient resources are all occupied.

Finally, in the third part of this dissertation we consider a multi-server queueing system with impatient customers, where the customers may change status over time. In many service systems, customers in the queue might leave without being served if their wait exceeds their tolerance, where the tolerance is usually called the "lifetime" or the "patience" of the queueing customer. There is a growing literature on optimal scheduling of impatient customers, most of which assumes customers belong to independent classes. In this dissertation, we have a further assumption that customers belong to different stages, and their stages could change over time either in service or queue. For example, we reconsider the ICU admission and discharge problem. We assume that the patients in the general ward could get readmitted to ICU when there is an available bed. Then, the ICU beds can be considered as servers and the patients in the general ward can be considered as queueing customers. Patients in ICU as well as those waiting for ICU beds in the general ward may leave the system or may change their health stages. We assume that a positive reward is obtained at each service completion, depending on the stage of the leaving customer, and we would like to investigate how priority should be assigned to maximize the long-run average reward.

CHAPTER 2: PRIORITY ASSIGNMENT IN AN M/G/1 QUEUE WITH NONLINEAR WAITING COSTS

In this Chapter, we consider a single server queueing system with two different types of customers. Each customer incurs a cost depending on its type and waiting time in the queue, which is assumed to be non-decreasing and nonlinear. We would like to compare the performances of several static policies.

2.1 Introduction

We consider a queueing system with two types of customers. For example, consider an emergency room with patients who are more severely injured and those who only have small problems. We assume that each customer incurs a waiting cost that depends on its type and the amount of time it spends in the queue before its service starts. We also assume that the cost function is increasing in time spent in the queue. The inter-arrival times and the service times depend on the type as well. Our objective is to determine the best "static" policy that minimizes the long-run average cost. By static policies, we mean policies that are independent of the state of the system, such as priority policies that give priority to a certain type of customers or non-priority policies like first-come-first-serve (FCFS) or last-come-first-serve (LCFS).

The study of priority queueing systems dates back to Cobham (1954), who considered a single server queueing system with customers belonging to several priority classes and the service is non-preemptive. In that work, the author formulated an M/M/1 queueing system with customers belonging to multiple priority classes, and derived the long-run average waiting time in the queue for each class of customer. In these priority queues, the position of a customer in a waiting line is determined by its priority class rather than by its time of arrival in the line as in FCFS. We can refer to Miller (1960) and Jaiswal (1968) for a review of such earlier work on priority queues.

Then the problem that which type of customers should be prioritized was studied. For linear cost functions when the cost is proportional to the delay, Cox and Smith (1961) first established the optimality of the so called "$c\mu$-rule" for M/G/1. According to the $c\mu$-rule, customers with larger $c_i\mu_i$ index are assigned higher priority, where c_i is the per unit time waiting cost and μ_i is the service rate for type i customers, to minimize the long-run average waiting cost in the system. Many other work has been done under the assumption of linear cost functions.

The consideration of nonlinear costs is necessary because approximating customers' delay sensitives with a linear function may not be reasonable in practice. For example, when perishable items wait for service, or landing aircraft wait in the air for landing, etc., the cost of a unit delay cannot be constant as delay increases, because in the former case, the items lose value, and in the latter case, the aircraft may fail for lack of fuel. Dewan and Mendelson (1990) provided a more detailed discussion about the delay cost structures. A number of papers have considered the generalized cost functions in their queueing models. Aféche and Mendelson (2004) compared the revenue-maximizing and socially optimal equilibria under uniform pricing, preemptive, and nonpreemptive priority auctions with an admission price and a generalized delay cost structure. Hassin et al. (2009) showed that relative priorities in an n-class queueing system can reduce customer waiting costs in a single server Markovian model where the goal is to minimize a non-linear cost function of class expected waiting times. Rothkopf and Smith (1984) conjectured that no static priority policy would minimize the expected delay costs when delay cost functions are nonlinear. Haji and Newell (1971) showed that when the cost functions are convex increasing functions, the optimal strategy will always involve serving customers of the same type according to the FCFS discipline. Van Mieghem (1995) proved that when waiting costs are convex in time, a generalized version of the $c\mu$-rule is asymptotically optimal under heavy traffic. Mandelbaum and Stolyar (2004) also proved the heavy-traffic optimality of the generalized $c\mu$-rule under more general settings.

The generalized $c\mu$-rule is actually a dynamic policy that gives priority to the customer who has the largest $C_i'(t)\mu_i$ value in the system at every service completion epoch, where $C_i(t)$ is the cost function of holding a type i customer in the queue for t units of time and $C_i'(t)$ is its first-order derivative. Hence, to implement the generalized $c\mu$-rule, we need to keep track of the waiting times of all customers in the system. In this article, we are interested in finding "good" policies that do not need such information under nonlinear costs. More specifically, we consider static queueing

policies, such as the FCFS and LCFS policies and priority policies that give priority to customers according to their types without considering their waiting times in the system.

The remainder of this article is organized as follows. We describe our model in Section 2.2, then in Section 2.3 we provide analytical comparisons of the three commonly used static policies, namely F, PF_1 and PF_2, where F denotes first-come-first-serve (FCFS) and PF_i denotes the priority policy that prioritizes type i customers and employs FCFS within each type for $i = 1, 2$. We apply our results for polynomial cost functions in Section 2.4 and for exponential convex function in Section 2.5, then we give a theoretical result that shows when the cost functions are convex, we only need to compare F, PF_1 and PF_2 in Section 2.6. We provide similar results when comparing policies L, PL_1 and PL_2, where L denotes last-come-first-serve (LCFS), and PL_i denotes the priority policy that prioritizes type i customers and employs LCFS within each type for $i = 1, 2$, and proved that it suffices to compare only L, PL_1 and PL_2 when the cost functions are concave. By means of a numerical study, we compare the best static policy that we found in earlier sections with the generalized $c\mu$-rule in Section 2.8. Finally, Section 2.9 provides our conclusions.

2.2 Model description

Consider a single server queueing system with two types of customers. Customers arrive to the system according to a Poisson process with rate $\lambda > 0$, and each arriving customer belongs to type $i \in \{1, 2\}$ with probability $p_i > 0$, where $p_1 + p_2 = 1$, independent of the arrival process. Service times for type $i \in \{1, 2\}$ customers are independent and identically distributed (i.i.d.) with mean $\tau_i > 0$ and second moment $\xi_i > 0$. We define $\rho_i \equiv p_i \lambda \tau_i$ and $\rho \equiv \rho_1 + \rho_2$, which we call the system load, and we assume that $\rho < 1$ for stability. Each type i customer incurs a waiting cost $C_i(t)$ when its waiting time in the queue is t, for $t \geq 0$ and $i = 1, 2$. We assume that $C_i(t)$ is first-order differentiable and non-decreasing in t for fixed i.

For such a queueing system, we consider *non-idling* and *non-preemptive* queueing policies that require information only on the type and the order of arrival of all customers in the system. These policies are non-idling and non-preemptive in the sense that the server does not idle as long as there is a customer in the system and that service of a customer who has been taken into service has to

be completed without any preemption before the server moves onto serving another customer. We let Π denote the set of all such queueing policies.

For any policy $\pi \in \Pi$, define the long-run average cost as

$$C_\pi \equiv \lim_{t\to\infty} \frac{\sum_{i=1}^{2}\sum_{k=1}^{n_i(t)} C_i(V_{i,k}^{\pi,x_0})}{t}, \tag{2.1}$$

where $n_i(t)$ is the number of type i customers that has arrived to the system by time t and $V_{i,k}^{\pi,x_0}$ is the waiting time of the kth arriving type i customer under policy π and initial state x_0. Our objective is to identify policies that provide the smallest long-run average waiting cost C_π in the policy set Π for any given initial state x_0. Let W_i^π denote the steady-state waiting time of a type i customer under policy π, then we will show in Appendix A that if $E\Big[\big|C_i(W_i^\pi)\big|\Big]$ exists for both $i \in \{1,2\}$, C_π defined in (2.1) satisfies

$$C_\pi = \lambda p_1 E\Big[C_1(W_1^\pi)\Big] + \lambda p_2 E\Big[C_2(W_2^\pi)\Big]. \tag{2.2}$$

2.3 Comparison of FCFS and fixed priority policies – general cost structures

In this section, we present the comparison of three commonly used policies within Π, namely F, PF_1 and PF_2. In order to compare C_F, C_{PF_1} and C_{PF_2}, we need several definitions and lemmas.

Definition 2.1. (E.g., Shaked and Shanthikumar (2007)). Let X and Y be two random variables with corresponding cumulative distribution functions $F_X(\cdot)$ and $F_Y(\cdot)$. If $F_X(x) \geq F_Y(x)$ for all $x \in (-\infty, \infty)$, then X is said to be smaller than Y in the usual stochastic ordering (denoted by $X \leq_{st} Y$).

Definition 2.2. (Di Crescenzo (1999)). Let X and Y be two non-negative random variables with $X \leq_{st} Y$ and $E[X] < E[Y] < \infty$. Then, we write $Z \equiv \Psi(X,Y)$ to mean that Z is a random variable with probability density function

$$f_Z(x) = \frac{F_X(x) - F_Y(x)}{E[Y] - E[X]}, x \geq 0, \tag{2.3}$$

where $F_X(\cdot)$ and $F_Y(\cdot)$ are the cumulative distribution functions of X and Y, respectively. Di Crescenzo (1999) shows that $f_Z(x)$ is a probability density function.

Lemma 2.1. *(Theorem 4.1 of Di Crescenzo (1999)) Let X and Y be two non-negative random variables satisfying $X \leq_{st} Y$ and $E[X] < E[Y] < \infty$, and let $Z = \Psi(X, Y)$. Let also g be a measurable and differentiable function such that $E[g(X)]$ and $E[g(Y)]$ are finite, and let its derivative g' be measurable and Riemann-integrable on the interval $[x, y]$ for all $0 \leq x \leq y$. Then, $E\big[g'(Z)\big]$ is finite and*

$$E[g(Y)] - E[g(X)] = E[g'(Z)]\big(E[Y] - E[X]\big).$$

Lemma 2.1 presents a probabilistic analogue of the mean value theorem, where Z is a random variable that can be considered as the "mean value" of X and Y. However, unlike for the (deterministic) mean value theorem, Z does not change with the function g, and $Z = \Psi(X, Y)$ is not necessarily ordered (in some stochastic sense) between X and Y. For example, when X and Y are exponential random variables with distinct rates, $Z =_{st} X + Y$ (see Example 3.1 in Di Crescenzo (1999)).

We will use Lemma 2.1 in several of our results including our main result that compares C_F, C_{PF_1} and C_{PF_2}. Before we present this result, we need two more lemmas for the comparison of W_i^F, $W_i^{PF_i}$, $W_{3-i}^{PF_i}$ for $i = 1, 2$.

Lemma 2.2. *(E.g., Gross et al. (2008) and Miller (1960)) For an M/G/1 queueing sysem, the expected steady-state waiting times under FCFS and PF_i are given as follows:*

$$E[W^F] = \frac{\lambda\bar{\xi}}{2(1-\rho)}, \quad E[W_i^{PF_i}] = \frac{\lambda\bar{\xi}}{2(1-\rho_i)}, \quad E[W_{3-i}^{PF_i}] = \frac{\lambda\bar{\xi}}{2(1-\rho_i)(1-\rho)},$$

where $\bar{\xi} = p_1\xi_1 + p_2\xi_2$, and we drop the subscript in W_i^F since the distribution of W^F does not depend on i.

Lemma 2.3. *For fixed $i = 1, 2$, we have $W_i^{PF_i} \leq_{st} W^F \leq_{st} W_{3-i}^{PF_i}$.*

The order of $W_i^{PF_j}$ and W^F for $i, j \in \{1, 2\}$, given in Lemma 2.3 and proved in Appendix A, makes intuitive sense. The steady-state waiting times under FCFS are stochastically less than those

for the non-priority type under a priority policy and greater than those for the priority type. Lemma 2.3 specifies a type of stochastic ordering between these three steady-state random variables.

Based on Lemmas 2.2 and 2.3, for $i \in \{1,2\}$, we define the following random variables,

$$U_i^{PF_i} \equiv \Psi(W_i^{PF_i}, W^F) \text{ and } U_{3-i}^{PF_i} \equiv \Psi(W^F, W_{3-i}^{PF_i}).$$

Note that $U_j^{PF_i}$ is well defined for $i,j \in \{1,2\}$ because $W_i^{PF_i} \leq_{st} W^F \leq_{st} W_i^{PF_{3-i}}$ according to Lemma 2.3, and when $\rho < 1$ and $p_i > 0$, we have $E\big[W_i^{PF_i}\big] < E\big[W^F\big] < E\big[W_i^{PF_{3-i}}\big] < \infty$, for $i = 1,2$ by Lemma 2.2.

Our main results are all stated under the following assumption.

Assumption 2.1. For fixed $i \in \{1,2\}$, $E\Big[\big|C_i(W_i^{\pi})\big|\Big]$ exists for $\pi \in \{F, PF_1, PF_2\}$.

We are now ready to present our main result and an immediate corollary

Theorem 2.1. *Under Assumption 2.1, we have,*

(a) for $i = 1,2$, $C_F \leq C_{PF_i}$ if and only if $a_i \leq b_i$, where

$$a_i \equiv \frac{E\big[C_i'(U_i^{PF_i})\big]}{\tau_i},\ b_i \equiv \frac{E\big[C_{3-i}'(U_{3-i}^{PF_i})\big]}{\tau_{3-i}}. \tag{2.4}$$

(b) $C_{PF_1} \leq C_{PF_2}$ if and only if $(1-\rho_1)(a_2 - b_2) \leq (1-\rho_2)(a_1 - b_1)$.

Corollary 2.1.

(a) If $a_1 \leq b_1$ and $a_2 \leq b_2$, then $C_F \leq C_{PF_1}$ and $C_F \leq C_{PF_2}$.

(b) For fixed $i \in \{1,2\}$, if $a_i \geq b_i$ and

$$\frac{a_i - b_i}{1-\rho_i} \geq \frac{a_{3-i} - b_{3-i}}{1-\rho_{3-i}},$$

then $C_{PF_i} \leq C_F$ and $C_{PF_i} \leq C_{PF_{3-i}}$.

Corollary 2.1 provides necessary and sufficient conditions for the optimality of F, PF_1 and PF_2 within the set of these three policies. These conditions require computation of a_i and b_i for $i = 1,2$. We demonstrate how these computations can be performed for polynomial functions in Section 2.4 and exponential cost functions in Section 2.5.

In order to compute a_i and b_i in Theorem 2.1 and Corollary 2.1, we need to obtain $E\left[C_i'(U_i^{PF_j})\right]$ for $i, j \in \{1, 2\}$. In some situations, the cost functions may be simple for one type and complicated for the other type. For example, we may assume that the cost function for type 2 customers is linear or quadratic, and the cost function for type 1 customers has a more complicated structure. In this case, we can use the next two results to order C_F, C_{PF_1} and C_{PF_2} without computing $E\left[C_1'(U_1^{PF_j})\right]$, but by computing $E\left[C_2'(U_i^{PF_j})\right]$ for $i, j \in \{1, 2\}$.

Corollary 2.2.

(a) If $C_1'(t) \geq \tau_1 \max\{a_2, b_1\}$ *for all* $t \geq 0$, *then* $C_{PF_1} \leq C_F \leq C_{PF_2}$.

(b) If $C_1'(t) \leq \tau_1 \min\{a_2, b_1\}$ *for all* $t \geq 0$, *then* $C_{PF_2} \leq C_F \leq C_{PF_1}$.

(c) If $\tau_1 a_2 \leq C_1'(t) \leq \tau_1 b_1$ *for all* $t \geq 0$, *then* $C_F \leq C_{PF_1}$ *and* $C_F \leq C_{PF_2}$.

Corollary 2.3. *If* $E[C_2'(U_1^{PF_2})] \neq 0$ *and* $E[C_2'(U_1^{PF_1})] \neq 0$, *define*

$$\alpha \equiv \frac{\tau_1 E[C_2'(U_2^{PF_2})]}{\tau_2 E[C_2'(U_1^{PF_2})]} \quad \textit{and} \quad \beta \equiv \frac{\tau_1 E[C_2'(U_2^{PF_1})]}{\tau_2 E[C_2'(U_1^{PF_1})]}.$$

(a) If $C_1'(t) \geq \max\{\alpha, \beta\} C_2'(t)$ *for all* $t \geq 0$, *then* $C_{PF_1} \leq C_F \leq C_{PF_2}$.

(b) If $C_1'(t) \leq \min\{\alpha, \beta\} C_2'(t)$ *for all* $t \geq 0$, *then* $C_{PF_2} \leq C_F \leq C_{PF_1}$.

(c) If $\alpha C_2'(t) \leq C_1'(t) \leq \beta C_2'(t)$ *for all* $t \geq 0$, *then* $C_F \leq C_{PF_1}$ *and* $C_F \leq C_{PF_2}$.

Corollary 2.2 compares $C_1'(t)$ with two fixed quantities, $\tau_1 a_2$ and $\tau_1 b_1$, for all $t \geq 0$. Hence $C_1'(t)$ has to be bounded from either above or below to be able to apply this result as in the case of a linear or logarithmic cost function for type 1 customers. On the other hand, in Corollary 2.3, we compare $C_1'(t)$ with two time-varying quantities, $\alpha C_2'(t)$ and $\beta C_2'(t)$, and hence $C_1'(t)$ does not need to be bounded. However, in Corollary 3, we require that $E\left[C_2'(U_1^{PF_i})\right]$ for $i = 1, 2$ be non-zero, which is satisfied when $C_2(\cdot)$ is a strictly increasing function. When $C_2'(t)$ is a constant, i.e., $C_2(t)$ is linear, it can be shown that $\tau_1 a_2 = \tau_1 b_1 = \alpha C_2'(t) = \beta C_2'(t)$, and hence, these two corollaries reduce to one another. We will demonstrate how these two corollaries can be applied in the case of polynomial and exponential functions in Sections 2.4 and 2.5, respectively.

2.4 Comparison of FCFS and fixed priority policies – polynomial cost functions

In this section, we focus on the case where the cost function for at least one type of customer is polynomial. In particular, suppose that for some $i = 1, 2$,

$$C_i(t) = \sum_{l=1}^{j(i)} h_l(i) t^l, \tag{2.5}$$

where $j(i)$ is the (finite) degree of the polynomial function $C_i(t)$, and $h_l(i)$ are some real numbers such that $C_i'(t) \geq 0$ for all $t \geq 0$. We first provide conditions under which Assumption 2.1 holds the polynomial cost functions in Lemma 2.4 (the proof is provided in Appendix A).

Lemma 2.4. *Assumption 2.1 is satisfied for $C_i(t)$ that takes the form of (2.5) if $\rho < 1$, and the first $(j(i)+1)$ moments of service times for both types are finite.*

In order to apply Theorem 2.1 and Corollaries 2.1, 2.2 and 2.3, we need to compute $E\left[C_i'(U_k^{PF_m})\right]$ for some $i, k, m \in \{1, 2\}$, where

$$E\left[C_i'(U_k^{PF_m})\right] = \sum_{l=1}^{j(i)} l h_l(i) E\left[\left(U_k^{PF_m}\right)^{l-1}\right]. \tag{2.6}$$

Here, $E\left[\left(U_k^{PF_m}\right)^{l-1}\right]$ for $l = 1, \ldots, j(i)$ can be computed as

$$E\left[\left(U_k^{PF_m}\right)^{l-1}\right] = \frac{E\big[(W^F)^l\big] - E\big[(W_k^{PF_m})^l\big]}{l\left(E[W^F] - E[W_k^{PF_m}]\right)}, \tag{2.7}$$

by letting $g(x) = x^l/l$ in Lemma 2.1. See Appendix A for the proof of (2.6).

To demonstrate and to gain insights, in the remainder of this section, we consider polynomial cost functions with an order of at most two.

2.4.1 Quadratic cost functions for both customer types

Suppose that $C_i(t) = k_i t^2 + h_i t$, where $k_i, h_i \geq 0$ and $i \in \{1, 2\}$. Let ζ_i denote the third moment of the service times for type $i \in \{1, 2\}$ and $\bar{\zeta} \equiv p_1\zeta_1 + p_2\zeta_2$. Then, we show in Appendix A that

$$a_i = \frac{k_i}{\tau_i}\left[\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_i\xi_i}{1-\rho_i} + \frac{\xi_{3-i}}{\tau_{3-i}}\right] + \frac{h_i}{\tau_i}, \tag{2.8}$$

$$b_i = \frac{k_{3-i}}{\tau_{3-i}}\left[\frac{2\bar{\zeta}}{3\bar{\xi}}\left(1 + \frac{1}{1-\rho_i}\right) + \frac{\lambda\bar{\xi}}{1-\rho}\left(1 + \frac{1}{1-\rho_i}\right) + \frac{\xi_i}{\tau_i(1-\rho_i)^2}\right] + \frac{h_{3-i}}{\tau_{3-i}}. \tag{2.9}$$

When we compare a_i's and b_i's given above, we found that for $i = 1, 2$,

$$b_{3-i} - a_i = \left(\frac{k_i}{\tau_i}\right)\left[\frac{2\bar{\zeta}}{3\bar{\xi}(1-\rho_{3-i})} + \frac{\lambda\bar{\xi}\rho_{3-i}}{1-\rho}\left(\frac{1}{1-\rho_{3-i}} + \frac{1}{1-\rho_i}\right)\right.$$
$$\left. + \lambda p_{3-i}\xi_{3-i}\left(\frac{1}{(1-\rho_{3-i})^2} + \frac{1}{1-\rho_i} + \frac{1}{1-\rho_{3-i}}\right)\right] \geq 0, \tag{2.10}$$

because $\rho, \rho_1, \rho_2 < 1$ and all moments of service times are non-negative. Therefore, when both cost functions are quadratic, if $a_i > b_i$ for some $i = 1, 2$ (and thus $a_{3-i} \leq b_i < a_i \leq b_{3-i}$), then PF_i is the best among F, PF_1 and PF_2 according to Corollary 2.1(b); otherwise ($a_1 \leq b_1$ and $a_2 \leq b_2$), F is the best among F, PF_1 and PF_2 by Corollary 2.1(a). By comparing a_i and b_i for fixed $i \in \{1, 2\}$, we find that $a_i \geq b_i$ (and thus PF_i is the best), if and only if

$$\frac{k_i}{\tau_i} \geq \frac{\frac{k_{3-i}}{\tau_{3-i}}\left[\frac{2-\rho_i}{1-\rho_i}\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_i\xi_i}{1-\rho_i}\right) + \frac{\xi_i}{\tau_i}\right] + \frac{h_{3-i}}{\tau_{3-i}} - \frac{h_i}{\tau_i}}{\left[\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_i\xi_i}{1-\rho_i} + \frac{\xi_{3-i}}{\tau_{3-i}}\right]}. \tag{2.11}$$

Hence, using Corollary 2.1 and Equations (2.8) and (2.9), we were able to obtain a complete order of policies F, PF_1 and PF_2 when cost functions are quadratic. To gain further insights, we next consider the case that $h_1/\tau_1 = h_2/\tau_2$, e.g., when $C_i(t) = k_i t^2$ for $i = 1, 2$, in the remainder of this section.

Proposition 2.1. *For quadratic cost functions, when $h_1/\tau_1 = h_2/\tau_2$, the best policy among PF_1, PF_2 and F is characterized as follows: PF_2 is the best if $k_1/k_2 < A\tau_1/\tau_2$, PF_1 is the best if*

$k_1/k_2 > B\tau_1/\tau_2$, and F is the best if $A\tau_1/\tau_2 \le k_1/k_2 \le B\tau_1/\tau_2$, where

$$A \equiv \frac{\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_2\xi_2}{1-\rho_2} + \frac{\xi_1}{\tau_1}}{\frac{2-\rho_2}{1-\rho_2}\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_2\xi_2}{1-\rho_2}\right) + \frac{\xi_2}{\tau_2}} < \frac{\frac{2-\rho_1}{1-\rho_1}\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_1\xi_1}{1-\rho_1}\right) + \frac{\xi_1}{\tau_1}}{\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_1\xi_1}{1-\rho_1} + \frac{\xi_2}{\tau_2}} \equiv B.$$

Proposition 2.1 completely characterized the best policy among F, PF_1 and PF_2 for quadratic cost functions with $h_1/\tau_1 = h_2/\tau_2$. In particular, it provides thresholds on k_1/k_2 such that PF_1 is the best if k_1/k_2 is large, PF_2 is the best if k_1/k_2 is small and F is the best if k_1/k_2 is medium. Proposition 2.1 also provides some useful insights. For example, when $\lambda \to 1/\bar{\tau}$, where $\bar{\tau} = p_1\tau_1 + p_2\tau_2$ is the mean service time, we find that $A \to \frac{1-\rho_2}{2-\rho_2}$ and $B \to \frac{2-\rho_1}{1-\rho_1}$, then in heavy traffic F is preferred when k_1 and k_2 are not significantly different.

Figure 2.1 provides plots of $\ln A$ and $\ln B$ with respect to λ, from which we can observe the above insights. We assume that the service times are exponentially distributed, and $\tau_1 = 1$.

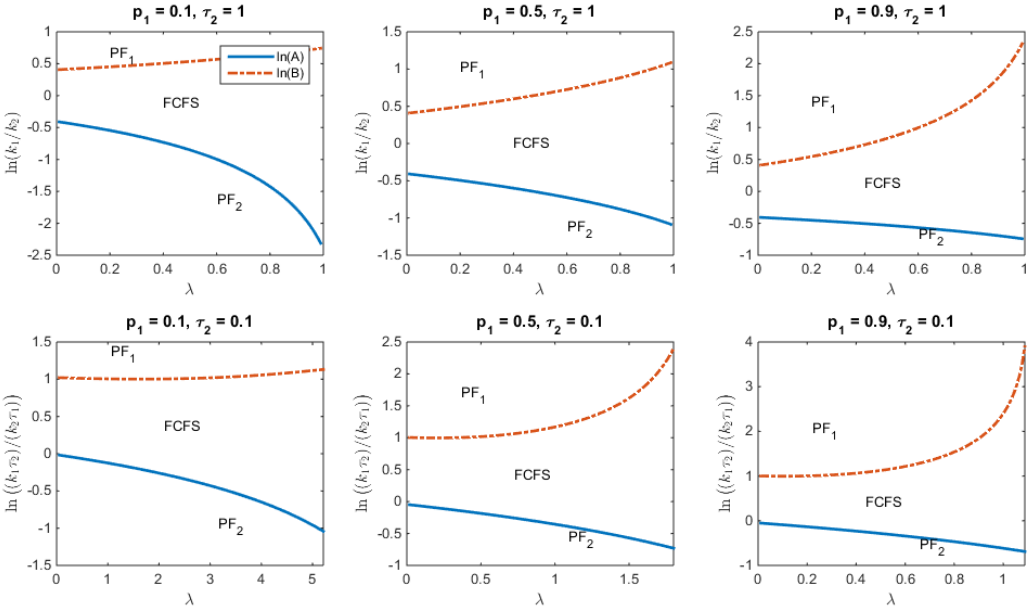

Figure 2.1: Plots of A and B with respect to λ for exponential service times

From Figure 2.1, we also notice that A and B change monotonically in λ, and our next result shows that it is not necessary always the case.

Proposition 2.2. *A decreases in λ if and only if*

$$\frac{\xi_2}{\tau_2} - \frac{(2-\rho_2)\xi_1}{(1-\rho_2)\tau_1} < \frac{\frac{p_2\tau_2}{(1-\rho_2)^2}\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_2\xi_2}{1-\rho_2}\right)\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_2\xi_2}{1-\rho_2} + \frac{\xi_1}{\tau_1}\right)}{\frac{\bar{\xi}}{(1-\rho)^2} + \frac{p_2\xi_2}{(1-\rho_2)^2}}, \tag{2.12}$$

and B increases in λ if and only if

$$\frac{\xi_1}{\tau_1} - \frac{(2-\rho_1)\xi_2}{(1-\rho_1)\tau_2} < \frac{\frac{p_1\tau_1}{(1-\rho_1)^2}\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_1\xi_1}{1-\rho_1}\right)\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_1\xi_1}{1-\rho_1} + \frac{\xi_2}{\tau_2}\right)}{\frac{\bar{\xi}}{(1-\rho)^2} + \frac{p_1\xi_1}{(1-\rho_1)^2}}. \tag{2.13}$$

Proposition 2.2 provides conditions under which the thresholds for the optimality of these policies monotonically change with λ. Note that the right-hand side of (2.12) and (2.13) are both nonnegative, which then give a sufficient condition, i.e., if $\frac{\xi_1}{\tau_1} > \frac{\xi_2(1-\rho_2)}{\tau_2(2-\rho_2)}$, then A decreases in λ and if $\frac{\xi_1}{\tau_1} < \frac{(2-\rho_1)\xi_2}{(1-\rho_1)\tau_2}$, then B increases in λ. If $\frac{\xi_1}{\tau_1} < \frac{\xi_2(1-\rho_2)}{\tau_2(2-\rho_2)}$, the left-hand side of (2.12) is positive and the right-hand converges to 0 as $p_2 \to 0$, thus (2.12) may not hold for very small p_2 and A increases in λ, which means PF_2 is preferred for a larger range of values of k_1/k_2 as λ increases. Similarly if $\frac{\xi_1}{\tau_1} > \frac{(2-\rho_1)\xi_2}{(1-\rho_1)\tau_2}$, we notice that B decreases in λ when p_1 is sufficiently small, and thus PF_1 is preferred for a larger range of values of k_1/k_2 as λ increases. *Thus, we conclude that if the proportion of one type of customers is sufficiently small, and the ratio of ξ_i/τ_i for the same type is sufficiently large, then prioritizing that type becomes more preferable as λ increases.*

On the other hand, if the ratios ξ_1/τ_1 and ξ_2/τ_2 are similar, e.g., when service times are i.i.d. for all customers, then the interval (A, B) enlarges as λ increases, which indicates that F is more preferable as the system becomes busier. However, this result is only to compare F with the fixed priority policies, under which prioritization is always given to one type of customers irrespective of the waiting times of customers. Under such a static priority rule, as the arrival rate increases, the non-priority customers will wait much longer, resulting in a significant increase in waiting costs as the cost increases at a higher rate for longer waiting. If we consider dynamic prioritization that takes into account the waiting times of customers, then F may not perform as well as a dynamic priority rule as we illustrate later by a numerical study in Section 2.8.

Figure 2.2 provides plots of how A and B change with respect to p_1. We assume that the service times are exponentially distributed, and $\tau_1 = 1$.

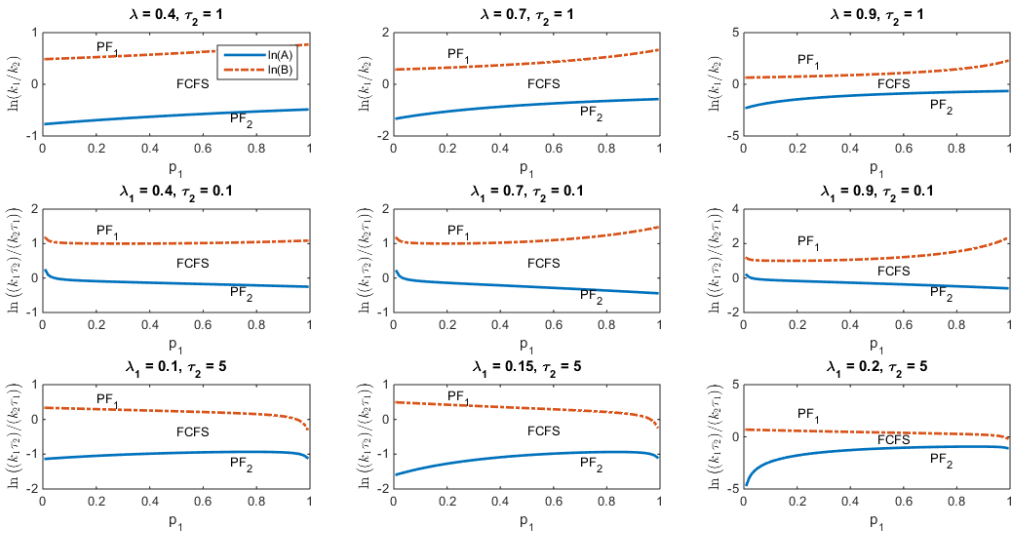

Figure 2.2: Plots of A and B with respect to p_1 for exponential service times

We notice from Figure 2.2 that when the mean service times are different, A and B do not always change monotonically in p_1. It is difficult to provide an algebraical analysis of how A and B change with respect to p_i, given that $\bar{\xi}$, $\bar{\zeta}$, ρ and ρ_i all change in p_i. We can first look at the heavy traffic setting, i.e., when $\lambda \to 1/\bar{\tau}$. We have,

$$\lim_{\substack{p_1\to 0\\ \lambda\to 1/\tau_2}} A = 0, \quad \lim_{\substack{p_1\to 0\\ \lambda\to 1/\tau_2}} B = 2, \quad \lim_{\substack{p_1\to 1\\ \lambda\to 1/\tau_1}} A = \frac{1}{2}, \quad \lim_{\substack{p_1\to 1\\ \lambda\to 1/\tau_1}} B = \infty.$$

In words, we would not like to prioritize type i customers if the proportion of this type is close to one under heavy traffic; on the other hand, if the proportion of type i customers is close to 0, we would prioritize this type is preferred if $k_i/\tau_i > 2k_{3-i}/\tau_{3-i}$, and otherwise choose F under heavy traffic.

Furthermore, if we assume that the service times are i.i.d., then, $\bar{\xi}$, $\bar{\zeta}$, and ρ no longer depend on p_i, and only ρ_i will change in p_i. Our next result analyzes the monotonicity of A and B in p_i under such a setting.

Proposition 2.3. *When the service times are i.i.d. for all customers, A and B both increase in p_1 (and hence decrease in p_2).*

Proposition 2.3 indicates that as the proportion of one type increases, giving priority to that type is preferred for a smaller range of k_1/k_2 values, while prioritizing the other type is preferred for a wider range of k_1/k_2 values. Intuitively, we would not like to prioritize type i customers if their total load on the system is too large under quadratic waiting cost, since by prioritizing such a large group, type $3-i$ customers will wait much longer, resulting in a significant increase in waiting costs. This intuition only works when service times are identical for all customers, and we have numerical examples in Section 2.8 show that A and B do not necessarily increase in p_1 when service times are different for the two types of customers.

When the service times are i.i.d., we have, A decreases in λ and increases in p_1, and B increases in both λ and p_1 from Propositions 2.2 and 2.3. Hence, we could obtain the range of values of A and B from their monotonicity. More specifically, we have

$$A > \lim_{\substack{p_1\to 0\\ \lambda\to 1/\tau}} A = 0, \text{ and } A < \lim_{\substack{p_1\to 1\\ \lambda\to 0}} A = \frac{\frac{2\zeta}{3\xi}+\frac{\xi}{\tau}}{\frac{4\zeta}{3\xi}+\frac{\xi}{\tau}} = \frac{2\tau\zeta+3\xi^2}{4\tau\zeta+3\xi^2},$$

and

$$B > \lim_{\substack{p_1\to 0\\ \lambda\to 0}} B = \frac{4\tau\zeta+3\xi^2}{2\tau\zeta+3\xi^2}, \text{ and } B < \lim_{\substack{p_1\to 1\\ \lambda\to 1/\tau}} B = \infty.$$

In the end of this section, we compare the values of A and B under two different service time distributions. Let $A_{exp}[B_{exp}]$ and $A_{det}[B_{det}]$ denote the values of $A[B]$ under exponential and deterministic service times, respectively.

Proposition 2.4.

(a) $A_{exp} \le A_{det}$ *if and only if* $\tau_2 \le \tau_1(2-\rho_2)/(1-\rho_2)$.

(b) $B_{exp} \ge B_{det}$ *if and only if* $\tau_2 \ge \tau_1(1-\rho_1)/(2-\rho_1)$.

Proposition 2.4 implies that if $\frac{\tau_1}{\tau_2} \in \left(\frac{1-\rho_2}{2-\rho_2}, \frac{2-\rho_1}{1-\rho_1}\right)$, then $A_{exp} \le A_{det} < B_{det} \le B_{exp}$, and hence, when the expected service times are not significantly different for the two types of customers, F is preferable for a wider range of values of k_1/k_2 under exponential service times than under deterministic service times. In other words, *when service time variability is high and the two types are not too different in terms of mean service times, we are more likely to select F since any fixed priority policy may increase the waiting times significantly in this case.* On the other hand, if one

type is sufficiently faster to serve in the mean sense, say, $\tau_1/\tau_2 > (2-\rho_1)/(1-\rho_1)$, then $A_{exp} \leq A_{det}$ and $B_{exp} \leq B_{det}$, which means that under exponential service times, PF_2 (prioritizing the slower type) is preferred for a smaller range of values of k_1/k_2, and PF_1 (prioritizing the faster type) is preferred for a wider range of values of k_1/k_2 than that under deterministic service times.

2.4.2 Quadratic cost for one type and general cost for the other type

Suppose one type of customers has a quadratic cost function (say, $C_2(t) = k_2t^2 + h_2t$ for $h_2, k_2 \geq 0$), and the other type has a general cost function. In this section, we will demonstrate how Corollaries 2.2 and 2.3 can be applied in such a case.

We first focus on Corollary 2.2. When $C_2(t)$ is a quadratic function, a_2 and b_1 are given by (2.8) and (2.9), respectively, and hence we have $a_2 \leq b_1$ from (2.10). Therefore, in Corollary 2.2, we can replace $\max\{a_2, b_1\}$ with b_1 and $\min\{a_2, b_1\}$ with a_2. Furthermore, since $a_2 \leq b_1$, we know that the interval $(\tau_1 a_2, \tau_1 b_1)$ is not necessarily an empty set and hence part (c) of Corollary 2.2 could be applicable. Consequently, Corollary 2.2 implies that if the smallest marginal increase in $C_1(t)$ is at least $\tau_1 b_1$, then type 1 customers should be prioritized; if the largest marginal increase in $C_1(t)$ is at most $\tau_1 a_2$, then type 2 customers should be prioritized; and if the marginal increase in $C_1(t)$ lies between $\tau_1 a_2$ and $\tau_1 b_1$ at all times, then $FCFS$ should be employed. Furthermore, by Equations (2.8), (2.9) and (2.10), we notice that a_i, b_i and the difference $b_{3-i} - a_i$ all increase in λ for $i \in \{1, 2\}$ since ρ_i, ρ_{3-i}, $1/(1-\rho)$, $1/(1-\rho_i)$ and $1/(1-\rho_{3-i})$ all increase in λ. This implies that the bounds $\tau_1 a_2$ and $\tau_1 b_1$, and the length of the interval $(\tau_1 a_2, \tau_1 b_1)$ are all increasing as λ becomes larger. Furthermore, both a_2 and b_1 go to infinity as λ approaches $\bar{\tau}^{-1}$. Combining this with Corollary 2.2 leads to an important conclusion: *if the marginal increase in the cost function of one type is bounded and the other type has a quadratic cost function, then it is better to prioritize the type with quadratic cost under heavy traffic no matter what the service time and cost parameters are.* We illustrate these observations further in Example 2.1.

Example 2.1.

(i) When $C_1(t) = h_1 t$ for $t \geq 0$, where h_1 is positive, $C_1'(t) = h_1$ is bounded. Then, PF_1 is preferred if $h_1 \geq \tau_1 b_1$, PF_2 is preferred if $h_1 \leq \tau_1 a_2$, and F is preferred otherwise. Hence, Corollary 2.2 leads to a complete characterization of the best policy among F, PF_1 and PF_2

in this case. We notice that as λ increases, PF_1 becomes less preferable, the range of h_1 values for which F is preferred is shifting up and becoming wider, and PF_2 is also preferred for a wider range of h_1 values. Furthermore, since $a_2 \to \infty$ as $\lambda \to 1/\bar{\tau}$, PF_2 is preferred for any finite h_1. *This means that when type 1 customers have linear and type 2 customers have quadratic waiting costs, prioritizing type 2 customers will reduce the cost in heavy traffic no matter what the cost and service time parameters are.*

(ii) When $C_1(t) = h_1(e^{\alpha_1 t} - 1)$ for $t \geq 0$ and positive constants h_1 and α_1, we have $C_1'(t) \geq h_1\alpha_1$ for all $t \geq 0$, and hence C_{PF_1} is the smallest if $h_1\alpha_1 \geq \tau_1 b_1$.

(iii) When $C_1(t) = h_1 \ln(t+1)$ for $t \geq 0$ and positive constant h_1, we have $C_1'(t) \leq h_1$ for all $t \geq 0$, and hence C_{PF_2} is the smallest if $h_1 \leq \tau_1 a_2$. As $\lambda \to 1/\bar{\tau}$, the bound $\tau_1 a_2$ goes to infinity, which indicates that PF_2 is the best for any h_1 under heavy traffic.

$\Diamond$

We next apply Corollary 2.3 to the case where the waiting cost for type 2 customers is a quadratic function.

Proposition 2.5. *When $C_2(t)$ is a quadratic function, α and β in Corollary 2.3 are given as*

$$\alpha = \left(\frac{\tau_1}{\tau_2}\right) \frac{k_2\left[\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}(2-\rho-\rho_2)}{(1-\rho)(1-\rho_2)} + \frac{\lambda p_1\xi_1(1-\rho)}{\rho_1(1-\rho_2)}\right] + h_2}{k_2\left[\frac{2\bar{\zeta}(2-\rho_2)}{3\bar{\xi}(1-\rho_2)} + \frac{\lambda\bar{\xi}(2-\rho_2)}{(1-\rho)(1-\rho_2)} + \frac{\lambda p_2\xi_2}{\rho_2(1-\rho_2)^2}\right] + h_2},$$

$$\beta = \left(\frac{\tau_1}{\tau_2}\right) \frac{k_2\left[\frac{2\bar{\zeta}(2-\rho_1)}{3\bar{\xi}(1-\rho_1)} + \frac{\lambda\bar{\xi}(2-\rho_1)}{(1-\rho)(1-\rho_1)} + \frac{\lambda p_1\xi_1}{\rho_1(1-\rho_1)^2}\right] + h_2}{k_2\left[\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}(2-\rho-\rho_1)}{(1-\rho)(1-\rho_1)} + \frac{\lambda p_2\xi_2(1-\rho)}{\rho_2(1-\rho_1)}\right] + h_2}.$$

Furthermore, we have the following:

(a) $\alpha < \beta$.

(b) If (2.12) *holds, then α decreases in λ, and if* (2.13) *holds, then β increases in λ. (When $h_2 = 0$,* (2.12) *and* (2.13) *are also necessary for the respective results). Furthermore,*

$$\lim_{\lambda\to 1/\bar{\tau}} \alpha = \frac{\tau_1}{\tau_2} \cdot \frac{1-\rho_2}{2-\rho_2}, \quad \lim_{\lambda\to 1/\bar{\tau}} \beta = \frac{\tau_1}{\tau_2} \cdot \frac{2-\rho_1}{1-\rho_1}.$$

(c) When the service times are i.i.d. for all customers, α and β both increase in p_1 (and hence decrease in p_2). Additionally, we have,

$$\lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 0}} \alpha = 0, \quad \lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 1}} \alpha = \frac{\tau_1}{2\tau_2}, \text{ and } \lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 0}} \beta = \frac{2\tau_1}{\tau_2}, \quad \lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 0}} \beta = \infty.$$

By Proposition 2.5(a), we can replace $\max\{\alpha, \beta\}$ with β and $\min\{\alpha, \beta\}$ with α in Corollary 2.3. Besides, when the service times are i.i.d., as λ increases α decreases and β increases which implies that when the system becomes more congested, the region where F is preferred becomes larger. Note that the conditions given by Corollary 2.3 are sufficient but not necessary. For example, PF_2 is the best among three if $C_1'(t) \leq \alpha C_2'(t)$ for all t, while it does not mean that PF_2 is not the best if the condition does not hold. Hence, the fact that α increases in p_1 only implies that Corollary 2.3 can provide a wider range of $C_1'(t)/C_2'(t)$ under which PF_2 is the best.

We demonstrate how Corollary 2.3 could be applied for functions given in Example 2.1, and to discuss the difference of this result from Corollary 2.2. In particular, we show that both Corollaries 2.2 and 2.3 could be useful in different situations.

Example 2.2.

(i) Let $C_1(t) = h_1 t$ for $t \geq 0$, where h_1 is positive. We compare $C_1'(t) = h_1$ with $\alpha C_2'(t)$ and $\beta C_2'(t)$ for all t, where $C_2'(t) = 2k_2 t + h_2$. Since $C_1'(t)$ is fixed and $C_2'(t)$ is increasing without any bound, the only possible case is that $C_1'(t) \leq \alpha C_2'(t)$ for all t, which is true if and only if $h_1 \leq \alpha h_2$. Hence, by applying Corollary 2.3 we have PF_2 is better than F and PF_1 if $h_1 \leq \alpha h_2$. Hence, Corollary 2.3 provides a partial comparison of the three policies. On the other hand for this case, Example 2.1(i) showed that Corollary 2.2 lead to a complete characterization. (Indeed, one can show that $\alpha h_2 < \tau_1 a_2$.) Hence, Corollary 2.2 is more useful for this example.

(ii) When $C_1(t) = h_1(e^{\alpha_1 t} - 1)$ for $t \geq 0$ and positive constants h_1 and α_1, then $C_1'(t) = h_1\alpha_1 e^{\alpha_1 t}$. Since $C_1'(t)$ is exponential and $C_2'(t)$ is linear, $C_1'(t)$ will be greater than $C_2'(t)$ for sufficiently large t, and thus the only possible case from Corollary 2.3 is that $C_1'(t) \geq \beta C_2'(t)$ for all $t \geq 0$, which is true if and only if the following condition holds (the proof is provided in Appendix

A):

$$\begin{cases} h_1 \geq \frac{h_2\beta}{\alpha_1}, & \text{if } h_2 \geq \frac{2k_2}{\alpha_1}, \\ h_1 \geq \frac{2k_2\beta}{\alpha_1^2} e^{\left(\frac{h_2\alpha_1}{2k_2}-1\right)}, & \text{otherwise.} \end{cases} \tag{2.14}$$

In this example, both Corollaries 2.2 and 2.3 provide conditions under which PF_1 is the best. Whether Corollary 2.2 or 2.3 is better depends on the system parameters.

Given we can show that $\tau_1 b_1 > h_2\beta$, which bound is better depends on the order of $\frac{\tau_1 b_1}{\alpha_1}$ and $\frac{2k_2\beta}{\alpha_1^2} e^{\left(\frac{h_2\alpha_1}{2k_2}-1\right)}$. To be more specific, Corollary 2.2 is more useful if $\tau_1 b_1 \alpha_1 < 2k_2\beta e^{\left(\frac{h_2\alpha_1}{2k_2}-1\right)}$ and $h_2 < \frac{2k_2}{\alpha_1}$, and Corollary 2.3 is more useful otherwise.

(iii) When $C_1(t) = h_1 \ln(t+1)$ for $t \geq 0$ and positive constant h_1, we have $C_1'(t) = h_1/(t+1)$. As $t \to \infty$, we have $C_1'(t) \to 0$ and $C_2'(t) \to \infty$. Hence, the only possible case from Corollary 2.3 is that $C_1'(t) \leq \alpha C_2'(t)$ for all $t \geq 0$, which is true if and only if $h_1 \leq \alpha h_2$. By applying Corollary 2.3, we find that if $h_1 \leq \alpha h_2$, then C_{PF_2} is the smallest. In this example, both Corollaries 2.2 and 2.3 provide upper bounds on h_1 when PF_2 is the best, and since we can show that $\tau_1 a_2 > h_2\alpha$, the bound from Corollary 2.2 is better than that from Corollary 2.3.

$\Diamond$

2.4.3 Linear cost for one type and general cost for the other type

Suppose that one type of customers has a linear cost function (say, $C_2(t) = h_2 t$ for $t \geq 0$ and $h_2 > 0$). Then $a_2 = b_1 = h_2/\tau_2$ and $\alpha = \beta = \tau_1/\tau_2$. Then, Corollaries 2.2 and 2.3 reduce to the same result:

(a) If $C_1'(t) \geq \frac{h_2\tau_1}{\tau_2}$ for all $t \geq 0$, then $C_{PF_1} \leq C_F \leq C_{PF_2}$.

(b) If $C_1'(t) \leq \frac{h_2\tau_1}{\tau_2}$ for all $t \geq 0$, then $C_{PF_2} \leq C_F \leq C_{PF_1}$.

We next observe what this result implies for three forms of $C_1(t)$:

Example 2.3. (i) When $C_1(t) = h_1 t$ for $t \geq 0$ and positive constant h_1, PF_1 is preferred if $h_1/\tau_1 \geq h_2/\tau_2$ and PF_2 is preferred otherwise. This is consistent with the well-known $c\mu$ rule, which indicates that under linear cost functions we should give priority to the type with the larger $c\mu$ value, where h_i is c_i and $\tau_i = 1/\mu_i$ in our case (Cox and Smith (1961)).

(ii) When $C_1(t) = h_1(e^{\alpha_1 t} - 1)$ for $t \geq 0$ and positive constants h_1 and α_1, C_{PF_1} is the smallest if $h_1\alpha_1/\tau_1 \geq h_2/\tau_2$.

(iii) When $C_1(t) = h_1 \ln(t+1)$ for $t \geq 0$ and positive constant h_1, C_{PF_2} is the smallest if $h_1/\tau_1 \leq h_2/\tau_2$.

$\Diamond$

Based on Example 3, one may wonder if F could be the best policy if one of the cost function is linear. Indeed, in Example 1(i), we showed that F could be better if one cost is linear and the other is quadratic (since a_2 is strictly less than b_1 from (2.10)).

2.4.4 Discussions

We present some important insights we gained from Section 2.4 here.

We first show that the average waiting cost is bounded when the first $n+1$ moments of the service times exist and $\rho < 1$ for if $C_i(\cdot)$ has a polynomial form with n degree.

When both cost functions are quadratic, we can completely characterize the best policy among F, PF_1 and PF_2 in terms of the first three moments of the service times and the linear and quadratic term coefficients of the two types. When the quadratic term coefficients are both zero, i.e., both functions are linear, only the mean service times and the linear coefficients will affect the characterization, and the best policy could always be in one of the two priority policies, i.e., F is never the unique best policy. On the other hand, when the linear term coefficients are both zero, the complete characterization compares k_1/k_2 with two threshold values, which are computed by the first three moments of the service times. F is the unique best policy when k_1/k_2 lies between the two threshold values. A priority policy is the best when the quadratic coefficients are significantly different and the priority should be given to the type with larger quadratic coefficient. However, this does not mean that F is the best when the difference between the quadratic coefficients are insignificant. For example, if the quadratic coefficients are the same, then PF_i is the best if $\tau_{3-i} > \frac{\tau_i(2-\rho_i)}{1-\rho_i}$ and $\xi_{3-i} > \xi_i$, i.e., priority should be given to the type with sufficiently small mean service times and smaller second moment of service times.

Service time distributions affect the selection of the best policy since the threshold values depend on the first three moments of the service time distributions. If the mean service times are

not significantly different, F is more preferable under exponential service distribution as opposed to deterministic service distribution, since larger variability will increase the risk of prioritization. However, if one type is sufficiently faster to serve than the other type, then with larger service variability, prioritizing the faster type is more preferable, while prioritizing the slower type is less preferable.

Assume service times are identical for all customers so that we can focus on how other system parameters, such as the arrival rate and the proportion of each type, affect the best policy. We find that as the system becomes more congested, both priority policies are preferred for a smaller range of parameters while F becomes more preferable. As the proportion of one type of customers increases, prioritizing this type becomes less preferable.

Corollaries 2.2 and 2.3 both provide sufficient conditions to select the best policy. When applying these two results to the case when the cost function for one type of customers is quadratic, the interval of parameters when F is the unique best policy always exists from both corollaries when the quadratic coefficient is strictly positive, while when the quadratic term is zero, the two corollaries reduce to one another, from which we could not obtain sufficient conditions under which F is the unique best. When the service times are i.i.d. for all customers and as the system becomes more congested, both corollaries provide a larger range of parameters for F to be preferred, while the two priority policies are less preferred. As the proportion of one type of customers increases, both corollaries provide a smaller range of parameters for which prioritizing that type of customer are preferred. We compare the bounds provided by the two results and find that in some cases Corollary 2.2 would provide a better bound, while there are also situations when the specific parameters determines which one is better.

2.5 Comparison of FCFS and fixed priority policies – exponential convex cost functions

In this section, we focus on the case where the cost function for type i customers has the following form:

$$C_i(t) = k_i(e^{h_i t} - 1), \text{ for } t \geq 0 \tag{2.15}$$

where $k_i, h_i \geq 0$ for some $i \in \{1,2\}$. For Assumption 2.1 to hold for an exponential cost function of the form 2.15, we need $\widetilde{W}_i^{\pi}(-h_i)$ exists for $\pi \in \{PF_1, PF_2, F\}$, where $\widetilde{W}_i^{\pi}(s)$ denote the LST of W_i^{π}. In the remainder, we first assume that Assumption 2.1 holds for both types, and later in the section we illustrate how to find the bound of h_i in the example of i.i.d. exponential service times.

2.5.1 General service time distributions

For $i = 1, 2$, let $\tilde{S}_i(s)$ denote the LST of the service time distribution for type i customers, and $\tilde{S}(s) \equiv p_1\tilde{S}_1(s) + p_2\tilde{S}_2(s)$. We first compute $E\left[C_i'(U_j^{PF_m})\right]$ for some $i, j, m \in \{1,2\}$ before applying the results from Section 2.3. For $i = 1, 2$, we have,

$$E\left[C_i'(U_j^{PF_m})\right] = k_i h_i \widetilde{U}_j^{PF_m}(-h_i), \tag{2.16}$$

where $\widetilde{U}_j^{PF_m}$ denotes the LST of $U_j^{PF_m}$ for $j, m \in \{1,2\}$, which can be computed by letting $g(x) = -e^{-sx}/s$ in Lemma 2.1 as

$$\widetilde{U}_j^{PF_m}(s) = E\left[g'\left(U_j^{PF_m}(s)\right)\right] = \frac{\widetilde{W}_j^{PF_m}(s) - \widetilde{W}^F(s)}{s\left(E[W^F] - E[W_j^{PF_m}]\right)},$$

under Assumption 2.1. Then, under Assumption 2.1, we have for $i = 1, 2$,

$$E\left[C_i'(U_j^{PF_j})\right] = \frac{2k_i(1-\rho_j)(1-\rho)\left(1-\tilde{S}_{3-j}(-h_i)\right)\left[\rho h_i + \lambda\left(1-\tilde{S}(-h_i)\right)\right]}{\lambda\bar{\xi}\tau_{3-j}\left(-h_i - \lambda + \lambda\tilde{S}(-h_i)\right)\left(-h_i - \lambda p_j + \lambda p_j \tilde{S}_j(-h_i)\right)}, \tag{2.17}$$

$$E\left[C_i'(U_j^{PF_{3-j}})\right] = \frac{2k_i(1-\rho_{3-j})(1-\rho)^2\left[-h_i\left(1-\tilde{S}\left(f_{3-j}(-h_i)\right)\right) - f_{3-j}(-h_i)\left(1-\tilde{S}(-h_i)\right)\right]}{\lambda\bar{\xi}p_{3-j}\tau_{3-j}\left[-h_i - \lambda + \lambda\tilde{S}(-h_i)\right]\left[f_{3-j}(-h_i) - \lambda + \lambda\tilde{S}\left(f_{3-j}(-h_i)\right)\right]}, \tag{2.18}$$

where $f_{3-j}(s) = \lambda p_{3-j}(1 - B_{3-j}(s)) + s$. (The derivation of Equations (2.17) and (2.18) are given in Appendix.)

If the cost functions for both types are of the form 2.15 with parameters k_i and h_i for $i = 1, 2$, we can apply Theorem 2.1 and Corollary 2.1 to characterize the best policy among PF_1, PF_2 and F. For $i = 1, 2$, we can obtain a_i and b_{3-i} in Theorem 2.1 and Corollaries 2.1 and 2.2 by plugging $E\left[C_i'(U_i^{PF_i})\right]$

and $E\left[C_i'(U_i^{PF_{3-i}})\right]$ from (2.17) and (2.18):

$$a_i = \frac{2k_i(1-\rho_i)(1-\rho)\left(1-\tilde{S}_{3-i}(-h_i)\right)\left[\rho h_i + \lambda\left(1-\tilde{S}(-h_i)\right)\right]}{\lambda\bar{\xi}\tau_{3-i}\tau_i\left(-h_i-\lambda+\lambda\tilde{S}(-h_i)\right)\left(-h_i-\lambda p_i+\lambda p_i\tilde{S}_i(-h_i)\right)},$$

$$b_{3-i} = \frac{2k_i(1-\rho_{3-i})(1-\rho)^2\left[-h_i\left(1-\tilde{S}\left(f_{3-j}(-h_i)\right)\right)-f_{3-j}(-h_i)\left(1-\tilde{S}(-h_i)\right)\right]}{\lambda\bar{\xi}p_{3-i}\tau_{3-i}\tau_i\left[-h_i-\lambda+\lambda\tilde{S}(-h_i)\right]\left[f_{3-j}(-h_i)-\lambda+\lambda\tilde{S}\left(f_{3-j}(-h_i)\right)\right]}.$$

If we assume only the cost function for type 2 customers has the convex exponential form, then we can apply Corollary 2.3 and compute α and β as:

$$\alpha = \frac{p_2\left(1-\tilde{S}_1(-h_2)\right)\left[\rho h_2+\lambda\left(1-\tilde{S}(-h_2)\right)\right]\left[f_2(-h_2)-\lambda+\lambda\tilde{S}\left(f_2(-h_2)\right)\right]}{(1-\rho)\left(-h_2-\lambda p_2+\lambda p_2\tilde{S}_2(-h_2)\right)\left[-h_2\left(1-\tilde{S}\left(f_2(-h_2)\right)\right)-f_2(-h_2)\left(1-\tilde{S}(-h_2)\right)\right]},$$

$$\beta = \frac{(1-\rho)\left(-h_2-\lambda p_1+\lambda p_1\tilde{S}_1(-h_2)\right)\left[-h_2\left(1-\tilde{S}\left(f_1(-h_2)\right)\right)-f_1(-h_2)\left(1-\tilde{S}(-h_2)\right)\right]}{p_1\left(1-\tilde{S}_2(-h_2)\right)\left[\rho h_1+\lambda\left(1-\tilde{S}(-h_2)\right)\right]\left[f_1(-h_2)-\lambda+\lambda\tilde{S}\left(f_1(-h_2)\right)\right]}.$$

To simplify these expressions and gain some insights, we next consider the case when the service times are i.i.d. with exponential distribution with mean τ for all customers.

2.5.2 Exponential service times with identical rates for both types

Under the assumption that all service times are i.i.d. exponentially distributed with mean τ, we have $\tilde{S}_i(s) = \tilde{S}(s) = \mu/(\mu+s)$ for $s > -\mu$ and $i = 1, 2$, where $\mu = 1/\tau$, and $\bar{\xi}\tau^2 = 2/\mu^4$. Then, from Lemma A.2 in Appendix A, we have

$$\widetilde{W}_i^{PF_i}(s) = \frac{1-\rho+\frac{\lambda p_{3-i}}{\mu+s}}{1-\frac{\lambda p_i}{\mu+s}} = \frac{(1-\rho)(\mu+s)+\lambda p_{3-i}}{\mu+s-\lambda p_i}, \tag{2.19}$$

$$\widetilde{W}^F(s) = \frac{1-\rho}{1-\frac{\lambda}{s}\left(1-\tilde{S}(s)\right)} = \frac{1-\rho}{1-\frac{\lambda}{\mu+s}} = \frac{(1-\rho)(\mu+s)}{\mu+s-\lambda}, \tag{2.20}$$

$$\widetilde{W}_{3-i}^{PF_i}(s) = \widetilde{W}^F\big(\lambda p_i(1-B_i(s))+s\big). \tag{2.21}$$

where $B_i(s) = \left(\mu+s+\lambda p_i - \sqrt{(\mu+s+\lambda p_i)^2-4\lambda p_i\mu}\right)/(2\lambda p_i)$. Note, although the LSTs $\widetilde{W}_i^{\pi}$ in Lemma A.2, and hence (2.19), (2.20) and (2.21) are given for $s > 0$, these expressions should also hold for certain negative values of s. Next, we look at the upper bounds on $h_i > 0$ that could ensure $\widetilde{W}_i^{\pi}(-h_i)$ exist.

Proposition 2.6. *When the service times are identical exponentially distributed, $\widetilde{W}_i^{\pi}(-h_i)$ exists (and hence Assumption 2.1 holds for type i) if*

$$\begin{cases} h_i \leq \mu(1-\sqrt{\rho_{3-i}})^2 & \text{if } p_i \leq 1-\rho, \\ h_i < \mu p_i(1-\rho), & \text{otherwise.} \end{cases} \tag{2.22}$$

As ρ increases, the upper bound decreases. When $\rho \to 1$, the upper bound on h_i goes to 0. Under heavy traffic, the expected cost for type i customers would be finite only for very small values of h_i when $C_i(t)$ has the form of (2.15).

When both cost functions are of the form (2.15), and $\widetilde{W}_i^{\pi}$ exists for both $i \in \{1,2\}$, we have,

$$\begin{aligned} a_i &= \frac{2k_i(1-\rho_i)(1-\rho)\left(\frac{-h_i}{\mu-h_i}\right)\left[\rho h_i + \frac{-\lambda h_i}{\mu-h_i}\right]}{\lambda\bar{\xi}\tau_{3-i}\tau_i\left(-h_i - \frac{-\lambda h_i}{\mu-h_i}\right)\left(-h_i - \frac{-\lambda p_i h_i}{\mu-h_i}\right)} = \frac{\mu^4 k_i(1-\rho_i)(1-\rho)\left(\frac{1}{\mu-h_i}\right)\left[-\frac{1}{\mu} + \frac{1}{\mu-h_i}\right]}{\left(1-\frac{\lambda}{\mu-h_i}\right)\left(1-\frac{\lambda p_i}{\mu-h_i}\right)} \\ &= \frac{\mu^3 k_i(1-\rho_i)(1-\rho)h_i}{(\mu-h_i-\lambda)(\mu-h_i-\lambda p_i)}, \end{aligned}$$

and

$$\begin{aligned} b_{3-i} &= \frac{2k_i(1-\rho_{3-i})(1-\rho)^2\left[-h_i\left(\frac{f_{3-i}(-h_i)}{\mu+f_{3-i}(-h_i)}\right) - f_{3-i}(-h_i)\left(\frac{-h_i}{\mu-h_i}\right)\right]}{\lambda\bar{\xi}p_{3-i}\tau_{3-i}\tau_i\left[f_{3-i}(-h_i) - \lambda\left(\frac{f_{3-i}(-h_i)}{\mu+f_{3-i}(-h_i)}\right)\right]\left[-h_i - \lambda\left(\frac{-h_i}{\mu-h_i}\right)\right]} \\ &= \frac{\mu^4 k_i(1-\rho_{3-i})(1-\rho)^2(-h_i - f_{3-i}(-h_i))}{\lambda p_{3-i}(\mu + f_{3-i}(-h_i) - \lambda)(\mu - h_i - \lambda)} \\ &= \frac{\mu^4 k_i(1-\rho_{3-i})(1-\rho)^2(B_{3-i}(-h_i)-1)}{(\mu - \lambda - h_i - \lambda p_{3-i}(B_{3-i}(-h_i)-1))(\mu-h_i-\lambda)}. \\ &= \frac{\mu^4 k_i(1-\rho_{3-i})(1-\rho)^2}{\left(\frac{\mu-h_i-\lambda}{B_{3-i}(-h_i)-1} - \lambda p_{3-i}\right)(\mu-h_i-\lambda)}. \end{aligned}$$

a_i and b_{3-i} are both positive when h_2 satisfies Proposition 2.6. The complicated expression of $B_{3-i}(-h_i)$ makes it difficult to compare the values of a_i and b_{3-i} analytically. Then, we conduct a numerical comparison of a_2 and b_1 where we set $\mu = 1$, and we fixed the value of h_2 to satisfy the conditions in Proposition 2.6 for all possible values of p_2 and ρ ($h_2 = 0.0028$). We plot the values of a_2 and b_1 with respect to ρ and p_2, respectively.

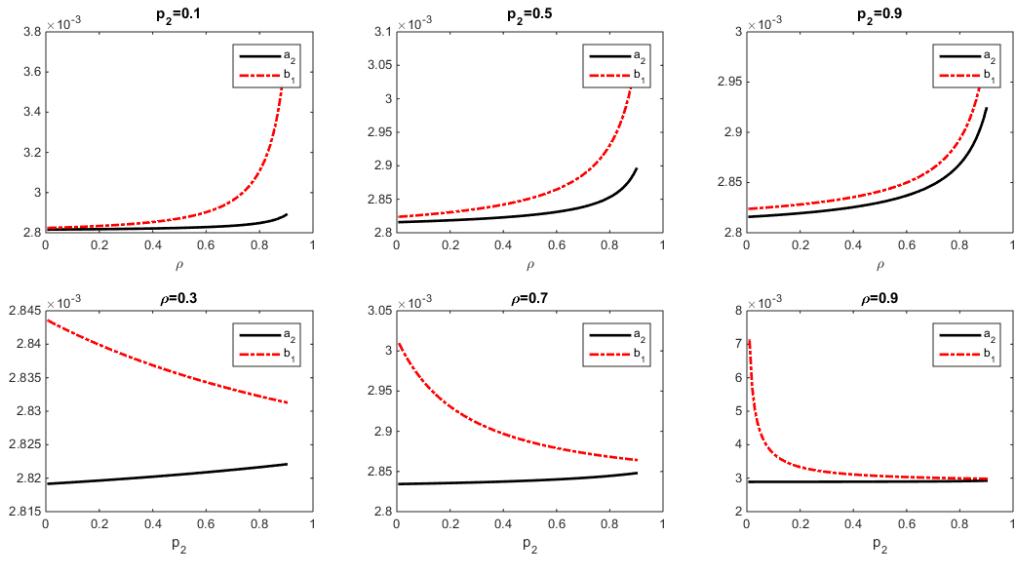

Figure 2.3: Plots of a_2 and b_1 with respect to ρ and p_2

Although we only provide plots of a_2 and b_1 above, the plots for a_1 and b_2 will be symmetric with respect to p_1 and ρ and h_1. From the plots, we find that $a_i \leq b_{3-i}$, and they both increases as ρ increases, and a_i increases and b_{3-i} decreases as p_i increases. The next plot shows how a_2 and b_1 changes in h_2, where we fixed $p_1 = p_2 = 0.5$.

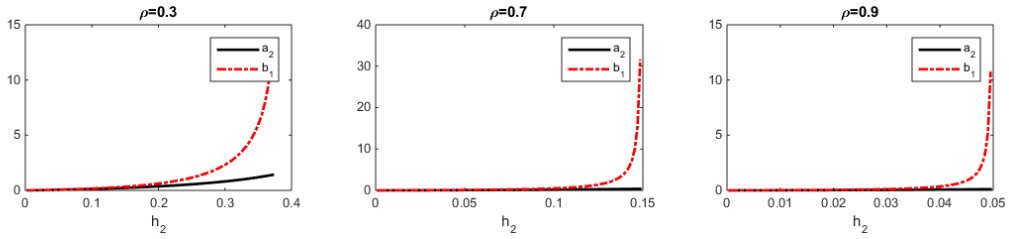

Figure 2.4: Plots of a_2 and b_1 with respect to h_2

With the facts that plots of a_1 and b_2 are symmetric with Figure 2.4, we find that as h_i goes to the upper bound given in Proposition 2.6, a_i is finite and b_{3-i} goes to infinity. Hence, if h_i is close to the upper bound, then $a_{3-i} \leq b_{3-i}$, and hence PF_{3-i} will be dominated by F from Theorem 2.1 (a). The intuition behind this result is that if type i customers have an exponential cost function with parameters near the upper bound, then prioritize type $3-i$ will make type i customers wait longer and hence is worse than F. If both h_1 and h_2 are close to their bounds, then F will be the best policy of the three from Theorem 2.1.

From Corollary 2.2, when type 2 has the exponential convex form we can compute α and β as follows:

$$\begin{aligned}\alpha =&\frac{p_2\left(1-\tilde{S}_1(-h_2)\right)\left[\rho h_2+\lambda\left(1-\tilde{S}(-h_2)\right)\right]\left[f_2(-h_2)-\lambda+\lambda\tilde{S}\left(f_2(-h_2)\right)\right]}{(1-\rho)\left(-h_2-\lambda p_2+\lambda p_2\tilde{S}_2(-h_2)\right)\left[-h_2\left(1-\tilde{S}\left(f_2(-h_2)\right)\right)-f_2(-h_2)\left(1-\tilde{S}(-h_2)\right)\right]}\\ =&\frac{p_2\frac{-h_2}{\mu-h_2}\left[\rho h_2+\lambda\frac{-h_2}{\mu-h_2}\right]\left[f_2(-h_2)-\lambda\frac{f_2(-h_2)}{\mu+f_2(-h_2)}\right]}{(1-\rho)\left(-h_2-\lambda p_2\frac{-h_2}{\mu-h_2}\right)\left[-h_2\frac{f_2(-h_2)}{\mu+f_2(-h_2)}-f_2(-h_2)\frac{-h_2}{\mu-h_2}\right]}\\ =&\frac{p_2\rho h_2\left[\mu+f_2(-h_2)-\lambda\right]}{(1-\rho)\left(\mu-h_2-\lambda p_2\right)\left(-h_2-f_2(-h_2)\right)}\\ =&\frac{p_2\rho h_2\left[\mu-\lambda-h_2-\lambda p_2(B_2(-h_2)-1)\right]}{(1-\rho)\left(\mu-h_2-\lambda p_2\right)\lambda p_2(B_2(-h_2)-1)}=\frac{h_2\left[\frac{\mu-\lambda-h_2}{B_2(-h_2)-1}-\lambda p_2\right]}{\mu(1-\rho)\left(\mu-h_2-\lambda p_2\right)},\end{aligned}$$

and similarly,

$$\beta=\frac{\mu(1-\rho)\left(\mu-h_2-\lambda p_1\right)}{h_2\left[\frac{\mu-\lambda-h_2}{B_1(-h_2)-1}-\lambda p_1\right]}.$$

When computing α and β, we assume h_2 is such that $\widetilde{W}_1^{PF_2}(-h_2)$ and $\widetilde{W}_2^{PF_1}(-h_2)$ exist. Then, from Proposition 2.6 we can obtain the bound on h_2.

2.6 Convex cost functions

In this section, we focus on convex $C_1(\cdot)$ and $C_2(\cdot)$. (Throughout the paper, we use convexity/concavity in the nonstrict sense.) We first state a result that shows it is sufficient to only compare F and PF_i if $C_i(\cdot)$ is a convex function for both $i=1,2$. Here, let L denotes last come first serve (LCFS) and PL_i denotes the priority policy that prioritizes type i customers and employs LCFS within each type for $i=1,2$.

Let Π_{NP} denote the set of non-idling and non-preemptive queueing policies that only depend on the order of arrival of customers but not on their type, and let Π_{P_i} denote the set of non-idling and non-preemptive priority policies that prioritize type i customers for $i=1,2$. Note that $\Pi=\Pi_{NP}\cup\Pi_{P_1}\cup\Pi_{P_2}$.

Proposition 2.7. *If $C_1(t)$ and $C_2(t)$ are both convex functions, then*

(a) *$C_F\le C_\pi\le C_L$ for any $\pi\in\Pi_{NP}$;*

(b) *$C_{PF_i}\le C_\pi\le C_{PL_i}$ for any $\pi\in\Pi_{P_i}$ and fixed $i=1,2$.*

Proposition 2.7 imply that when the cost functions for both types are convex, it is sufficient to consider Π_{cx}, instead of Π. Part (a) of Proposition 2.7 follows directly from Theorem 2 in Vasicek (1977). The first inequality in part (b) follows from Theorem 1 in Haji and Newell (1971). One can use the same argument used in the proof of Theorem 1 in Haji and Newell (1971) to prove the second inequality in part (b) of Proposition 2.7.

In Corollary 2.3, we use "$\max\{\alpha, \beta\}$" and "$\min\{\alpha, \beta\}$" since we do not know their order under general settings. However, when the service times are identical for both types of customers we have the following result.

Proposition 2.8. *Suppose the service times are i.i.d. for both types of customers and $C_2(\cdot)$ is convex, then $\alpha \leq 1 \leq \beta$.*

2.7 Concave cost functions

If $C_1(\cdot)$ and $C_2(\cdot)$ are both concave, then we can have the following result.

Corollary 2.4. *If $C_1(t)$ and $C_2(t)$ are both concave functions, then the inequalities in Proposition 2.7 hold in the opposite direction.*

Corollary 2.4 follows from Proposition 2.7 using the fact that if $C_i(t)$ is concave for $i = 1, 2$, then $-C_i(t)$ is convex.

2.7.1 Comparison of LCFS and fixed priority policies

The order of service within each type of customer will not affect the expected waiting time, and hence we have $E[W^L] = E[W^F]$ and $E[W_j^{PL_i}] = E[W_j^{PF_i}]$ for $i, j \in \{1, 2\}$, where the expectations have been given in Lemma 2.2.

Lemma 2.5. *For fixed $i = 1, 2$, $W_i^{PL_i} \leq_{st} W^L \leq_{st} W_{3-i}^{PL_i}$.*

Then, the following random variables are well-defined based on Lemma 2.5.

$$U_i^{PL_i} \equiv \Psi(W_i^{PL_i}, W^L),\ U_{3-i}^{PL_i} \equiv \Psi(W^L, W_{3-i}^{PL_i}).$$

Then, we obtain similar results as Proposition 2.1 by comparing policies in $\Pi_{cv} = \{L, PL_1, PL_2\}$.

Theorem 2.2.

(a) For $i = 1, 2$, $C_L \leq C_{PL_i}$ if and only if $a_i^L \leq b_i^L$ where

$$a_i^L = \frac{E\big[C_i'(U_i^{PL_i})\big]}{\tau_i},\ b_i^L = \frac{E\big[C_{3-i}'(U_{3-i}^{PL_i})\big]}{\tau_{3-i}}. \tag{2.23}$$

(b) $C_{PL_1} \leq C_{PL_2}$ if and only if $(1-\rho_1)(a_2^L - b_2^L) \leq (1-\rho_2)(a_1^L - b_1^L)$.

Corollary 2.5.

(a) If $C_1'(t) \geq \tau_1 \max\{a_2^L, b_1^L\}$ *for all* $t \geq 0$, *then* $C_{PL_1} \leq C_L \leq C_{PL_2}$.

(b) If $C_1'(t) \leq \tau_1 \min\{a_2^L, b_1^L\}$ *for all* $t \geq 0$, *then* $C_{PL_2} \leq C_L \leq C_{PL_1}$.

(c) If $\tau_1 a_2^L \leq C_1'(t) \leq \tau_1 b_1^L$ *for all* $t \geq 0$, *then* $C_L \leq C_{PL_1}$ *and* $C_L \leq C_{PL_2}$

Corollary 2.6. *If* $E[C_2'(U_1^{PL_2})] \neq 0$ *and* $E[C_2'(U_1^{PL_1})] \neq 0$, *define*

$$\alpha^L \equiv \frac{\tau_1 E[C_2'(U_2^{PL_2})]}{\tau_2 E[C_2'(U_1^{PL_2})]} \quad \textit{and} \quad \beta^L \equiv \frac{\tau_1 E[C_2'(U_2^{PL_1})]}{\tau_2 E[C_2'(U_1^{PL_1})]}.$$

(a) If $C_1'(t) \geq \max\{\alpha^L, \beta^L\} C_2'(t)$ *for all* $t \geq 0$, *then* $C_{PL_1} \leq C_L \leq C_{PL_2}$.

(b) If $C_1'(t) \leq \min\{\alpha^L, \beta^L\} C_2'(t)$ *for all* $t \geq 0$, *then* $C_{PL_2} \leq C_L \leq C_{PL_1}$.

(c) If $\alpha^L C_2'(t) \leq C_1'(t) \leq \beta^L C_2'(t)$ *for all* $t \geq 0$, *then* $C_L \leq C_{PL_1}$ *and* $C_L \leq C_{PL_2}$.

Proposition 2.9. *Suppose the service times are i.i.d. for both types of customers and* $C_2(\cdot)$ *is concave, then* $\alpha^L \geq 1 \geq \beta^L$.

The proofs of the above results are similar to Proposition 2.1 and Corollary 2.3, and hence are omitted.

2.7.2 Waiting time LSTs

Before we apply results from the previous section to compare C_{PL_1}, C_{PL_2} and C_L, we first establish the LST of the waiting times under policy L and PL_i. As far as we know, the LST under PL_i has not been studied in literature.

We assume the LST of the service time distribution for type i customers is $\tilde{S}_i(s)$ for $i = 1, 2$, and let $\tilde{S}(s) = p_1 \tilde{S}_1(s) + p_2 \tilde{S}_2(s)$ be the LST of the general service time.

We define a "busy period" of the M/G/1 queue as the time until the server becomes idle for the first time, starting when one customer enters an empty system. Then, according to Wishart (1960), when $\rho < 1$, the LST of the busy period, denoted by $B(s)$, is given by the unique solution to

$$B(s) = \tilde{S}\big(s + \lambda - \lambda B(s)\big).$$

The LST of the steady-state waiting times in an M/G/1 queue under LCFS policy, denoted by $\widetilde{W}^L(s)$, can be obtained as follows.

Lemma 2.6. *(E.g., Wishart (1960)) In an M/G/1 queueing sysem with LCFS queueing discipline, the LST of steady-state waiting times is given by*

$$\widetilde{W}^L(s) \quad = \quad 1 - \rho + \frac{\lambda - \lambda B(s)}{s + \lambda - \lambda B(s)}. \tag{2.24}$$

Let $B_i(s)$ denote the busy period of the M/G/1 queue with only type i arrivals, which is given by the unique solution to

$$B_i(s) = \tilde{S}_i\big(s + \lambda p_i - \lambda p_i B_i(s)\big).$$

Then, we will compute $\widetilde{W}_i^{PL_i}(s)$ and $\widetilde{W}_{3-i}^{PL_i}(s)$, which is the respective LST of $W_i^{PL_i}(s)$ and $W_i^{PL_i}(s)$, for $i \in \{1, 2\}$.

Lemma 2.7. *In an M/G/1 queue under queueing discipline PL_i, the LST of the waiting time for type i customers is given by*

$$\widetilde{W}_i^{PL_i} = 1 - \rho + \frac{\lambda\big[1 - \tilde{S}(s + \lambda p_i - \lambda p_i B_i(s))\big]}{s + \lambda p_i - \lambda p_i B_i(s)},$$

and the LST of the waiting time for type $3 - i$ customers is given by

$$\widetilde{W}_{3-i}^{PL_i} = \frac{(1-\rho)(s + \lambda - \lambda B(s)) + \lambda p_{3-i}\left[1 - \tilde{S}_{3-i}(s + \lambda - \lambda B(s))\right]}{s + \lambda - \lambda B(s) - \lambda p_i\left[1 - \tilde{S}_i(s + \lambda - \lambda B(s))\right]}.$$

The proof of Lemma 2.7 is given in Appendix A.

Letting $g(x) = -e^{-sx}/s$ for some real s in Lemma 2.1 we have, the LST of $U_i^{PL_j}$, denoted by $\widetilde{U}_i^{PL_j}$ can be computed as

$$\widetilde{U}_i^{PL_j}(s) = E\left[g'\left(U_i^{PL_j}(s)\right)\right] = \frac{\widetilde{W}_i^{PL_j}(s) - \widetilde{W}^L(s)}{s\left(E\left[W^L\right] - E\left[W_i^{PL_j}\right]\right)}.$$

With the knowledge of $\widetilde{U}_i^{PL_j}(s)$, we can compute a_i^L and b_i^L under certain cost functions. We will analyze a special form of function in Section 2.7.3.

2.7.3 Exponential cost function

We assume that $C_2(t) = \alpha_2(1 - e^{-h_2 t})$ for $\alpha_2 > 0$ and $h_2 > 0$, then $C_2'(t) = \alpha_2 h_2 e^{-h_2 t}$. Then we apply results from Section 2.7.1 to compare C_{PL_1}, C_{PL_2} and C_L. First, we compute a_2^L and b_1^L as

$$a_2^L = \frac{1}{\tau_2} E\left[C_2'\left(U_2^{PL_2}\right)\right] = \frac{\alpha_2 h_2}{\tau_2}\widetilde{U}_2^{PL_2}(h_2),\ b_1^L = \frac{1}{\tau_2} E\left[C_2'\left(U_2^{PL_1}\right)\right] = \frac{\alpha_2 h_2}{\tau_2}\widetilde{U}_2^{PL_1}(h_2).$$

Applying Corollary 2.5, we can order C_{PL_1}, C_{PL_2} and C_L by comparing $C_1'(t)$ with a_2 and b_1. Next we consider the special case when the service times for both types of customers are exponentially distributed with mean τ, to obtain close-form expressions for a_2 and b_1. In this case, $\tilde{S}_1(s) = \tilde{S}_2(s) = \tilde{S}(s) = \tau^{-1}/(\tau^{-1} + s)$,

and hence

$$B_i(s) = \frac{\tau^{-1}}{\tau^{-1} + s + \lambda p_i - \lambda p_i B_i(s)} \Leftrightarrow s + \lambda p_i - \lambda p_i B_i(s) = \frac{\tau^{-1}}{B_i(s)} - \tau^{-1} = \frac{1 - B_i(s)}{\tau B_i(s)}.$$

Then,

$$\widetilde{W}_i^{PL_i} = 1 - \rho + \frac{\lambda[1 - B_i(s)]}{[1 - B_i(s)]/[\tau B_i(s)]} = 1 - \rho + \rho B_i(s),$$

and similarly, $s + \lambda - \lambda B(s) = [1 - B(s)]/[\tau B(s)]$, and hence

$$\begin{aligned}\widetilde{W}_{3-i}^{PL_i} &= \frac{(1-\rho)[1 - B(s)]/[\tau B(s)] + \lambda p_{3-i}\,[1 - B(s)]}{[1 - B(s)]/[\tau B(s)] - \lambda p_i\,[1 - B(s)]} = \frac{1 - \rho + \rho_{3-i} B(s)}{1 - \rho_i B(s)} \\ &= \frac{1 - \rho + \rho B(s) - \rho_i B(s)}{1 - \rho_i B(s)} = 1 - \rho\left(\frac{1 - B(s)}{1 - \rho_i B(s)}\right),\end{aligned}$$

and

$$\widetilde{W}^L(s) = 1 - \rho + \frac{\lambda[1 - B(s)]}{[1 - B(s)]/[\tau B(s)]} = 1 - \rho + \rho B(s).$$

where $$B_i(s) = \left(1 + \tau s + \rho_i - \sqrt{(1 + \tau s + \rho_i)^2 - 4\rho_i}\right) / (2\rho_i)$$

and $$B(s) = \left(1 + \tau s + \rho - \sqrt{(1 + \tau s + \rho)^2 - 4\rho}\right) /(2\rho).$$

Then,

$$\begin{aligned}\widetilde{W}_i^{PL_i}(s) - \widetilde{W}^L(s) &= \rho[B_i(s) - B(s)], \\ \widetilde{W}^L(s) - \widetilde{W}_{3-i}^{PL_i}(s) &= \rho[1 - B(s)]\left[\frac{1}{1 - \rho_i B(s)} - 1\right] = \frac{\rho_i \rho B(s)[1 - B(s)]}{1 - \rho_i B(s)}\end{aligned}$$

Then,

$$a_2^L = \left(\frac{\alpha_2 h_2}{\tau}\right)\left(\frac{\widetilde{W}_2^{PL_2}(h_2) - \widetilde{W}^L(h_2)}{h_2\left(E\left[W^L\right] - E\left[W_2^{PL_2}\right]\right)}\right) = \left(\frac{\alpha_2(1-\rho_2)(1-\rho)}{p_1\lambda\tau^3}\right)[B_2(s) - B(s)],$$

$$b_1^L = \left(\frac{\alpha_2 h_2}{\tau}\right)\left(\frac{\widetilde{W}^L(h_2) - \widetilde{W}_2^{PL_1}(h_2)}{h_2\left(E\left[W_2^{PL_1}\right] - E\left[W^L\right]\right)}\right) = \left(\frac{\alpha_2(1-\rho_1)(1-\rho)}{p_1\lambda\tau^3}\right)\left(\frac{\rho_1 B(s)[1 - B(s)]}{1 - \rho_1 B(s)}\right)$$

We can also apply Corollary 2.6 and compute α^L and β^L as:

$$\begin{aligned}\alpha^L &= \left(\frac{p_2}{p_1}\right)\left(\frac{[B_2(h_2)-B(h_2)]\,[1-\rho_2 B(h_2)]}{\rho_2 B(h_2)[1-B(h_2)]}\right) = \frac{[B_2(h_2)-B(h_2)]\,[1-\rho_2 B(h_2)]}{\rho_1 B(h_2)[1-B(h_2)]},\\ \beta^L &= \left(\frac{p_2}{p_1}\right)\left(\frac{\rho_1 B(h_2)[1-B(h_2)]}{[B_1(h_2)-B(h_2)]\,[1-\rho_1 B(h_2)]}\right) = \frac{\rho_2 B(h_2)[1-B(h_2)]}{[B_1(h_2)-B(h_2)][1-\rho_1 B(h_2)]}.\end{aligned}$$

2.8 Simulation results

We conduct a simulation study to compare the performances of the optimal policies within Π with the performance of applying the generalized $c\mu$ rule under different cost functions. We have the analytic expressions in the previous sections for the long-run average cost of policies in Π_{cx} and Π_{cv}, and we use simulation to obtain a 95% confidence interval on the long-run average cost of the generalized $c\mu$ rule.

We use Arena13 simulation software, in which we run 10 replications of each simulation with simulating time 50000 minutes and truncated the first 5000 minutes based on a warm-up period analysis. We consider the arrival process is a Poisson process with rate $\lambda \in \{0.3, 0.7, 0.9\}$, the proportion of type 1 patients $p_1 \in \{0.1, 0.5, 0.9\}$. We assume the service times are exponentially distributed, and we first consider the scenarios that all customers have the same service rates, and then we consider the scenarios when service rates are different for different types.

We consider two types of cost functions. First we consider convex cost functions in quadratic form, where we assume that $C_1(t) = kt^2$ and $C_2(t) = t^2$ for different values of k, and then we consider concave cost functions in exponential form, where we assume that $C_1(t) = h(1-e^{-t})$ and $C_2(t) = 1-e^{-t}$ for different values of h.

2.8.1 Same service rates

We consider an M/M/1 queue with the same mean service time $\tau = 1$ for both types.

2.8.1.1 Quadratic cost functions

We consider 27 different scenarios generated by all combinations of $\lambda \in \{0.3, 0.7, 0.9\}$, $p_1 \in \{0.1, 0.5, 0.9\}$ and $k \in \{0.1, 0.9, 5\}$. From Proposition 2.1, we can compute the values of A and B for each combination of λ and p_1 as follows.

Table 2.1: The threshold values (A and B) to characterize the best policy among F, PF_1 and PF_2.

λ	0.3			0.7			0.9		
p_1	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
A	0.532	0.580	0.620	0.307	0.450	0.546	0.169	0.375	0.500
B	1.612	1.725	1.880	1.831	2.223	3.255	2.000	2.664	5.918

We choose the policy with the smallest cost among PF_1, PF_2 and $FCFS$, which can be achieved by comparing the value of k with A and B. To be more specific, PF_2 has the smallest cost if $k < A$; PF_1 has the smallest cost if $k > B$ and F has the smallest cost if $A < k < B$. We find that for all combinations of λ and p_1, PF_2 is the best policy when $k = 0.1$ and F is the best when $k = 0.9$. When $k = 6$, PF_1 is the best when λ takes values 0.3 and 0.7 and F is the best when $\lambda = 0.9$ (as we have illustrated in Section 2.4.1 that F becomes more preferable as λ increases). Next, we compare the smallest cost from the three static policies with the cost of the generalized $c\mu$ rule, which can be obtained from the simulation. In the comparison, if the confidence intervals of the generalized $c\mu$ rule does not contain the value of the cost of the best static policy, then we confirmed that there is a statistical difference between these two policies at a significance level of 95% and the difference is in favor of the policy with the smaller performance.

Table 2.2: Compare the best static policy with G-$c\mu$ rule under quadratic cost

		$k = 5$		$k = 0.9$		$k = 0.1$	
ρ	p_1	Best static	G-$c\mu$ C.I.	Best static	G-$c\mu$ C.I.	Best static	G-$c\mu$ C.I.
0.3	0.1	1.52	(1.53 ± 0.06)	1.21	(1.21 ± 0.05)	1.04	(1.05 ± 0.01)
	0.5	3.03	(3.08 ± 0.11)	1.16	(1.16 ± 0.04)	0.51	(0.52 ± 0.02)
	0.9	5.36	(5.36 ± 0.20)	1.11	(1.11 ± 0.04)	0.18	(0.19 ± 0.04)
0.7	0.1	17.36	(17.36 ± 0.92)	15.40	(15.36 ± 0.85)	10.92	(11.03 ± 0.10)
	0.5	29.67	(29.49 ± 1.53)	14.78	(14.72 ± 0.8)	3.80	(3.91 ± 0.17)
	0.9	63.19	(60.96 ± 3.29)	14.16	(14.13 ± 0.77)	1.82	(1.84 ± 0.51)
0.9	0.1	198.65	(194 ± 18.97)	178.20	(175.8 ± 17.44)	**115.99**	**(103 ± 1.97)**
	0.5	**336.74**	**(300.4 ± 29.34)**	171.00	(168.5 ± 16.72)	35.16	(34.32 ± 3.23)
	0.9	**828.00**	**(642.8 ± 61.04)**	163.80	(161.7 ± 16.05)	19.97	(19.79 ± 9.96)

The generalized $c\mu$ rule is asymptotically optimal under heavy traffic. From Table 2.2, we find that the differences between the best static policy we derived and the generalized $c\mu$ rule are statistically insignificant in most scenarios. We noticed that when the traffic intensity is high, the proportion of the priority group is large and the difference in costs between two types of customers are large, the generalized $c\mu$ rule performs

significantly better than the best statical policy. We also observe that there is one scenario that the best static policy performs significantly better than the generalized $c\mu$ rule when the traffic intensity is not high.

When $p_1 = 0.9$ and $k = 5$, we find that the best static policy changes from PF_1 to F as λ increases from 0.7 to 0.9, which verifies the fact that we incline to employ F rather than a static priority policy under heavy traffic. However, F is not performing as well as the dynamic priority policy in this scenario.

2.8.1.2 Exponential cost functions

We use the same model as the previous section with different cost functions. By computing the value of parameters in Corollary 2.6, we find that when $h = 5$, PL_1 is the best policy in Π, when $h = 0.9$, LCFS is the best policy in Π and when $h = 0.1$, PL_2 is the best policy in Π. We computed the cost for these three static policies numerically, then compare the cost of the best static policy with the 95% confidence interval of the cost of the generalized $c\mu$ rule conducted by simulation. As shown in Table 2.3, under most scenarios the performance difference of the best static policy and the generalized $c\mu$ rule is insignificant at level 0.05. There are four scenarios when there is a significant difference, which happens when systems have medium or high traffic intensity and when the proportion of the priority group is medium or large. Even though there is a significant difference in these scenarios, we can see that the difference is less than 3% for all scenarios.

Table 2.3: Compare the best static policy with G-$c\mu$ rule under exponential cost (10^{-2})

		$h = 5$		$h = 0.9$		$h = 0.1$	
ρ	p_1	PL_1	G-$c\mu$ C.I.	$LCFS$	G-$c\mu$ C.I.	$PL2$	G-$c\mu$ C.I.
0.30	0.10	22.27	(21.45, 22.78)	15.96	(15.51, 16.29)	14.59	(14.20, 14.89)
	0.50	47.56	(46.09, 48.28)	15.31	(14.87, 15.64)	8.65	(8.38, 8.89)
	0.90	73.88	(71.58, 75.34)	14.67	(14.25, 14.97)	2.98	(2.91, 3.05)
0.70	0.10	55.76	(54.45, 55.80)	40.54	(39.91, 40.68)	36.90	(36.32, 37.00)
	0.50	119.05	(116.46, 118.82)	38.90	(38.28, 39.00)	**21.42**	**(20.97, 21.37)**
	0.90	187.26	(183.76, 187.26)	37.26	(36.67, 37.40)	7.36	(7.25, 7.46)
0.90	0.10	74.00	(72.80, 74.30)	54.11	(53.49, 54.46)	49.19	(48.59, 49.37)
	0.50	**158.20**	**(154.58, 157.61)**	51.93	(51.35, 52.20)	28.32	(27.91, 28.38)
	0.90	**249.97**	**(245.89, 249.96)**	49.74	(49.21, 50.09)	9.70	(9.60, 9.84)

2.8.2 Different service rates

In this section, we assume that the service times are exponentially distributed with different rates. Assume the service rate for type 2 customers is 1, and for type 1 customers is τ. We consider the system traffic $\rho \in \{0.3, 0.7, 0.9\}$ and type 1 proportion $p_1 \in \{0.1, 0.5, 0.9\}$, and then computed $\lambda = \rho/(p_1\tau + 1 - p_1)$ for each scenario. Note that for fixed p_1 and τ, the change of ρ can only be resulted by changing λ.

2.8.2.1 Quadratic cost functions

We consider $\tau = 5$ and $\tau = 0.2$, and for each case we first compute the values of A and B from Proposition 2.1 to determine the smallest cost obtained from the three policies PF_1, PF_2 and F, then we compare the smallest cost with the cost of the generalized $c\mu$ rule obtained from the simulation.

When $\tau = 5$ and $\tau = 0.2$, the values of A and B are computed and shown in Table 2.4.

Table 2.4: The threshold values (A and B) for different service rates.

		$\tau = 5$		$\tau = 0.2$	
ρ	p_1	A	B	A	B
0.3	0.1	0.805	2.558	0.375	1.260
	0.5	0.791	2.618	0.382	1.265
	0.9	0.794	2.670	0.391	1.242
0.7	0.1	0.508	2.551	0.238	1.592
	0.5	0.606	3.470	0.288	1.651
	0.9	0.628	4.202	0.392	1.970
0.9	0.1	0.347	2.560	0.107	1.852
	0.5	0.506	5.000	0.200	1.975
	0.9	0.540	9.313	0.391	2.880

We find that when the service rates are different, A is not always decreasing in λ (see the case $\tau = 0.2$ and $p_1 = 0.9$) and B is not always increasing in λ (see the case $\tau = 5$ and $p_1 = 0.1$). Besides, A and B are not necessarily increasing in p_1 (e.g., when $\tau = 5, \rho = 0.3$ or when $\tau = 0.2, \rho = 0.3$). By the values of A and B, we can identify the best static policy for each scenario. More specifically, we compare k with the values of $A\tau$ and $B\tau$, and PF_2 is the best if $k \leq A\tau$; PF_1 is the best if $k \geq B\tau$ and F is the best if $A\tau \leq k \leq B\tau$. Next, we compare the performance of the best static policy with the generalized $c\mu$ rule.

When $\tau = 5$, we find that PF_2 is the best of PF_1, PF_2 and F for all scenarios when $k = 0.1$ and $k = 0.9$, and F is the best for all scenarios when $k = 5$.

Table 2.5: Compare the best static policy with G-$c\mu$ rule under quadratic costs with $\tau = 5$

		$k = 5$		$k = 0.9$		$k = 0.1$	
ρ	p_1	Best Static †	G-$c\mu$ C.I.	Best Static ⋆	G-$c\mu$ C.I.	Best Static ⋆	G-$c\mu$ C.I.
0.30	0.10	14.52	(15.83 ± 1.82)	8.29	(10.03 ± 1.26)	6.96	(6.80 ± 0.91)
	0.50	74.69	(76.62 ± 8.28)	19.22	(22.33 ± 2.36)	8.09	(8.21 ± 0.65)
	0.90	137.30	(139.20 ± 19.26)	26.00	(26.98 ± 3.63)	4.20	(4.27 ± 0.57)
0.70	0.10	152.44	(162.23 ± 23.47)	61.63	(92.80 ± 13.67)	30.39	(33.64 ± 4.24)
	0.50	907.41	(1155.24 ± 302.80)	193.07	(279.74 ± 63.57)	36.65	(42.34 ± 7.07)
	0.90	1734.51	(1814.01 ± 502.50)	318.72	(321.85 ± 84.97)	38.52	(37.24 ± 8.15)
0.90	0.10	1578.86	(1707.63 ± 657.90)	651.51	(1016.78 ± 408.80)	116.49	(174.73 ± 51.89)
	0.50	10260.00	(15260.33 ± 8015.00)	2161.72	(2364.30 ± 575.70)	260.68	(246.06 ± 64.28)
	0.90	20014.43	(23865.87 ± 11380.00)	3672.34	(3964.99 ± 1620.00)	412.05	(409.11 ± 136.80)

When $\tau = 0.2$, we find that F is the best for all scenarios when $k = 0.1$ and PF_1 is the best for all scenarios when $k = 0.9$ and $k = 5$.

Table 2.6: Compare the best static policy with G-$c\mu$ rule under quadratic costs with $\tau = 0.2$

		$k = 5$		$k = 0.9$		$k = 0.1$	
ρ	p_1	Best Static ⋆	G-$c\mu$ C.I.	Best Static ⋆	G-$c\mu$ C.I.	Best Static †	G-$c\mu$ C.I.
0.30	0.10	1.39	(1.41 ± 0.14)	1.14	(1.17 ± 0.11)	1.09	(1.09 ± 0.11)
	0.50	1.90	(1.95 ± 0.05)	0.80	(0.93 ± 0.04)	0.55	(0.54 ± 0.03)
	0.90	1.43	(1.45 ± 0.06)	0.31	(0.39 ± 0.02)	0.08	(0.09 ± 0.01)
0.70	0.10	14.71	(14.73 ± 1.51)	14.14	(14.55 ± 1.09)	13.73	(13.69 ± 1.46)
	0.50	11.24	(13.27 ± 0.88)	8.44	(12.02 ± 0.80)	6.65	(7.48 ± 0.76)
	0.90	6.86	(9.34 ± 0.48)	2.52	(4.28 ± 0.26)	0.83	(0.89 ± 0.07)
0.90	0.10	163.92	(161.46 ± 29.96)	163.18	(175.76 ± 33.08)	158.38	(182.09 ± 37.50)
	0.50	99.66	(128.23 ± 17.48)	95.88	(138.21 ± 10.36)	75.24	(109.12 ± 16.67)
	0.90	36.67	(65.40 ± 7.36)	28.54	(48.86 ± 6.63)	8.57	(9.02 ± 1.03)

2.9 Conclusions

When the waiting costs are nonlinear functions, the optimal queueing policy would involve dynamically determining the priorities of customers in the system according to the customers' types and their waiting times in the queue. It would be very costly to keep track of the waiting times of all customers in the system at any decision epochs, so our problem arises that whether we can find the best static queueing discipline which either gives priority to a certain type of customers or using no-priority policy.

We first compare the cost under F, PF_1 and PF_2, and demonstrate the comparisons by examples of polynomial and exponential functions. Our results in Section 2.6 shows that when the cost functions are

convex, then it is suffice to compare only F, PF_1 and PF_2 to find the best policy within Π. Similarly we present the comparisons of L, PL_1 and PL_2 and show in Corollary 2.4 that if the cost functions are concave, then the best policy within Π should be one of L, PL_1 and PL_2.

We conduct simulation study to verify the results and to compare the performance of the best static policy with the generalized $c\mu$ rule. The simulations show that the best static policy performs very well under medium and light traffic, or when the proportion of the priority group is small. The generalized $c\mu$ rule performs worse than our best optimal policy under light traffic or under concave cost functions. Actually the worst performance of the generalized $c\mu$ rule happens under the heavy traffic with the concave cost functions.

CHAPTER 3: ALLOCATION OF INTENSIVE CARE UNIT BEDS DURING PERIODS OF HIGH DEMANDS

In this chapter, we consider a stylized, discrete-time model for an ICU in which patients' health conditions change over time according to a Markov chain. Patients are assumed to be in one of the two health stages, and we would like to make admission and discharge decisions s to minimize the mortality rate.

3.1 Introduction

Efficient management of ICU beds has long been a topic of interest in practice as well as academia. Simply put, an ICU bed is a very expensive resource and the number of available ICU beds frequently falls short of the existing demand in many hospitals. Therefore, it is important to make the best use of these beds via intelligent admission and discharge decisions. There is wide agreement that during times of high demand, beds should not be given to patients who have little to benefit from intensive care treatment. However, when it comes to choosing among patients who can potentially benefit from such treatment, there do not appear to be easy answers. Even if one can quantify the ICU benefit at the individual patient level and there is agreement on some utilitarian objective such as maximizing the number of survivors, it is not difficult to see that allocating beds to those with the highest potential to benefit is not necessarily the "right" thing to do. For example, if this potential benefit can only be realized at the expense of a long length of stay, which is likely to prevent the use of the bed for treating other patients, then it is difficult to weigh the "benefits" with the "costs." In short, making patient admission and discharge decisions for a particular patient, especially when overall demand is high, is a complex task that requires careful consideration of not only the health condition of that particular patient in isolation but a collective assessment of the health conditions and operational requirements of all the patients in the ICU as well as the mix of patients the ICU expects to see in the near future. The objective of this chapter is to provide insights into this complex decision problem using mathematical modeling and analysis.

A key factor in deciding whether or not a patient should be admitted to the ICU is the chance of survival for the patient or more generally how much benefit the patient would likely get from being treated in the ICU as opposed to a non-ICU setting like a general ward at the hospital. Thus, quantification of the expected benefits (e.g., change in probability of survival or readmission) given patient health conditions,

comorbidities, and other patient characteristics such as age, gender, etc. is essential. Many studies have contributed to that effort (see, e.g., Sinuff et al. (2004), Shmueli and Sprung (2005), Kim et al. (2014)) and provided some understanding as to what kind of patients would benefit most from ICU treatment. With more research in this area, we will likely have an increasingly more precise quantification of those benefits. The natural question then is how exactly to use this information in making ICU admission and discharge decisions. This chapter contributes to the relatively limited but recent literature that aims to provide an answer to this question. It is important to note that our goal in this chapter is not to develop a highly realistic model and propose a support tool that can be used readily to make decisions but rather to develop a stylized formulation and analyze it with the goal of providing insight into this difficult question.

While our analysis provides insights into "optimal" ICU admission/discharge decisions in general, it is particularly relevant to situations where the ICU experiences a severely-high-demand period that lasts for a number of weeks if not longer. Such conditions would arise for instance in case of an influenza epidemic or pandemic, which would require a significantly increased percentage of the local population to be admitted to ICUs within a relatively short period of time. For example, using the models of the Centers for Disease Control, Christian et al. (2006) estimated the potential impact of a pandemic on the Ontario population and found that over a 6-week period, the demand for ICU beds for influenza patients alone, at its peak, would reach 171% of the existing ICU bed capacity. With such heavy demand sustained over a long period of time, the need to develop a framework for patient triage and prioritization, which aims to "do the greatest good for the greatest number," is clearer. To the best of our knowledge, there is not a formally adopted triage protocol which is meant to be used in case of an influenza epidemic or pandemic except for a protocol developed by Christian et al. (2006) in response to a request of the steering committee of the Ontario Health Plan. According to this protocol, arriving patients are put in one of the four triage classes. Patients classified as "red" are given a higher priority than patients classified as "yellow" while patients in the other two classes are typically not accepted to the ICU. The protocol also recognizes the fact that patients' health conditions would change over time and thus it requires reclassification of each patient at the 48th and the 120th hours after admission. The main tool used for classification is the SOFA (Sequential Organ-Failure Assessment) score.

In parallel with the triage protocol proposed by Christian et al. (2006), we consider a discrete-time model in which patients who use the ICU are in one of two health stages and each patient's health condition changes over time. Specifically, in our model there are two Markov chains with one representing the evolution of the patients in the ICU and one representing the evolution of the patients in the general ward. Each Markov chain has four states corresponding to *death*, *highly critical*, *critical*, and *survival*. Death and survival states are absorbing states and can more broadly be interpreted as "bad" and "good" outcomes, respectively,

depending on the objective. Consequently, we assume that the system incurs a unit cost every time a patient, regardless of whether s/he is in the ICU or in the general ward, hits the death state. There is no reward or cost associated with any one of the other states. As soon as a patient enters the death state or the survival state, s/he leaves the system vacating the bed s/he has been occupying. In each time period, a patient arrives with some probability and a decision needs to be made as to whether or not to admit the patient and/or early discharge any of the highly critical or critical patients to the general ward. The objective is to minimize the long-run average cost, or equivalently, the long-run average number of deaths.

We first consider an extreme setting, where the ICU has a single bed. The main goal in analyzing this hypothetical setting is to take advantage of its relatively simple mathematical formulation and develop insights into how bed allocation decisions should be made when ICU beds are limited; however, we also use the results of this section later in the chapter to develop a heuristic bed allocation policy, which is a simpler alternative to the optimal policy. Our analysis of the single-bed setting reveals that the decision of which patient to admit to the ICU depends on how much benefit the patients are expected to get from ICU treatment and how long they are expected to stay in the ICU, and furthermore, we find that which one of these two factors is more dominant depends on the overall demand level on the ICU. We then consider the general setting, where the ICU has some arbitrary but finite number of ICU beds. We formulate the decision problem as a Markov decision process (MDP) where the system state is described by a vector providing the number of patients (among those in the ICU together with the patient who has just arrived) in each health stage. We prove that the optimal admission/early discharge policy depends on the mix of patients currently present in the ICU, specifically how many patients there are in each health stage. In other words, whether we want to keep Patient A or Patient B in the ICU is not just about Patient A and B but also about all the other patients. Specifically, we prove that the optimal policy is of threshold-type: If the number of highly critical patients is above a particular value then we early discharge one of the highly critical patients; otherwise we early discharge one of the critical patients. Finally, we carry out a numerical study to investigate the benefits that one would get by using the optimal policy as opposed to simpler, state-independent alternatives, and we find that a policy we propose in particular performs quite well.

3.2 Literature review

In the medical literature, there has been a long line of research on quantifying the benefits of ICU care and providing empirical and mathematical support for making more sound ICU admission/discharge decisions. Most of this work has concentrated on predicting patient mortality in the ICU, estimating the benefits of ICU care, and more generally developing patient severity scores. We do not attempt to provide a

thorough review of this literature here as it is extensive and is not directly related to this chapter but only highlight a few papers as examples.

Strand and Flaatten (2008) provides a review of some of the severity scoring systems that have been proposed and used over the years. Among these scoring systems are APACHE (Acute Physiology and Chronic Health Evaluation) I, II, III, and IV (Zimmerman et al. (2006)), SAPS (Simplified Acute Physiology Score) I, II, and III (Moreno et al. (2005)), and SOFA (Vincent et al. (1996)). One of the objectives behind the development of these scoring systems is to obtain a tool that can reliably predict patient mortality, which has been the subject of many other articles that aimed to improve upon the predictive power of the proposed scoring systems (see, e.g., Rocker et al. (2004), Gortzis et al. (2008), and Ghassemi et al. (2014)).

A number of papers studied the benefits of ICU care and the effects of rationing beds in times of limited availability. Sinuff et al. (2004) reviewed past studies on bed rationing and found that admission to the ICU is associated with lower mortality. Shmueli and Sprung (2005) studied the potential survival benefit for patients of different types and severity (measured by APACHE II score) and more recently Kim et al. (2014) quantified the cost of ICU admission denial on a number of patient outcomes including mortality, readmission rate, and hospital length of stay using a large data set. Kim et al. (2014) also carried out a simulation study to test various patient admission policies and found that a threshold type policy which takes into account the patient severity and ICU occupancy level has the potential to significantly improve overall performance.

Studies found that delayed admission to or early discharge from ICUs, which are both common, affect patients outcomes. For example, Chalfin et al. (2007) and Cardoso et al. (2011) studied patients immediately admitted to ICU and those who had delayed admissions (i.e., waited longer than 6 hours for admission) and concluded that the patients in the latter group are associated with longer length of stay and higher ICU and hospital mortality. Wagner et al. (2013) and Kc and Terwiesch (2012) found patients were discharged more quickly when ICU occupancy was high, and such patients were associated with increased mortality rate and readmission probability.

In addition to Kim et al. (2014), which we have already mentioned above, a number of papers from the operations literature developed and analyzed models with the goal of generating insights into how patient admission and discharge decisions should be made at ICUs. Modeling the ICU as an M/M/c/c queue, Shmueli et al. (2003) compare three different patient admission policies and find that restricting admission to those whose expected benefit is above a certain threshold (which may or may not depend on the number of occupied beds in the ICU) brings sizeable improvements in the expected number of survivors. Dobson et al. (2010), on the other hand, develop a model in which patients are bumped out of (early discharged from) the ICU and show how this model can be used to predict performance measures like the probability of

being bumped for a randomly chosen patient. The model assumes that each patient's length of stay can be observed upon arrival and when a patient needs to be bumped because of lack of beds, the patient with the shortest remaining length of stay is bumped out of the ICU. Chan et al. (2014) develop a fluid formulation in which service rate can be increased (which can be seen as patient early discharge) at the expense of increased probability of readmission. The authors identify scenarios under which taking such action is and is not helpful.

To our knowledge, within the operations literature on ICUs, the paper that is closest to our work is Chan et al. (2012). The authors consider a discrete-time MDP in which a decision needs to be made as to which patient to early discharge (with a cost) every time a new patient arrives for admission to the ICU. They show that the greedy policy, which discharges the class with the smallest discharge cost, is optimal when patient types can be ordered so that the types with smaller discharge costs have shorter expected length of stay and provide bounds on the performance of this policy for cases when such ordering is not possible. Despite some similarities, our formulation and analysis have some important differences. We assume that patients can be in one of two health stages, can transition from one stage to the other during their stay, and they eventually either die or survive. On the other hand, Chan et al. (2012) allow for multiple types of patients whose health status can also change over time but their model does not permit a patient to return to a state s/he has already visited. The reason why these differences are mainly important is that the analysis of the two models leads to two different sets of results which complement each other. In particular, our formulation allows us to push the analytical results and optimal policy characterizations further and thereby provide deeper insights into optimal ICU admission and discharge decisions. For example, we provide a characterization of the optimal policy not only when patients with higher benefits from ICU have shorter length of stay but also when higher benefits can only come at the expense of longer length of stay in the ICU.

Our analysis in this chapter can also be seen as a contribution to the classical queueing control literature where arriving jobs are admitted or rejected according to some reward or cost criteria. More specifically, because jobs in our model do not queue, it can be seen as a loss system (see, e.g., Örmeci et al. (2001), Örmeci and Burnetas (2005), Ulukus et al. (2011) and references therein). Within this literature, Ulukus et al. (2011) appears to be the closest to our work. This chapter considers a model in which the decision is not only whether or not an arriving job should be admitted but also whether any of the jobs in service should be terminated. This termination action can be seen as the early discharge action in our model. However, despite this similarity, there are some important differences. First, Ulukus et al. (2011) consider a more general form for the termination cost. However, they assume that there is no cost associated with rejecting jobs. While this assumption would be reasonable in many service settings, that is mostly not true in our

context since this would imply that it is better to never admit a patient than to admit and early discharge. Second, Ulukus et al. (2011) do not allow the possibility of jobs changing types during service. There are also important differences in the results. Just as we do in this chapter, Ulukus et al. (2011) also provide conditions under which one of the two types should be preferred over the other at all times. However, our formulation makes it possible for us to provide optimal policy characterizations at a more detailed level and mathematically establish some of the numerical observations made by Ulukus et al. (2011) regarding the threshold structure of the optimal policy.

3.3 Model Description

We consider an ICU with a capacity of b beds, where b is a finite positive integer. Patients arriving to this system are assumed to have conditions that require treatment in an ICU. However, there is also the option of admitting these patients to what we refer to as the *general ward*, where the patient may be provided a different level of service. It is also possible that a patient who was previously admitted to the ICU can be early discharged to the general ward in order to accommodate another patient. We assume that the general ward has infinite bed capacity. It is important to note that we use the general ward to represent any non-ICU care unit, which includes actual hospital wards, other hospital units, nursing homes, and any other facility that can accommodate the patients but cannot provide an ICU-level service to the patients.

Arriving patients are assumed to be in one of two health stages with stage 1 representing a *highly critical* condition and stage 2 representing a *critical* condition. We consider discrete time periods during which at most one patient arrives. Let $\lambda_i > 0$ denote the probability that a stage i patient will arrive in each period for $i = 1, 2$ and let $\bar{\lambda} = 1 - \lambda_1 - \lambda_2$ denote the probability of no arrival, where we assume $\bar{\lambda}$ is nonnegative. During their stay, in the ICU or in the general ward, patients' health conditions change according to a Markov chain and they eventually either enter stage 0 or stage 3. Stage 0 represents an undesired outcome, which might correspond to the death of the patient, or hospital readmission shortly after the patient's discharge (e.g., within thirty days). On the other hand, stage 3 represents an ideal outcome, such as the patient's survival or at least not being readmitted for a period of time long enough to count the patient's treatment as success. As soon as a patient hits either stage 0 or 3, the patient leaves the system vacating the bed s/he has been occupying. We assume that the system incurs a unit cost every time a patient leaves in stage 0 while there is no cost or reward associated with other stages.

Patients currently in stage $i \in \{1, 2\}$ can enter stage $i+1$ or $i-1$ in the next time period with probabilities that depend on where they are being treated: ICU or general ward. A stage i patient in the ICU either jumps to stage $i + 1$ with probability p_i, jumps to stage $i - 1$ with probability q_i, or stays in stage i with probability $r_i = 1 - p_i - q_i$. The respective probabilities for the general ward are p_i^G, q_i^G and r_i^G. We assume

that p_i, q_i, p_i^G, q_i^G are all strictly positive while r_i and r_i^G are non-negative. The transition diagram of patient evolution is shown in Figure 3.1.

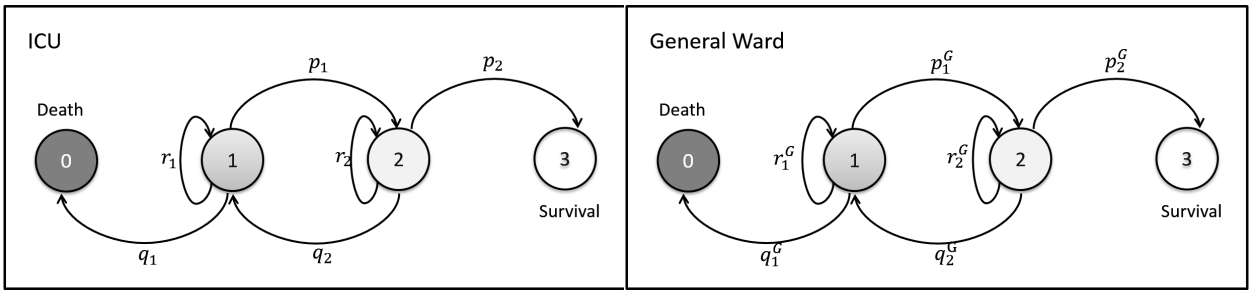

Figure 3.1: Transition diagram of patient evolution in the ICU and general ward

In some respects, assuming that sick patients can only be in one of two health stages can be seen as a significant simplification of reality. While it is true that it is difficult to capture the full spectrum of patient diversity with a two-stage model, the assumption helps us capture the reality that patients' health conditions change over time at least in some stylized way without rendering the analysis impossibly difficult. More importantly, the assumption can in fact be perfectly justified in some contexts because even in practice such simplification is made to bring highly complex decision problems to a manageable level. When managing patient demand under highly resource restrictive environments, particularly in case of epidemics and mass-casualty events, practitioners typically choose to employ prioritization policies that keep the number of triage classes at minimum in an effort to make the policies simpler and easier to implement. For example, the ICU triage protocol developed by Christian et al. (2006), places patients in need of ICU treatment into one of two priority classes based on the patients' SOFA scores. The proposed protocol also calls for patient reassessments recognizing the possibility that there could be changes in the patients' health conditions.

At each time period, the decision maker needs to make the following decisions: (i) if there is an arrival, whether the patient should be admitted to the ICU or the general ward, and (ii) which patients in the ICU (if any) should be early discharged to the general ward regardless of whether there is a new arrival or not. Note that if all b beds are occupied at the time a stage i patient arrives, admitting the patient will mean early discharging at least one stage $3-i$ patient to the general ward. To keep the presentation simple, we will call both the decision of discharging an existing patient from the ICU to the general ward and admitting a new arrival to the general ward *discharge* even though the latter action does not in fact correspond to a discharge but direct admission to the general ward.

We formulate this problem as an MDP. We denote the system state by $\mathbf{x} = (x_1, x_2)$, where x_i represents the number of stage i patients. Note that any new arrival is included either in x_1 or x_2 since there is no need to distinguish between new and existing patients. Since the ICU has a capacity of b and at most 1 patient

arrives in each time period, the state space is:

$$\mathcal{S} = \{(x_1, x_2) : x_1, x_2 \geq 0 \text{ and } x_1 + x_2 \leq b + 1\}.$$

The decision at each epoch can be described by action $a = (a_1, a_2)$, where a_i is the number of stage i patients to be discharged. The action space is defined as $\mathcal{A} = \{(a_1, a_2) : a_1, a_2 \geq 0, \text{ and } a_1 + a_2 \leq b + 1\}$. Then in any state $(x_1, x_2) \in \mathcal{S}$, the feasible action set is

$$\mathcal{A}(x_1, x_2) = \{(a_1, a_2) : 0 \leq a_i \leq x_i, \text{ for } i = 1, 2, \text{ and } x_1 + x_2 - a_1 - a_2 \leq b\}.$$

Let ϕ_i^G denote the probability that a patient who is discharged to the general ward in stage i will end up in stage 0 for $i = 1, 2$. Then, ϕ_i^G can be computed by solving the following equations

$$\phi_1^G = q_1^G + r_1^G \phi_1^G + p_1^G \phi_2^G, \ \phi_2^G = q_2^G \phi_1^G + r_2^G \phi_2^G.$$

Letting $\beta_i^G = q_i^G / p_i^G$ for $i = 1, 2$, we can show that

$$\phi_1^G = \frac{\beta_1^G + \beta_1^G \beta_2^G}{1 + \beta_1^G + \beta_1^G \beta_2^G}, \quad \phi_2^G = \frac{\beta_1^G \beta_2^G}{1 + \beta_1^G + \beta_1^G \beta_2^G}. \tag{3.1}$$

Similarly, for $i = 1, 2$, let ϕ_i denote the probability that a patient who is admitted to the ICU in stage i will end up in stage 0 under the condition that the patient will never be early discharged to the general ward. Then, ϕ_i can similarly be computed as

$$\phi_1 = \frac{\beta_1 + \beta_1 \beta_2}{1 + \beta_1 + \beta_1 \beta_2}, \quad \phi_2 = \frac{\beta_1 \beta_2}{1 + \beta_1 + \beta_1 \beta_2}. \tag{3.2}$$

where $\beta_i = q_i / p_i$ for $i = 1, 2$.

Let $c(x_1, x_2, a_1, a_2)$ denote the immediate expected cost of taking action (a_1, a_2) in state (x_1, x_2). The expected cost for the patients who will occupy the ICU during the next period is equal to the expected number of ICU patients who will transition to state 0 in the next time period, i.e., $(x_1 - a_1)q_1$. The expected cost for the discharged stage i patients is $a_i \phi_i^G$ since each discharged patient will end up in stage 0 with probability ϕ_i^G. Note that this second portion of the cost is the expected lump-sum cost of discharging stage i patients, the expected cost that will eventually incur, not the immediate cost. However, for our analysis, we can equivalently assume that this cost will incur immediately since we know that if the patient enters state 0 eventually, this will happen within some finite time period with probability 1. The total immediate

expected cost then can be written as

$$c(x_1, x_2, a_1, a_2) = a_1\phi_1^G + a_2\phi_2^G + q_1(x_1 - a_1).$$

Let $P_{(a_1,a_2)}(x_1, x_2, y_1, y_2)$ denote the transition probability from state (x_1, x_2) to state (y_1, y_2) when action (a_1, a_2) is chosen. Then, we have $P_{(a_1,a_2)}(x_1, x_2, y_1, y_2) = P(y_1, y_2|x_1-a_1, x_2-a_2)$, where $P(y_1, y_2|x_1, x_2)$ denotes the probability that given that there are x_1 stage 1 patients and x_2 stage 2 patients at a decision epoch after that epoch's action is taken, there will be y_1 stage 1 patients and y_2 stage 2 patients at the beginning of the next decision epoch. Specifically,

$$\begin{aligned}
P(y_1, y_2|x_1, x_2) = \bar{\lambda} &\sum_{u=0}^{x_1}\sum_{d=0}^{x_1-u} \bar{P}_1\{x_1, u, d\}\bar{P}_2\{x_2, x_1 + x_2 - d - y_1 - y_2, y_1 - (x_1 - u - d)\} \\
+ \lambda_1 &\sum_{u=0}^{x_1}\sum_{d=0}^{x_1-u} \bar{P}_1\{x_1, u, d\}\bar{P}_2\{x_2, x_1 + x_2 - d - (y_1 - 1) - y_2, (y_1 - 1) - (x_1 - u - d)\} \\
+ \lambda_2 &\sum_{u=0}^{x_1}\sum_{d=0}^{x_1-u} \bar{P}_1\{x_1, u, d\}\bar{P}_2\{x_2, x_1 + x_2 - d - y_1 - (y_2 - 1), y_1 - (x_1 - u - d)\},
\end{aligned}$$

where $\bar{P}_i\{x_i, u, d\}$ is the probability that of the x_i stage i patients, u of them will transition to stage $i+1$ and d of them will transition to stage $i-1$, i.e.,

$$\bar{P}_i\{x_i, u, d\} = \begin{cases} \binom{x_i}{u}\binom{x_i-u}{d} p_i^u q_i^d r_i^{x_i-u-d}, & \text{for } u, d \geq 0 \text{ and } u + d \leq x_i \\ 0, & \text{otherwise.} \end{cases}$$

A policy π maps the state space $\mathcal{S}$ to the action space $\mathcal{A}$. We use Π to denote the set of feasible stationary discharge policies. Let $N_\pi(t)$ and $N_\pi^G(t)$ respectively denote the number of patients who enter stage 0 by time t in the ICU and in the general ward. Then $J^\pi(\mathbf{x})$, the expected long-run average cost under policy π given the initial state x, can be expressed as

$$J^\pi(\mathbf{x}) = \lim_{t\to\infty} \frac{1}{t} E\left[N_\pi(t) + N_\pi^G(t)|\mathbf{x}\right].$$

Our objective is to obtain an optimal policy π^* such that $J^{\pi^*}(\mathbf{x}) \leq J^\pi(\mathbf{x})$ for any $\pi \in \Pi$ and $\mathbf{x} \in \mathcal{S}$. Note that this MDP is unichain with finite state and action spaces, hence the above limit exists and is independent of the initial state $\mathbf{x}$ (see, e.g., Theorem 8.4.5 of Puterman (2005)). We also know that there exists a bounded

function $h(x_1, x_2)$ for $(x_1, x_2) \in \mathcal{S}$ and a constant g satisfying the optimality equation

$$h(x_1,x_2) \;+\; g \quad = \quad \min_{(a_1,a_2)\in\mathcal{A}(x_1,x_2)} \Big\{ c(x_1,x_2,a_1,a_2) \;+\; \sum_{(y_1,y_2)\in\mathcal{S}} P_{(a_1,a_2)}(x_1,x_2,y_1,y_2) h(y_1,y_2) \Big\}, \quad (3.3)$$

and there exists an optimal stationary policy π^* such that $g = J^{\pi^*}(\mathbf{x})$ and π^* chooses an action that maximizes the right-hand side of (3.3) for each $(x_1, x_2) \in \mathcal{S}$.

3.4 Single-bed ICU

In this section, we consider the case where $b = 1$, i.e., there is a single ICU bed. The objective of this analysis is to generate insights into situations where ICU capacity is severely limited. As we will later see in Section 3.6, this analysis will also help us develop a simple heuristic method that can be used in the general case, where b is any finite integer.

When $b = 1$, at any decision epoch there are at most two patients under consideration, the patient who is currently occupying the bed (if there is one) and the patient who has just arrived for possible admission (if there was an arrival). Restricting ourselves to non-idling policies, (i.e., the bed is never left empty when there is demand), we investigate the question of which of the two patients to admit to the ICU. (An implicit assumption here is that ICU is the preferred environment for the patients. This is a reasonable assumption to make, but nevertheless in the next section, we identify conditions under which this is true in our mathematical formulation.) Specifically, there are two stationary policies to compare, $\bar{\pi}_1$, the policy that discharges the stage 1 patient and $\bar{\pi}_2$, the policy that discharges the stage 2 patient when the choice is between a stage 1 and a stage 2 patient. Under any of the two policies, when there are two patients in the same stage, the choice between the two is arbitrary. Let $J^{\bar{\pi}_k}$ for $k \in \{1, 2\}$ denote the long-run average cost under policy $\bar{\pi}_k$.

The following proposition provides a comparison of the performances of the two policies, which accounts for both the incremental survival benefit and the required ICU length of stay (LOS) when making prioritization decisions. We first let L_i denote the expected ICU LOS for a patient admitted to the ICU in stage i and is never early discharged in either stage 1 or 2. Then, L_i can be obtained by solving the equations $L_1 = 1 + r_1 L_1 + p_1 L_2$ and $L_2 = 1 + q_2 L_1 + r_2 L_2$, which gives us

$$L_1 = \frac{p_1 + p_2 + q_2}{p_1 p_2 + q_1 p_2 + q_1 q_2}, \; L_2 = \frac{p_1 + q_1 + q_2}{p_1 p_2 + q_1 p_2 + q_1 q_2}. \quad (3.4)$$

Proposition 3.1. *If $b = 1$, i.e., there is a single ICU bed, and the ICU admission decision is between a stage 1 and stage 2 patient, it is optimal to admit the stage 2 patient, i.e., $J^{\bar{\pi}_1} \leq J^{\bar{\pi}_2}$, if and only if*

$$\lambda \left[L_2(\phi_1^G - \phi_1) - L_1(\phi_2^G - \phi_2)\right] + (1-\lambda)\left[(\phi_1^G - \phi_1) - (\phi_2^G - \phi_2)\right] \leq 0 \tag{3.5}$$

where $\lambda = \lambda_1 + \lambda_2$.

While Proposition 3.1 provides a complete mathematical description of the regions under which the two policies are optimal, it does not give a clear insight into why one would prefer one patient over the other. The following corollary provides conditions that have clearer practical interpretations.

Corollary 3.1. *Suppose that $b = 1$, i.e., there is a single ICU bed, and the ICU admission decision is between a stage 1 and stage 2 patient. Also assume without loss of generality that $\phi_i^G - \phi_i \geq \phi_{3-i}^G - \phi_{3-i}$ for some fixed $i \in \{1, 2\}$. Then, we have*

(a) if $\frac{\phi_i^G - \phi_i}{L_i} \geq \frac{\phi_{3-i}^G - \phi_{3-i}}{L_{3-i}}$, then it is optimal to admit the patient in stage i, i.e., $J^{\bar{\pi}_i} \geq J^{\bar{\pi}_{3-i}}$;

(b) if $\frac{\phi_i^G - \phi_i}{L_i} < \frac{\phi_{3-i}^G - \phi_{3-i}}{L_{3-i}}$, then it is optimal to admit the patient in stage i, i.e., $J^{\bar{\pi}_i} \geq J^{\bar{\pi}_{3-i}}$, if and only if

$$\lambda \leq \frac{(\phi_i^G - \phi_i) - (\phi_{3-i}^G - \phi_{3-i})}{(\phi_i^G - \phi_i) - (\phi_{3-i}^G - \phi_{3-i}) + \left[L_i(\phi_{3-i}^G - \phi_{3-i}) - L_{3-i}(\phi_i^G - \phi_i)\right]}. \tag{3.6}$$

The difference $\phi_i^G - \phi_i$ can be seen as the benefit of staying in the ICU instead of the general ward for a stage i patient. From system optimization point of view, we can call the patients with larger $\phi_i^G - \phi_i$ as "high-value" patients. On the other hand, the ratio $\frac{\phi_i^G - \phi_i}{L_i}$ can roughly be seen as the per unit time benefit of keeping a patient who arrives in stage i in the ICU at all times and thus we can call the patients with the larger $\frac{\phi_i^G - \phi_i}{L_i}$ as "high-value-rate" patients. Then, according to Corollary 3.1 (a), if stage i patients are both high-value and high-value-rate patients, they should be preferred over stage $3-i$ patients. As Corollary 3.1 (b) implies, in order for stage i patients to be preferable, it is not sufficient for them to be high-value. If they are high-value patients but not high-value-rate, then they are preferable only if the arrival rate is sufficiently small. This is because when the arrival rate is small, having a limited bed capacity is less of a concern and thus in that case the value is the dominating factor. However, when the arrival rate is large, the lengths of stay are important as they would be a key factor in the availability of the ICU beds for new patients. As a result the rate with which the value incurs becomes the dominant factor.

These results point to the importance of taking into account the ICU load when making patient admission/early discharge decisions and prioritizing one patient over the other. In short, what may be the "right" thing to do for one particular ICU may not be right for another. For ICUs with relatively ample capacity, it might be best to focus on identifying patients who will benefit most from ICU care and admit them without

being overly concerned about how long they will stay. However, for highly loaded ICUs, the decision is more complicated and the anticipated length of stay should be part of the decision. In the following section, we investigate this question further by analyzing dynamic decisions in a model where the number of beds in the ICU can take any finite value.

3.5 Analysis of the multi-bed ICU model

In this section, we analyze the general case where b, the number of beds in the ICU is any finite integer. As in the previous section, our main objective is to minimize the long-run average cost. However, following an approach that is often used in long-run average analysis, we first analyze the system under the objective of minimizing expected total discounted cost and establish some analytical properties, which serve as a stepping stone to our main results for the long-run average case.

3.5.1 Formulation of the discounted model

sec:multi model discounted

The system states, actions and the transition probability $P_{(a_1,a_2)}(x_1, x_2, y_1, y_2)$ are defined to be the same as the long-run average formulation given in Section 3.3. Let $c_\alpha(x_1, x_2, a_1, a_2)$ denote the immediate expected cost of taking action (a_1, a_2) in state (x_1, x_2) with discount factor $\alpha \in (0, 1)$. The expected discounted cost for the patients who remain in the ICU is $\alpha(x_1 - a_1)q_1$, where $(x_1 - a_1)q_1$ is the expected number of ICU patients who will transition to state 0 in the next time period and these patients will incur a unit cost when they depart in stage 0 at the beginning of the next period. The expected cost for the discharged stage i patients is $a_i c_i^G$, where c_i^G is the expected discounted cost of discharging a stage i patient for $i = 1, 2$. The expression for c_i^G can be determined by solving the following equations:

$$c_1^G = \alpha(q_1^G + r_1^G c_1^G + p_1^G c_2^G),\ c_2^G = \alpha(q_2^G c_1^G + r_2^G c_2^G),$$

which are obtained by first-step analysis. Solving the above equations, we find

$$c_1^G = \frac{\alpha q_1^G(1 - \alpha r_2^G)}{(1 - \alpha r_1^G)(1 - \alpha r_2^G) - \alpha^2 p_1^G q_2^G},\ c_2^G = \frac{\alpha^2 q_1^G q_2^G}{(1 - \alpha r_1^G)(1 - \alpha r_2^G) - \alpha^2 p_1^G q_2^G}. \tag{3.7}$$

The total immediate expected cost then can be written as

$$c_\alpha(x_1, x_2, a_1, a_2) = a_1 c_1^G + a_2 c_2^G + \alpha q_1(x_1 - a_1).$$

Let $v_{\pi,\alpha}(\mathbf{x})$ denote the expected total discounted cost under policy π given initial state $\mathbf{x}_0$, then,

$$v_{\pi,\alpha}(\mathbf{x}_0) = E\left[\sum_{n=0}^{\infty} c_\alpha(\mathbf{x}_n, \mathbf{a}_n)\alpha^n | \mathbf{x}_0\right],$$

where $\mathbf{x}_n \in \mathcal{S},\ \mathbf{a}_n \in \mathcal{A}$ are bivariate vectors that denote the state and action at the nth decision epoch for $n = 0, 1, 2, \ldots$. The above quantity is well defined since $\mathcal{S}$ and $\mathcal{A}$ are finite and $c_\alpha(\mathbf{x}, \mathbf{a})$ is bounded for any $\mathbf{x} \in \mathcal{S},\ \mathbf{a} \in \mathcal{A}$.

Let $v_\alpha(x_1, x_2) = \min_\pi v_{\pi,\alpha}(x_1, x_2)$ denote the minimum expected total discounted cost over infinite horizon starting from state $\mathbf{x}_0 = (x_1, x_2)$. We would like to find a policy π^* that satisfies $v_{\pi^*,\alpha}(x_1, x_2) = v_\alpha(x_1, x_2)$ for all $(x_1, x_2) \in \mathcal{S}$.

From Theorem 6.2.5 in Puterman (2005), the optimal value function $v_\alpha(x_1, x_2)$ satisfies the following optimality equation:

$$v_\alpha(x_1, x_2) = \min_{(a_1,a_2)\in\mathcal{A}(x_1,x_2)} \{V_\alpha(x_1, x_2, a_1, a_2)\}, \tag{3.8}$$

where $V_\alpha(x_1, x_2, a_1, a_2)$ is the total expected discounted cost if we take action (a_1, a_2) for one step and then follow the optimal policy thereafter in state (x_1, x_2). Specifically,

$$V_\alpha(x_1, x_2, a_1, a_2) = a_1 c_1^G + a_2 c_2^G + \alpha\left[q_1(x_1 - a_1) + \Gamma v_\alpha(x_1 - a_1, x_2 - a_2)\right],$$

where Γ is an operator defined as follows:

Definition 3.1. For a function $w(x_1, x_2)$ with $x_1, x_2 \geq 0$ and $x_1 + x_2 \leq b$

$$\Gamma w(x_1, x_2) = \sum_{j_1=0}^{x_1+x_2+1} \sum_{j_2=0}^{x_1+x_2+1-j_1} P(j_1, j_2|x_1, x_2) w(j_1, j_2). \tag{3.9}$$

3.5.2 Main results for the discounted model

For the infinite-horizon expected total α-discounted cost problem (3.8), we denote the set of optimal actions in state (x_1, x_2) by $\mathcal{A}^*_\alpha(x_1, x_2)$, i.e.,

$$\mathcal{A}^*_\alpha(x_1, x_2) = \Big\{(\bar{a}_1, \bar{a}_2) \in \mathcal{A}(x_1, x_2) : \\ V_\alpha(x_1, x_2, \bar{a}_1, \bar{a}_2) \leq V_\alpha(x_1, x_2, a_1, a_2) \text{ for all } (a_1, a_2) \in A(x_1, x_2)\Big\}.$$

Since the state space and action space are finite and costs are bounded, $\mathcal{A}^*$ is non-empty. In general, the set $A^*_\alpha(x_1, x_2)$ can have more than one element. However, for convenience, we adopt the following convention for picking one action from the set and refer to it as *the* optimal action for state (x_1, x_2). Specifically, we define the optimal action $a^*_\alpha(x_1, x_2) = (a^*_{1\alpha}(x_1, x_2), a^*_{2\alpha}(x_1, x_2))$, where

$$a^*_{1\alpha}(x_1, x_2) = \min\{\bar{a}_1 : (\bar{a}_1, \bar{a}_2) \in A^*_\alpha(x_1, x_2)\}$$

and

$$a^*_{2\alpha}(x_1, x_2) = \min\{\bar{a}_2 : (a^*_{1\alpha}(x_1, x_2), \bar{a}_2) \in A^*_\alpha(x_1, x_2)\}.$$

Thus, if there are multiple actions for any given state, we choose the one that discharges as few stage 1 patients as possible; if there are multiple such actions then among those, we choose the one that discharges as few stage 2 patients as possible.

3.5.2.1 Optimality of non-idling ICU beds.

The non-idling policies are defined as the policies that will always allocate an ICU bed to a new arriving patient and never discharge an ICU patient to the general ward when there are ICU beds available. We first identify conditions under which there exists an optimal policy, which is non-idling.

Proposition 3.2. *There exists a non-idling optimal policy, i.e., there exists an optimal policy under which it is never optimal to leave an ICU bed empty whenever there is a patient in need of treatment, if*

$$\alpha(q_1 + r_1 c_1^G + p_1 c_2^G) \le c_1^G, \ \alpha(q_2 c_1^G + r_2 c_2^G) \le c_2^G. \tag{3.10}$$

The term $\alpha(q_1 + r_1 c_1^G + p_1 c_2^G)$ is the expected total discounted cost of discharging a stage 1 patient at the next time period, and thus $\alpha(q_1 + r_1 c_1^G + p_1 c_2^G) \le c_1^G$ means that the expected total discounted cost of keeping a stage 1 patient for one more time period is smaller than discharging the patient right now. Similarly, $\alpha(q_2 c_1^G + r_2 c_2^G) \le c_2^G$ means that the expected total discounted cost of keeping a stage 2 patient for one more time period is smaller than discharging the patient right now.

From Proposition 3.2, we know that if (3.10) holds, there exists an optimal policy that is non-idling. Thus, we can restrict ourselves to the set of policies that is non-idling under assumption (3.10). Then, the optimality equations (3.8) can be reduced to

$$v_\alpha(x_1, x_2) = T v_\alpha(x_1, x_2) \text{ for } (x_1, x_2) \in S, \tag{3.11}$$

where the optimality operator T is defined as follows:

Definition 3.2. Let $\mathcal{W}$ denote the space of bounded functions on $\mathcal{S}$. Then, for $w \in \mathcal{W}$, we define the operator T as

(i) for $x_1 + x_2 \leq b$,

$$Tw(x_1, x_2) \quad = \quad \alpha \left[q_1 x_1 + \Gamma w(x_1, x_2) \right], \tag{3.12}$$

(ii) for $x_1 = b + 1, x_2 = 0$,

$$Tw(x_1, x_2) \quad = \quad c_1^G + \alpha \left[q_1 (x_1 - 1) + \Gamma w(x_1 - 1, x_2) \right], \tag{3.13}$$

(iii) for $x_1 = 0, x_2 = b + 1$,

$$Tw(x_1, x_2) \quad = \quad c_2^G + \alpha \left[q_1 x_1 + \Gamma w(x_1, x_2 - 1) \right], \tag{3.14}$$

(iv) for $x_1 + x_2 = b + 1$ and $x_1, x_2 > 0$,

$$Tw(x_1, x_2) = \min \Big\{ c_1^G + \alpha \left[q_1 (x_1 - 1) + \Gamma w(x_1 - 1, x_2) \right], \\ c_2^G + \alpha \left[q_1 x_1 + \Gamma w(x_1, x_2 - 1) \right] \Big\}. \tag{3.15}$$

3.5.2.2 General structure of the optimal policy.

Since we restrict ourselves to the set of non-idling policies, which we know contains an optimal policy under assumption (3.10), we only need to investigate the optimal actions for states (x_1, x_2) such that $x_1 + x_2 = b + 1$ and $x_1, x_2 > 0$, i.e., when all ICU beds are currently occupied, a patient has just arrived, and there are patients from both stages (including the patient who has just arrived). As we describe in the following proposition, it turns out that the optimal decision has a threshold structure.

Proposition 3.3. *Suppose that (3.10) holds. Then, there exists a threshold $x_\alpha^* \in [1, b+1]$ such that for any state (x_1, x_2) with $x_1, x_2 > 0$ and $x_1 + x_2 = b + 1$, we have*

$$a_\alpha^*(x_1, x_2) = \begin{cases} (1, 0) & \text{if } x_1 \geq x_\alpha^*, \\ (0, 1) & \text{if } x_1 < x_\alpha^*. \end{cases}$$

Proposition 3.3 states that the decision of whether a stage 1 or stage 2 patient should be early discharged may depend on how many stage 1 patients and how many stage 2 patients there are in the ICU or waiting

for admission to the ICU. There exists a threshold x^*_α, which depends on the discount factor, so that if the number of stage 1 patients is below x^*_α then the optimal action is to discharge a stage 2 patient; otherwise, the optimal action is to discharge a stage 1 patient. We postpone further discussion on practical implications of this result until Section 3.5.3, where we establish a version of this result under the long-run average cost minimization criteria.

If the threshold $x^*_\alpha = 1$, the optimal decision is to discharge a stage 1 patient and if $x^*_\alpha = b + 1$, the optimal decision is to discharge a stage 2 patient regardless of how many stage 1 and stage 2 patients there are in the ICU. Thus, in some cases, where x^*_α takes one of the two end values, the optimal policy is simpler as one can designate one stage as having higher priority than the other regardless of system conditions. Such a policy would be easier to implement in practice. In the following section, we identify some conditions under which that would be the case.

3.5.2.3 Conditions for the optimality of a state independent policy.

We call a non-idling policy state-independent if the decision of whether a stage 1 patient or a stage 2 patient is kept in the ICU whenever all beds are occupied does not depend on the composition of the patients, i.e., the number of patients in each stage. In this section, we identify conditions under which the optimal policy has that simple structure.

Let c_i denote the expected total discounted cost for a stage i patient if the patient is not discharged when in state 1 or 2. Then, c_i can be obtained as

$$c_1 = \frac{\alpha q_1(1-\alpha r_2)}{(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2},\ c_2 = \frac{\alpha^2 q_1 q_2}{(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2}, \tag{3.16}$$

using the same argument used to obtain c^G_i's in (3.7).

The difference $c^G_i - c_i$ can be seen as the reduction in cost due to ICU treatment for stage i patients. Then, if stage i patients have smaller reduction in cost (compared with stage $3-i$ patients), one might conjecture that a *greedy* policy, which always chooses to discharge a stage i patient (as opposed to a stage $3-i$ patient) whenever there are more patients than beds, would be optimal. However, we have numerical examples which show that this conjecture is not correct. We next identify conditions under which such a policy is optimal.

Proposition 3.4. *Suppose that (3.10) holds. Then, for any state (x_1, x_2) such that $x_1 > 0$, $x_2 > 0$, and $x_1 + x_2 = b + 1$,*

(a) $a^*_\alpha(x_1, x_2) = (1, 0)$ if

$$c^G_1 - c_1 < c^G_2 - c_2 \text{ and } L_1 \geq L_2, \tag{3.17}$$

(b) $a^*_\alpha(x_1, x_2) = (0, 1)$ *if*

$$c_1^G - c_1 \geq c_2^G - c_2 \text{ and } L_1 \leq L_2, \tag{3.18}$$

where L_i *is given by (3.4) for* $i = 1, 2$.

Proposition 3.4 states that the optimality of the greedy policy is guaranteed, i.e., the optimal decision is to favor patients in the health stage that is associated with higher expected ICU benefit regardless of system conditions, if the expected length of ICU stay for those patients is also smaller when compared with that of patients in the other stage. In the next section, we establish a version of this result for the long-run average cost minimization case and provide further discussion on its practical implications.

3.5.3 Long-run average cost criteria

In this section, we consider the long-run average cost optimization problem with optimality equations given in (3.3). An optimal action in any particular state is the one that achieves the minimum in the optimality equation. We denote the set of optimal actions in state (x_1, x_2) by $\mathcal{A}^*(x_1, x_2)$, i.e.,

$$\mathcal{A}^*(x_1, x_2) = \Big\{(\bar{a}_1, \bar{a}_2) \in \mathcal{A}(x_1, x_2) : c(x_1, x_2, \bar{a}_1, \bar{a}_2) + \sum_{(y_1,y_2)\in\mathcal{S}} P_{(\bar{a}_1,\bar{a}_2)}(x_1, x_2, y_1, y_2)h(y_1, y_2) = \min_{(a_1,a_2)\in\mathcal{A}(x_1,x_2)} \Big\{c(x_1, x_2, a_1, a_2) + \sum_{(y_1,y_2)\in\mathcal{S}} P_{(a_1,a_2)}(x_1, x_2, y_1, y_2)h(y_1, y_2)\Big\}\Big\}.$$

In general, the set $\mathcal{A}^*_\alpha(x_1, x_2)$ can have more than one element. However, for convenience, as in the case of the discounted cost model, we adopt the following convention for picking one action from the set and refer to it as *the* optimal action for state (x_1, x_2). Specifically, we define the optimal action $a^*(x_1, x_2) = (a^*_1(x_1, x_2), a^*_2(x_1, x_2))$ where

$$a^*_1(x_1, x_2) = \min\{\bar{a}_1 : (\bar{a}_1, \bar{a}_2) \in A^*(x_1, x_2)\}$$

and

$$a^*_2(x_1, x_2) = \min\{\bar{a}_2 : (a^*_1(x_1, x_2), \bar{a}_2) \in A^*(x_1, x_2)\}.$$

Thus, if there are multiple actions for any given state, we choose the one that discharges as few stage 1 patients as possible; if there are multiple such actions then among those, we choose the one that discharges as few stage 2 patients as possible.

Theorems 3.1, 3.2, and 3.3 below respectively establish the long-run average versions of Propositions 3.2, 3.3, and 3.4. The proofs of these theorems utilize Theorem 6.4.2 of Sennott (1999), which essentially state that under some conditions (which we show to be true for our problem), the difference $h_\alpha(x_1, x_2) :=$

$v_\alpha(x_1, x_2) - v_\alpha(0, 0)$ converges to $h(x_1, x_2)$ as α approaches to 1. This limiting property makes it possible for us to utilize the three propositions for the discounted cost case in the analysis of the long-run average cost problem.

We start with conditions that ensure the existence of a non-idling optimal policy.

Theorem 3.1. *Suppose that $\beta_i < \beta_i^G$ for $i = 1, 2$. Then, there exists a stationary average-cost optimal policy, which is non-idling, i.e., a policy under which it is never optimal to leave an ICU bed empty whenever there is a patient in need of treatment.*

Comparing β_i with β_i^G can be seen as one way of assessing the potential benefit of ICU over the general ward for stage i patients. The condition $\beta_i < \beta_i^G$ for $i = 1, 2$ essentially means that the ratio of the probability of a patient getting worse to the probability of a patient getting better over the next time step is smaller in the ICU for all the patients. Theorem 3.1 states that this condition is sufficient to ensure the existence of an optimal policy that admits patients of either stage to the ICU as long as there is an available bed. Note that one can show that this condition is equivalent to (3.10), the non-idling condition for the discounted infinite-horizon case when the discount factor α is set to 1.

Under the non-idling condition of Theorem 3.1, we can in fact prove the existence of an optimal policy, which has some additional structural properties as described in the following two theorems.

Theorem 3.2. *Suppose that $\beta_i < \beta_i^G$ for $i = 1, 2$. Then, there exists a threshold $x^* \in [1, b+1]$ such that for any state (x_1, x_2) with $x_1, x_2 > 0$ and $x_1 + x_2 = b + 1$, we have*

$$a^*(x_1, x_2) = \begin{cases} (1, 0) & \text{if } x_1 \geq x^* \\ (0, 1) & \text{if } x_1 < x^*. \end{cases}$$

According to Theorem 3.2, when the non-idling condition holds and when the system conditions are so that one of the patients has to be admitted to the general ward because of a fully occupied ICU, whether or not that patient should be a stage 1 or stage 2 patient depends on the health conditions of all the patients in the ICU. Specifically, if the number of stage 1 (stage 2) patients in the ICU is above a particular threshold value, then one of the stage 1 (stage 2) patients should be admitted to the general ward. In other words, if there are sufficiently many stage 1 patients, the preference should be for a stage 2 patient; otherwise the preference should be for a stage 1 patient.

It is important to note that while x^* can take one of the boundary values of 1 or $b + 1$ (both of which would imply that the policy is in fact not dependent on the composition of the patients) there are examples that show that it can also take values in between. This means that there are indeed certain settings in which the optimal policy is state-dependent. This might seem somewhat surprising at first because the implication

is that if there are two specific patients, A and B, one of them being in stage 1 the other in stage 2, and only one of them can be admitted to the ICU, whether we choose A or B depends on the health stages of all the patients in the ICU, not just A and B. Given that this decision will not impact other patients' survival chances and patient A's and B's survival chances do not depend on the other patients in the ICU, why should our choice between A and B depend on the other patients?

To answer the question above, in light of our analysis of the single-bed case, consider the two important factors that go into the decision of which patient to admit: expected net ICU benefit, which we would like to be as high as possible and expected length of stay, which we would like to be as small as possible. The expected length of stay is important because it directly affects the bed availability for the future patients. In particular, it affects the probability that a bed will be available the next time there is a patient seeking admission to the ICU. However, whether or not a bed will be available for the next patient (and patients thereafter) depends on the length of stay for not just Patient A and Patient B but all the patients in the ICU.

Now, consider two extreme cases, one in which patients other than A and B all have very short expected lengths of stay and one in which they all have long expected lengths of stay. In the former case, there is a good chance for a bed to be available soon even if we ignore A and B, and this, when choosing between A and B, will make the expected lengths of stay for A and B far less important compared with the latter case. Thus, in the former case, whoever has the larger expected benefit, will be (most likely) admitted to the ICU. In the latter case, however, the decision is more complicated and in order to make a bed available for the next patient with a higher probability, it might actually be preferable to admit the patient with the smaller expected net benefit if that patient's expected length of stay is shorter. In general, one can then see that, as the composition of the patients in the ICU changes, future bed availability probability changes and this in turn results in shifting preferences for the patient to be admitted. More specifically, as Theorem 3.2 implies, there is an ideal mix of patients (a certain number of stage 1 patients and a certain number of stage 2 patients), which hits the "right" balance between the expected benefit and the future bed availability, and the optimal policy continuously strives to push the system to that level by employing a threshold-type policy.

Given the explanation above, it would be reasonable to expect that Patient A should always be preferred over Patient B regardless of the patient composition in the ICU if the expected benefit for Patient A is larger than that of Patient B and the expected length of stay for Patient A is smaller than that of Patient B. We can indeed prove that is the case as we formally state in the following theorem.

Theorem 3.3. *Suppose that $\beta_i < \beta_i^G$ for $i = 1, 2$, and for some fixed $k \in \{1, 2\}$*

$$\phi_k^G - \phi_k < \phi_{3-k}^G - \phi_{3-k} \text{ and } L_k \geq L_{3-k}. \tag{3.19}$$

Then, for any state (x_1, x_2) *such that* $x_1 + x_2 = b + 1$*, we have* $a^*(x_1, x_2) = (a_1^*, a_2^*)$ *with* $a_k^* = 1$ *and* $a_{3-k}^* = 0$.

Theorem 3.3 states that if a particular health stage is associated with a lower expected ICU benefit and longer expected length of ICU stay, then a patient from that health stage should be admitted to the general ward when the demand for the ICU exceeds the ICU bed capacity. In this case, the optimal policy is simple since one of the two stages can be designated as the higher priority stage regardless of the system state. The result makes sense intuitively. If Patient A will benefit more from the ICU bed compared to Patient B and Patient A will also vacate the bed more quickly for the use of the future patients, there is no reason why the bed should be given to Patient B.

3.6 Numerical study

The objective of our numerical study is twofold: (i) to investigate how much additional benefit there is in using a state-dependent admission/discharge policy as opposed to a simpler, state-independent alternative, and (ii) to investigate how different state-independent policies, one of which we propose in this section, compare with each other in regards to their performances.

To the best of our knowledge, there are no data which we can use to estimate the probabilities with which patients' health stages change over time. Clearly, even if one were to conduct a study to make such estimation, it would have been quite challenging since that would have first required agreeing on definitions for stage 1 and stage 2 patients. (Note that one possibility would be adopting the definitions of Christian et al. (2006).) Given the lack of data, one option for our study was to generate different scenarios completely randomly. However, that approach would have resulted in many scenarios that are unlikely to be reasonably close to what happens in practice. Therefore, instead of generating scenarios completely randomly, we generated them randomly while ensuring that the scenarios conform to what happens in practice based on what we know from prior studies. More specifically, the scenarios are generated so that the ICU survival probabilities and lengths of stay are in line with the numbers reported in the literature.

A number of articles in the literature provide estimates on ICU length of stay and survival probabilities. One could potentially choose the estimates in any one of these articles to construct different scenarios. However, in line with our focus on situations where the ICU experiences an extremely high demand over a long period of time, we chose to use the estimates that are provided by Kumar et al. (2009), which are based on data obtained in Canada during the 2009 H1N1 influenza outbreak. Kumar et al. (2009) found that the average mortality rate in the ICU was approximately 17% and the average length of stay in the ICU was 12 days. (In our numerical study, we assumed that a single day is divided into four time periods and thus the expected length of stay of 12 days corresponds to 48 time periods.) Therefore, we randomly generated

scenarios so that the ICU death probability is $\theta\phi_1 + (1-\theta)\phi_2 = 0.17$ and the expected ICU length of stay (with no early discharge) is $\theta L_1 + (1-\theta)L_2 = 48$, where θ is the probability that a random patient arrived to the ICU is in stage 1, ϕ_1 and ϕ_2 are as defined in (3.2), and L_1 and L_2 are as defined in (3.4). To randomly generate feasible scenarios that satisfy the above two equalities, we used the following procedure:

First, using Equations (3.2) and (3.4), we find that the transition probabilities need to satisfy

$$p_1 = \frac{1-\phi_1}{(1-\phi_2)L_1 - (1-\phi_1)L_2},\ q_1 = \frac{\phi_1 - \phi_2}{(1-\phi_2)L_1 - (1-\phi_1)L_2},$$
$$p_2 = \frac{\phi_1 - \phi_2}{\phi_1 L_2 - \phi_2 L_1},\ q_2 = \frac{\phi_2}{\phi_1 L_2 - \phi_2 L_1}.$$

Using the fact that $p_i, q_i > 0$ and $p_i + q_i \leq 1$ for $i = 1, 2$, one can show that the generated scenarios must satisfy

$$\frac{1-\phi_1}{1-\phi_2} \leq \frac{L_1 - 1}{L_2} < \frac{L_1}{L_2} < \frac{L_1}{L_2 - 1} \leq \frac{\phi_1}{\phi_2}. \tag{3.20}$$

Now, to construct a single scenario, we first generated ϕ_1 uniformly in $(0.2, 0.5)$ and chose the fraction ϕ_1/ϕ_2 uniformly in $(1, 10)$. Then, using the equation $\theta\phi_1 + (1-\theta)\phi_2 = 0.17$, we solved for θ, the proportion of stage 1 patients. Since we know from (3.20) that L_1/L_2 must be between $\frac{1-\phi_1}{1-\phi_2}$ and $\frac{\phi_1}{\phi_2}$, we chose L_1/L_2 uniformly in $(\frac{1-\phi_1}{1-\phi_2}, \frac{\phi_1}{\phi_2})$. Then, using the equation $\theta L_1 + (1-\theta)L_2 = 48$ and θ, whose value has already been determined, we determined L_1 and L_2. We checked whether the generated scenario satisfy all the feasibility conditions of (3.20). We discarded the scenario and restarted the random generation if the conditions did not hold; otherwise, we kept the scenario and proceeded with populating the rest of the parameters. Kumar et al. (2009) do not provide any estimates on what the survival probabilities for the ICU patients would be if they would have been treated outside the ICU. In the absence of such estimates, recognizing that the death probabilities outside the ICU would likely be higher than what they would be in the ICU and stage 1 patients would be more likely to die than stage 2 patients, we chose ϕ_1^G uniformly in the interval $(\phi_1, 1)$ and ϕ_2^G uniformly in the interval (ϕ_2, ϕ_1^G). Using this procedure we generated 1000 different scenarios.

For every one of the 1000 scenarios constructed as above, we considered, twelve different set of conditions regarding the number of beds in the ICU and ICU demand. This resulted in a total of 12000 scenarios tested. For the size of the ICU, we considered three different cases. We assumed it to be small ($b = 5$), medium ($b = 10$) or large ($b = 20$). We also set the arrival probability λ so that the system is busy at different levels. Specifically, we let the ICU utilization $\rho = \lambda(\theta L_1 + (1-\theta)L_2)/b$ to be either 0.8, 1, 1.2, or 2.4. (Note that once ρ is set, we can determine λ, $\lambda_1 = \theta\lambda$, and $\lambda_2 = (1-\theta)\lambda$ since $\theta,\ L_1,\ L_2$, and b are already set for each scenario.)

Next, we propose a new state-independent policy as an alternative to the state-dependent optimal policy and describe the other policies we tested in our numerical study.

3.6.1 A new policy and other benchmarks

A state-independent policy is one that always prioritizes one of the stages over the other regardless of the system state. Such policies have clear practical advantages and thus can still be preferred in practice even when they can in fact be potentially outperformed by state-dependent policies. Two immediate alternatives for such policies is the policy that always chooses a stage 1 patient when there is one and the policy that always chooses a stage 2 patient when there is one. Another obvious alternative is what we call the *greedy policy* (GP), which always chooses the patient in the health stage that is associated with the higher expected benefit independently of the system state. Specifically, if $\phi_1^G - \phi_1 \geq \phi_2^G - \phi_2$, the greedy policy chooses stage 1 patients over stage 2 patients; otherwise, the greedy policy chooses stage 2 patients. In our numerical study, we find that the greedy policy clearly outperforms both of the other two simple policies mentioned above and thus we do not discuss those policies in the rest of the chapter.

The greedy policy intuitively makes sense as it always chooses to allocate the beds to the patients who are likely to benefit more. However, it ignores the long-term effects of the decisions, more specifically, how long the beds will need to be occupied in expectation in order for the benefits to be realized. In Section 3.4, we provided a complete characterization of the optimal policy for a single-bed ICU. This characterization provided us with necessary and sufficient conditions for the optimality of the greedy policy and more generally a clear prescription for which stage patient to prefer at all times depending on various parameters including the ICU benefit, lengths-of-stay, and system load or more specifically ICU demand under the restriction that the ICU has a single bed. The main insight that came out of this analysis was the following: the greedy policy is optimal if the patients with higher expected benefits have shorter expected length of stay; however, if the patients with higher expected benefit have longer length of stay, which is likely to be the case in practice, the greedy policy is optimal only if the ICU demand (arrival probability) is sufficiently low.

In the case of a multi-bed ICU, it may not be realistic to expect that the exact same conditions will continue to have the exact same implications. It is however reasonable to expect that the factors like ICU benefit, length of stay, and system load will continue to have similar effects on the optimal actions. Thus, one can adapt Condition (3.6) of Corollary 3.1 to the multi-bed setting by simply rescaling the arrival probability, which represents the ICU demand, on the left-hand side of the condition accordingly and use the new condition to decide which one of the two stages to designate as high priority independently of the system state. This is the idea behind the heuristic method we propose, which we name *load-based state-independent policy* (LBSIP).

Specifically, LBSIP works as follows: Suppose that stage i patients have higher expected ICU benefit without loss of generality, i.e., $\phi_i^G - \phi_i \geq \phi_{3-i}^G - \phi_{3-i}$ for fixed $i \in \{1, 2\}$. Then, if $L_i \leq L_{3-i}$ LBSIP chooses stage i patients over stage $(3-i)$ patients. Otherwise (i.e., if $L_i \geq L_{3-i}$), LBSIP chooses stage i patients over stage $(3-i)$ patients if and only if

$$\frac{\lambda}{b} \leq \frac{(\phi_i^G - \phi_i) - (\phi_{3-i}^G - \phi_{3-i})}{(\phi_i^G - \phi_i) - (\phi_{3-i}^G - \phi_{3-i}) + \left[L_i(\phi_{3-i}^G - \phi_{3-i}) - L_{3-i}(\phi_i^G - \phi_i)\right]}. \tag{3.21}$$

It is important to note that if $b = 1$, LBSIP is the optimal policy since (3.21) reduces to (3.6). We also know from Theorem 3.3 that LBSIP is optimal even when $b > 1$ if $L_i \leq L_{3-i}$. We do not know, however, how LBSIP performs when $b > 1$ and $L_i > L_{3-i}$, i.e., conditions that are more likely to be true in practice. This is one of the questions we investigate in Section 3.6.2.

One potentially very useful property of Condition (3.21) (and consequently LBSIP) is that the condition depends on the transition probabilities only through the expected net benefits and expected lengths of stay. This means that in practice once there is an agreement on who is a stage 1 patient and who is a stage 2 patient, LBSIP only requires estimation of the expected net benefits and the expected lengths-of-stay for the two stages (in addition to the patient arrival rate), and in particular not the transition probabilities, which are much more difficult to estimate reliably. Furthermore, this conditional independence from transition probabilities suggests that LBSIP can be applied even if the patient health conditions evolve differently from what we assumed in this chapter. In other words, LBSIP provides a way out of the difficulty of identifying the "correct" transition formulation in practice since it only makes use of the very basic estimates for the patients.

3.6.2 Results of the numerical study

Let M^* denote the mortality rate, i.e., long-run fraction of patients who die, under the optimal policy. Also let M_{GP} and M_{LBSIP} denote the mortality rates under GP and LBSIP, respectively. These long-run average mortality rates are obtained numerically by a value iteration algorithm. Table 3.1 below summarizes the results of the numerical study. In the table, the first two columns indicate the number of beds b and the system load ρ. As described above, for each fixed value of b and ρ, 1000 different scenarios were generated. The third column provides the mean mortality rate, the 95% confidence interval (C.I.) half-length, and the maximum value of the mortality rate out of the 1000 scenarios under the optimal policy. The fourth and the fifth columns provide the mean (and the 95% C.I. half-length) and maximum increase in the mortality rate under GP and LBSIP, respectively out of the 1000 scenarios. Finally, the last column provides the same information for the comparison of GP and LBSIP with each other.

Table 3.1: Performance comparison of the optimal policy, GP, and LBSIP. (All numbers are in percentages. In the columns titled "Mean" the first number reported is the mean, the second number is the 95% C.I. half-width.)

b	ρ	M^*		$M_{GP} - M^*$		$M_{LBSIP} - M^*$		$M_{GP} - M_{LBSIP}$	
		Mean	Max	Mean	Max	Mean	Max	Mean	Max
5	0.8	22.21 (0.207)	31.32	0.30 (0.043)	4.63	0.01 (0.004)	0.95	0.29 (0.043)	4.63
	1	24.57 (0.300)	37.81	0.51 (0.070)	7.39	0.02 (0.006)	1.13	0.49 (0.070)	7.39
	1.2	26.67 (0.382)	43.60	0.72 (0.096)	9.98	0.02 (0.008)	1.38	0.70 (0.096)	9.98
	2.4	34.28 (0.677)	64.39	1.54 (0.187)	19.10	0.03 (0.012)	0.19	1.50 (0.226)	19.10
10	0.8	19.89 (0.118)	25.10	0.21 (0.030)	3.38	0.01 (0.003)	1.00	0.20 (0.030)	3.38
	1	22.33 (0.217)	31.95	0.48 (0.065)	7.09	0.03 (0.008)	1.50	0.45 (0.065)	7.09
	1.2	24.69 (0.312)	38.60	0.78 (0.100)	10.94	0.04 (0.012)	2.15	0.74 (0.101)	10.94
	2.4	33.39 (0.654)	62.59	1.83 (0.217)	21.80	0.06 (0.020)	3.01	0.22 (0.255)	21.80
20	0.8	18.25 (0.053)	20.55	0.12 (0.017)	1.98	0.01 (0.002)	0.71	0.11 (0.017)	1.98
	1	20.55 (0.149)	27.15	0.43 (0.058)	6.48	0.04 (0.009)	1.71	0.40 (0.058)	6.48
	1.2	23.13 (0.257)	34.57	0.85 (0.108)	12.58	0.07 (0.017)	2.92	0.79 (0.109)	12.58
	2.4	32.74 (0.639)	61.39	2.16 (0.252)	24.87	0.10 (0.028)	4.09	2.06 (0.254)	24.87

We can observe from Table 3.1 that the optimal policy performs better than both GP and LBSIP as it should. We also found that in all the twelve different scenarios the policies are tested, the mean mortality rate under the optimal policy is statistically smaller than the mean mortality rate under both GP and LBSIP. A quick look at the numbers in the table might nevertheless suggest that the benefit that one would get from using the optimal policy is somewhat small. However, while the numbers might be small, considering what they represent, a small difference might have a highly tangible benefit in practice. For example, consider the case where $b = 10$ and $\rho = 1$. The mean improvement that one would get by using the optimal policy as opposed to GP is 0.48%. This would mean that on average the optimal policy would approximately save one more patient out of every 208 patients in need of an ICU treatment. Furthermore, given that the maximum improvement is 7.09%, the difference in the expected number of survivors out of 100 patients can be as large as 7 patients.

When it comes to comparing the optimal policy and LBSIP, the policy we devised, the performance difference is much less significant. The largest difference is observed when the ICU capacity is large ($b = 20$) and the load is high ($\rho = 2.4$). In this case, the difference in the mortality rate is 0.10%, which corresponds to about saving on average one more patient out of every 1000 patients. If we look at the maximum improvement (4.09%), we see that out of every 100 patients, the optimal policy can potentially save almost four more patients. Four out of 100 patients is a sizable difference but it is worth recalling that this is the maximum improvement observed from the 1000 randomly chosen scenarios (i.e., transition probability values). As it is

also somewhat clear from the mean improvement, in very few scenarios, the improvement gets close to this level.

One can argue that no matter how small the difference is one should always use the optimal policy. The optimal policy always has the best performance after all, it can easily be determined, and is relatively simple (even though not as simple as GP or LBSIP). Given all that, why should one bother with policies that we know are suboptimal? The problem is that even though we can find the optimal policy for our mathematical optimization problem we do not know how exactly this optimal policy would perform in practice. One important challenge is posed by transition probabilities, which are needed in order to determine the optimal policy.

Our model, particularly the part that captures the evolution of patient criticality, is largely a stylized representation of what happens in practice. Even though in ICUs patient health conditions are frequently assessed and these assessments are used to make accept/discharge decisions there is not a commonly accepted protocol for classifying patients. Thus one way of interpreting our model is by seeing it as a rough conceptualization of what happens and how decisions are made in practice even if patients are not strictly put into two levels in reality. Another interpretation would be based on the triage protocol proposed by Christian et al. (2006) where there are two criticality levels and patients can transition from one level to the other just like we assume in our model. Under this interpretation, the model can be seen as somewhat less stylized since the patient criticality levels can actually be described precisely. With the first interpretation, it is not clear what exactly transition probabilities correspond to in practice. With the second interpretation, the meaning of transition probabilities is clear but nevertheless their reliable estimation is very difficult partially due to the coarseness of using only two levels to group patients. Therefore, it is highly doubtful how much of, if any, the potential benefits that we attribute to the "optimal" policy in our numerical study would actually be realized in practice. To be more precise, the benefits reported above would be realized only under the assumption that model parameters including transition probabilities are correctly estimated, which is not quite likely to happen in practice. Therefore, one might still prefer using a policy like LBSIP or even GP, which are not only simpler and easier to implement but also easier to determine because they only require estimation of the expected ICU benefit and length of stay, not the transition probabilities.

3.7 Conclusions

Many studies reported that the number of ICU beds in many parts of the US and the rest of the world are in short supply to sufficiently meet the daily ICU demand. It is frequently the case that a patient who is relatively in a less critical condition is discharged early to make room for another patient who is deemed more critical. While this bed shortage problem arises even under daily operating conditions it is natural to

expect the problem to get worse in case of an event like an influenza epidemic, which causes a significantly increased number of patients in need of an ICU bed. It is thus highly important to investigate how ICU capacity can be managed efficiently by allocating the available beds to the patients in a way the greatest good is achieved for the greatest number of the patients. Our goal in this chapter has been to provide insights into how such allocation decisions should be made.

What sets our work apart from prior work mainly is that in our model we allow the patients to move from one health stage to another while they are in the ICU and allocation decisions are made based on the patients' updated health conditions. This formulation captures an important feature of the actual problem at least in some stylized way and nicely fits with the triage protocol proposed by Christian et al. (2006). But more importantly, the model allowed us to push the analytical results further than it was possible for other models considered in prior work and in particular we were able to provide analytical results for the case where patients who have higher expected ICU benefits also have longer expected length of stay.

Our analysis of the single-bed scenario led to interesting insights into how optimal decisions depend on the patients' expected ICU benefit, expected length of stay, and the patient load on the system. We found that when patients who are expected to benefit more from ICU treatment also have longer expected length of stay, those patients should get higher priority only if the overall patient demand is below a certain level. This is because when beds are in high demand, prioritizing those patients (who are expected to occupy the beds longer) would require turning too many patients away from the ICU that it becomes more preferable to adopt a policy that has quicker bed turnaround times even though the expected net benefit is smaller for every admitted patient.

More generally, when the ICU has finitely many beds, we found that the optimal policy aims for an ideal mix in the ICU so as to hit the right balance between the overall expected net ICU benefit per patient and length of stay. That is, in general, the optimal policy for prioritizing among patients depends on the mix of customers in the ICU. Even though our formulation is stylized in nature, our finding that the optimal policy in general depends on the system state suggests that there could be benefits to developing sophisticated decision support tools that take into account patient characteristics and health conditions in the ICU. Thus, our results provide some support for research that aims to contribute to the development of such tools. However, this is also a highly challenging research avenue and one might wonder whether the potential benefits of such tools, which might be observed in numerical studies, can actually be realized in practice. This is because considering the variety of patients treated in the ICU, there could be significant obstacles to classifying patients well enough for the potential benefits to be realized. The main goal of our numerical study was to shed some light into this question by investigating whether the optimal policy brings significant benefits at least in our stylized setup. Our study revealed that the improvement with the optimal policy

may be statistically significant but somewhat small especially when its performance is compared with that of the state-independent policy we propose in this chapter. Furthermore, as we discussed in Section 3.6.2, it is easier to reliably estimate the parameters needed to implement our state-independent policy than the parameters needed to compute the optimal policy. In short, based on our analysis, it is difficult to make a strong case in support of state-dependent policies. Nevertheless, all this analysis is based on a single mathematical formulation, and thus it is prudent not to overgeneralize. It is possible that capturing patient health conditions at a level that is more detailed than considered in our model may lead to higher benefits of using state-dependent optimal policies. This points to an important research avenue for the future.

CHAPTER 4: PRIORITIZATION IN A MULTI-SERVER QUEUEING SYSTEM WITH IMPATIENT CUSTOMERS

In this chapter, we consider a multi-server queueing system with impatient customers. Customers are assumed to be in one of the two different stages, and their stages could change over time. A reward is earned upon each service completion depending on the stage of the customer. Our objective is to maximize the expected total discounted reward and the long-run average reward by prioritization.

4.1 Introduction

In many service systems, customers may become impatient after waiting for a long time and leave without being served. For example, in a call center, customers who wait longer than their tolerance may hang up before being answered. The same thing may also happen in health care systems, where patients waiting for medical resources (operating rooms, intensive care units) may no longer need the resource after a certain time. In the meanwhile, the customers in the system may be described in different status and their status could change. In the call center example, customer may change from a patient mode to an agitated one while waiting, and in the health care example, the health condition of the patients may become better or more severe while receiving and waiting for the treatment.

There are mainly two types of objectives in optimal scheduling of impatient customers: either minimize a cost related measures or maximize a certain reward. The models with the former objective, which we refer to as *cost models*, commonly consider a holding cost per customer per time unit for queueing customers and/or a penalty associated with each reneging customer, while the models with the later objective, which we refer to as *reward models*, commonly consider a reward associated with each service completion. Although there is no general proof of the equivalence of these two types of models, the optimality equations (after formulating as an MDP) under these two objectives have similar structures and could be analyzed in a similar way. One could refer to Down et al. (2011), who considered both cost and reward models, for the similarity in optimality equations and proofs of results. In our model, we will consider a reward model that customers who complete the service will gain a positive service reward upon departure, and the customers who abandon the queue will leave with no reward.

There is a growing literature on optimal scheduling in queueing systems with impatient customers, most of which assumed the heterogeneity exists among independent classes of customer. To be more specific, the customers are differentiated by different classes based on their service requirements, impatience, reward and/or cost structures, etc. Once a customer is classified, s/he will always be of the same class in the system. Accounting for the abandonment of the queueing customers increases the difficulty of analyzing optimal policy. Most of the existing literature provide some partial characterization of the optimal policy and propose heuristic and/or index-based priority policies, or analyzing the system under some heavy traffic/fluid approximation. In our model, we will first formulate the system as an MDP, and then provide optimal policies under certain conditions. Our results could reduced to many existing results in literature with special parameter settings, and we will demonstrate the similarity and difference by considering three special parameter settings. In the end, we propose several priority index policies based on our analytical results and compare their performances via a numerical study.

4.2 Literature Review

There are two types of literature that are related to our work. The first type is optimal scheduling of impatient customers. Argon et al. (2008) consider priority assignment in a clearing system with impatient jobs, where the first work considers a single-server formulation with two different types of jobs and the objective is to maximize the expected number of survivors. Jacobson et al. (2012) consider a more general formulation with type dependent service rewards. Atar et al. (2010) consider a cost model of a single server queue with abandonment under heavy traffic. They proposed a $c\mu/\theta$ index policy, by which priority is assigned to the class with highest value of $c_i\mu_i/\theta_i$, where c_i is the per unit waiting cost in queue, and μ_i and θ_i are the service rates and abandonment rates for class i customers, respectively. They show that such a policy is asymptotically optimal in the heavy traffic analysis. Down et al. (2011) study dynamic control of an M/M/1 queue with impatient customers that belongs to two different classes. They formulate two continuous time MDPs for each of the cost model and the reward model, and establish some sufficient conditions under which a state-independent prioritization policy is optimal for each model. However, the complete structure of the optimal policy is not discussed.

The second type of related work are models with customers who change status (denoted by either classes or stages) in the system. For example, Down and Lewis (2010) consider a multi-server system where low-priority customers will be upgraded to high-priority class if they have been in queue for some time. Cao and Xie (2015) consider a single server queue with customers that could transfer from one class to the other with a cost of transferring. He et al. (2012) study a priority queueing system with multiple classes of customers,

where customers can upgrade from their current class to the next more important class after a so-called upgrading time.

There is very limited literature combining both abandonment and customer status changing. Akan et al. (2012) study liver allocation to patients that belong to multiple classes (i.e., health levels) and the patients could switch between classes over time. They model the transplant waiting list as a multi-class fluid model of overloaded queues and analyze the bi-criteria objective of minimizing number of patient deaths and maximizing total quality-adjusted life years. Chapter 5.2 of Jacobson (2010) considers a single server queue with impatient customers, where the jobs waiting in the queue will go over multiple stages sequentially before abandonment, and provides a set of sufficient conditions under which it is optimal to prioritize jobs at the last stage.

4.3 Model description and the MDP formulation

In this section, we first provide a mathematical description of the system, and then formulate it as an MDP.

4.3.1 Model description

We consider a discrete-time setting, where there will be at most one customer arriving to the system at each time period. We describe the status of customers by their stages, and they could be in one of the two stages (labeled by stage 1 and stage 2). We next define the arrival process and the service and queueing dynamics of the customers.

Arrival process: At the beginning of each time period, there will be a customer arrival to the system in stage i, $i \in \{1,2\}$, with nonnegative probability λ_i, and there will be no arrival with probability $\lambda_0 = 1-\lambda_1-\lambda_1$. We assume $0 \leq \lambda_i \leq 1$ for all $i \in \{0,1,2\}$. Note that this system will reduce to a clearing model if $\lambda_0 = 1$.

Service and queueing processes: We assume that customers in the system change their stages according to a Markov chain, where they can leave the system, or stay at the same stage or change to the other stage with certain probabilities. The transition probabilities depend on the stage of the customer as well as whether s/he is in service or in queue. More specifically, for a stage i customer in service, s/he will leave the system (due to service completion) with probability p_{i0}, stay in stage i with probability $p_{i,i}$ and transit into stage $3-i$ with probability $p_{i,3-i}$. For a stage i customer waiting in queue, they will change stages similarly with parameters q_{ij} for $i,j \in \{1,2\}$ and abandon the queue with probability q_{i0}. All these transition probability parameters are assumed to be non-negative. The transitions of customers in service and queue are shown in Figure 4.1.

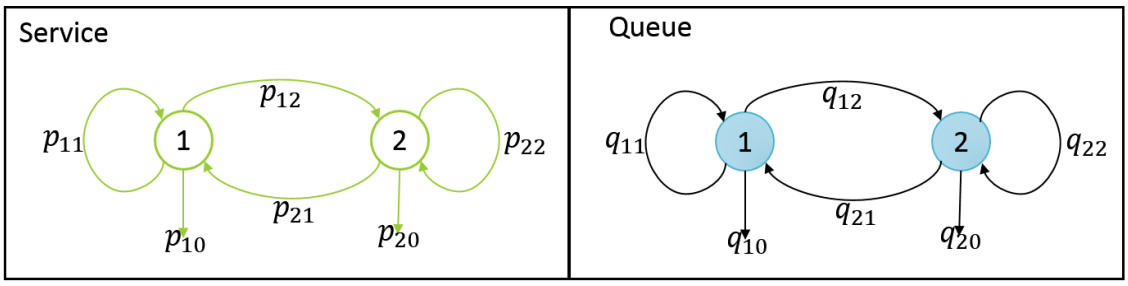

Figure 4.1: Customer Transition Diagrams

Furthermore, we need to make some assumptions so that the system is stable.

Assumption 4.1.

(i) $p_{11} < 1,\ p_{22} < 1$ and $p_{10} + p_{20} > 0$.

(ii) $q_{11} < 1,\ q_{22} < 1$ and $q_{10} + q_{20} > 0$.

The first part of Assumption 4.1 ensures the expected required length of service for any customer is finite, and the second part ensures that the expected waiting tolerance for any customer is finite.

Service rewards: There is a positive reward R_i collected at each service completion, depending on the stage the customer leaves the system in. Our objective is to maximize the expected total discounted reward and the long-run average reward over infinite-horizon.

At the beginning of each time period, we first observe the arrival process: whether or not there would be an arrival, and if yes, what stage the new arrival is in. Then, we can observe the current number of customers in each stage in the system, and we decide whom to keep in service and whom to put in the queue among all the customers in the system. At the end of this time period, all customers make transitions according to the corresponding Markov chain (service or queue).

4.3.2 Formulation as an MDP

We next formulate this system as a discrete-time MDP. For notational convenience, we use bold symbols to denote two-dimensional vectors in $\mathbb{N}^2$, where $\mathbb{N}$ denotes the set of natural numbers. For $\mathbf{x} \in \mathbb{N}^2$, let x_i denote the ith component of the vector $\mathbf{x}$ for $i = 1, 2$. We also define a partial ordering of the vectors as follows: for two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{N}^2$, $\mathbf{x}$ is said to be smaller than $\mathbf{y}$, denoted by $\mathbf{x} \leq \mathbf{y}$, if $x_1 \leq y_1$ and $x_2 \leq y_2$. Finally, we let $\mathbf{e_0} = (0, 0)$, $\mathbf{e_1} = (1, 0)$ and $\mathbf{e_2} = (0, 1)$.

System state: Let $\mathbf{x}_t = (x_{t,1}, x_{t,2})$ denote the system state at discrete time points $t = 0, 1, 2, \ldots$, where $x_{t,i}$ is the number of type i customers at time t for $i \in \{1, 2\}$, which includes all existing customers as well as the new arrival if there is one. Here we do not differentiate a new arriving customer and the existing

customers that belong to the same type due to the Markov property of customer transitions. The state space is $S = \mathbb{N}^2$.

Actions: The decisions are made at the beginning of each time period immediately after the customer arrival (if there is one), at which times we need to decide how to assign the servers to all presenting customers. Let a_i denote the number of servers we allocate to stage i customers, and $\mathbf{a} = (a_1, a_2)$ denote a possible action if $\mathbf{a} \in A$ where

$$A = \{\mathbf{a} : \mathbf{a} \in \mathbb{N}^2 \text{ and } a_1 + a_2 \leq b\}.$$

For any state $\mathbf{x}$, let $A(\mathbf{x})$ denote the set of all feasible actions where $A(\mathbf{x}) = \{\mathbf{a} : \mathbf{a} \in A \text{ and } \mathbf{a} \leq \mathbf{x}\}$. Let $A'(\mathbf{x})$ denote the set of all *feasible non-idling actions*, where $A'(\mathbf{x}) = \big\{\mathbf{a} : \mathbf{a} \in A(\mathbf{x}) \text{ and } a_1 + a_2 = \min\{x_1 + x_2, b\}\big\}$.

One step expected reward: When the process is in state $\mathbf{x} \in S$ and an action $\mathbf{a} \in A(\mathbf{x})$ is chosen, there will be an expected immediate reward $R(\mathbf{x}, \mathbf{a}) = \sum_{i=1}^{2} a_i R_i p_{i0}$, which is bounded since a_i, R_i are nonnegative and bounded.

Transition probabilities: Let $P_{\mathbf{a}}(\mathbf{x}, \mathbf{y})$ denote the probability that the process will transit to state $\mathbf{y} \in S$, starting from state $\mathbf{x} \in S$ given action $\mathbf{a} \in A(\mathbf{x})$. The transition probabilities can be computed by conditioning on how each customer evolves. We can also compute these probabilities recursively as follows: first we can obtain the transition probabilities in state $\mathbf{x} = \mathbf{e_0}$ with the only feasible action $\mathbf{a} = \mathbf{e_0}$:

$$P_{\mathbf{e_0}}(\mathbf{e_0}, \mathbf{e_0}) = \lambda_0, \;\; P_{\mathbf{e_0}}(\mathbf{e_0}, \mathbf{e_1}) = \lambda_1, \;\; P_{\mathbf{e_0}}(\mathbf{e_0}, \mathbf{e_2}) = \lambda_2, \;\; \text{and} \;\; P_{\mathbf{e_0}}(\mathbf{e_0}, \mathbf{y}) = 0 \text{ for } \mathbf{y} \notin \{\mathbf{e_0}, \mathbf{e_1}, \mathbf{e_2}\}.$$

Then, for $\mathbf{x} \in S$ and $\mathbf{x} \neq \mathbf{e_0}$, we have $\mathbf{x} \geq \mathbf{e_i}$ for at least one $i \in \{1, 2\}$. Using the fact that all patients evolve independently, $P_{\mathbf{a}}(\mathbf{x}, \mathbf{y})$ satisfies the following properties for $\mathbf{x} \geq \mathbf{e_i}$ and $\mathbf{a} \in A(\mathbf{x})$:

$$P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) = \sum_{j=0}^{2} p_{ij} I_{\{\mathbf{y} \geq \mathbf{e_j}\}} P_{\mathbf{a}-\mathbf{e_i}}(\mathbf{x} - \mathbf{e_i}, \mathbf{y} - \mathbf{e_j}), \quad \text{if } \mathbf{a} \geq \mathbf{e_i}, \tag{4.1}$$

$$P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) = \sum_{j=0}^{2} q_{ij} I_{\{\mathbf{y} \geq \mathbf{e_j}\}} P_{\mathbf{a}}(\mathbf{x} - \mathbf{e_i}, \mathbf{y} - \mathbf{e_j}), \quad \text{if } \mathbf{a} \leq \mathbf{x} - \mathbf{e_i}. \tag{4.2}$$

The intuition behind (4.1) is that, if in state $\mathbf{x}$ we take an action $\mathbf{a}$ that keeps at least one stage i customer in service, and we pick any one such customer (referred to as customer A), and then we compute the transition probability to state $\mathbf{y}$ by conditioning on how customer A evolves. The customer A jumps to stage j with probability p_{ij}, and the probability that the system will transition to state $\mathbf{y}$ equals to the probability that the remaining customers (we have $\mathbf{x} - \mathbf{e_i}$ remaining customers, among which we take action $\mathbf{a} - \mathbf{e_i}$) transit to state $\mathbf{y} - \mathbf{e_j}$ when $\mathbf{y} \geq \mathbf{e_j}$, which equals $P_{\mathbf{a}-\mathbf{e_i}}(\mathbf{x} - \mathbf{e_i}, \mathbf{y} - \mathbf{e_j})$. Similarly, if there is at least one type i

customer in the queue when we take action $\mathbf{a}$ in state $\mathbf{x}$, then we can obtain (4.2) by picking a stage i customer in queue and conditioning on how this customer evolves.

We will only need the these properties (4.1) and (4.2) later in the proofs of our analytical results, and we can obtain the values of the transition probabilities recursively with these two properties and the initial transition probabilities at $\mathbf{x} = \mathbf{e_0}$ in the numerical study.

Objectives: we consider two models with different optimality criteria, which we refer to as a *discounted model* and an *average model*, respectively. For the *discounted model*, we maximize the expected total discounted reward over an infinite-horizon, which can be expressed as

$$V_{\pi,\alpha}(\mathbf{x}) = E_\pi \left[\sum_{t=0}^{\infty} R(\mathbf{x}_t, \mathbf{a}_t)\alpha^t \Big| \mathbf{x}_0 = \mathbf{x} \right],$$

where $\alpha \in (0,1)$ is the discount factor. Note $V_{\pi,\alpha}(\mathbf{x})$ is well defined since $R(\mathbf{x},\mathbf{a})$ is bounded and $\alpha < 1$. Let

$$V_\alpha(\mathbf{x}) = \max_\pi V_{\pi,\alpha}(\mathbf{x}),$$

where the maximum is attainable since the action space is finite and $V_{\pi,\alpha}(\mathbf{x})$ is bounded. A policy π^* is said to be α-optimal if $V_{\pi^*,\alpha}(\mathbf{x}) = V_\alpha(\mathbf{x})$ for all $\mathbf{x} \in S$.

For the *average model*, we would like to maximize the long-run average reward, which can be expressed as

$$g_\pi(\mathbf{x}) = \liminf_{T\to\infty} \frac{E_\pi \left[\sum_{t=0}^{T} R(\mathbf{x}_t, \mathbf{a}_t) \Big| \mathbf{x}_0 = \mathbf{x} \right]}{T+1}.$$

A policy π^* is said to be average optimal if $g_{\pi^*}(\mathbf{x}) = \max_\pi g_\pi(\mathbf{x})$ for all $\mathbf{x} \in S$.

The following two Lemmas will provide the optimality equations for these two models.

Lemma 4.1. *(The optimality equation for the discounted model.)*

(a) $V_\alpha(\mathbf{x})$ *satisfies the following optimality equation:*

$$V_\alpha(\mathbf{x}) = \max_{\mathbf{a}\in A(\mathbf{x})} \left\{ R(\mathbf{x},\mathbf{a}) + \alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}}(\mathbf{x},\mathbf{y}) V_\alpha(\mathbf{y}) \right\}, \ \mathbf{x} \in S. \tag{4.3}$$

(b) The stationary policy that selects any action maximizing the right-hand side of (4.3) *in state* $\mathbf{x}$ *is* α*-optimal.*

Lemma 4.1 follows directly from Theorems 2.1 and 2.2 in Chapeter II of Ross (1983).

Lemma 4.2. *(The optimality equation for the average model.)*

(a) There exists a bounded function $h(\mathbf{x})$ and a constant g, where

$$h(\mathbf{x}) = \lim_{\alpha \to 1} [V_\alpha(\mathbf{x}) - V_\alpha(\mathbf{x}_0)] \text{ and } g = \lim_{\alpha \to 1} (1-\alpha) V_\alpha(\mathbf{x}_0)$$

for some $\mathbf{x}_0 \in S$, which satisfy the following optimality equation:

$$g + h(\mathbf{x}) = \max_{\mathbf{a} \in A(\mathbf{x})} \left\{ R(\mathbf{x}, \mathbf{a}) + \sum_{\mathbf{y} \in S} P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) h(\mathbf{y}) \right\}, \ \mathbf{x} \in S. \tag{4.4}$$

(b) There exists a stationary policy π^* that is average optimal and $g_{\pi^*}(\mathbf{x}) = g$ for all $\mathbf{x} \in S$, and π^* is any policy that selects an action maximizing the right-hand side of (4.4) in state $\mathbf{x}$.

Proof. For any policy π and any $\alpha < 1$, $V_{\pi,\alpha}(\mathbf{x})$ is bounded. Then, $V_\alpha(\mathbf{x})$ is bounded for any α by definition. Thus, for some $\mathbf{x}_0 \in S$,

$$|V_\alpha(\mathbf{x}) - V_\alpha(\mathbf{x}_0)| \leq |V_\alpha(\mathbf{x})| + |V_\alpha(\mathbf{x}_0)|$$

is bounded for all α and $\mathbf{x}$. Then, part (a) follows from Theorem 2.2 in Chapter V of Ross (1983) and part (b) follows from Theorem 2.1 of the same chapter. □

Before proceeding further, it is convenient to define the first-order difference operator in the vector form as follows:

Definition 4.1. For a real-valued function $w(\mathbf{x})$ defined on S, the first-order difference operator D_j is defined as

$$D_j w(\mathbf{x}) = w(\mathbf{x} + \mathbf{e_j}) - w(\mathbf{x}), \text{ for } j = 1, 2.$$

4.4 Main results for the discounted model

In this section, we start with the analysis of the optimal control of the α-discounted model. For a real-valued function $w(\mathbf{x})$ defined on the state space S, define the mapping $T_{\mathbf{a}}$ as follows.

$$T_{\mathbf{a}} w(\mathbf{x}) = \begin{cases} R(\mathbf{x}, \mathbf{a}) + \alpha \sum_{\mathbf{y} \in S} P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) w(\mathbf{y}), & \text{if } \mathbf{a} \in A(\mathbf{x}), \\ -\infty, & \text{otherwise.} \end{cases} \tag{4.5}$$

Then, the optimality equation (4.3) can be rewritten as

$$V_\alpha(\mathbf{x}) = \max_{\mathbf{a} \in A(\mathbf{x})} \Big\{ T_{\mathbf{a}} V_\alpha(\mathbf{x}) \Big\},$$

and $\mathbf{a}^*$ is an optimal action in state $\mathbf{x}$ if $T_{\mathbf{a}^*}V_\alpha(\mathbf{x}) \geq T_{\mathbf{a}}V_\alpha(\mathbf{x})$ for all $\mathbf{a}$.

We first propose a finite-horizon MDP model with the objective of maximizing the total expected discounted reward over n time periods. Then, we prove some structural properties for the optimal value function of the discounted model, V_α, by letting n go to infinity.

Let $V_n(\mathbf{x}) = \max_{\mathbf{a}\in A(\mathbf{x})} \left\{T_{\mathbf{a}}V_{n-1}(\mathbf{x})\right\}$ for $n \geq 1$, and $V_0(\mathbf{x})$ be a bounded nonnegative function. Then, we have the following result which follows directly from Proposition 3.1 of Chapter II in Ross (1983).

Lemma 4.3. *$V_n(\mathbf{x}) \to V_\alpha(\mathbf{x})$ uniformly in $\mathbf{x}$ as $n \to \infty$.*

As a result of Lemma 4.3, we can prove the structural properties of $V_\alpha(\mathbf{x})$ by induction on $V_n(\mathbf{x})$, starting from any bounded $V_0(\mathbf{x})$. For simplicity, we assume $V_0(\mathbf{x}) = 0$ for any $\mathbf{x} \in S$ from now on.

4.4.1 Conditions for optimality of non-idling policies

We start with an example that shows non-idling is not always optimal for this model. We state by the following simple example.

Example 4.1. Suppose $R_1 < R_2$, $p_{10} = p_{20}$, $p_{12} = p_{21} = 0$, $q_{12} = 1$ and $q_{22} = 1$. Then, for both discounted and average models, it is optimal to keep the servers idle when there are only stage 1 customers waiting in the queue. This is because customers leaving in stage 2 will obtain a larger reward with the same amount of service time (since $R_2 > R_1$, $p_{10} = p_{20}$). Hence, it is better to keep any stage 1 customers in the queue so that they will become stage 2 after waiting in the queue for one time period.

On the other hand, it is trivial to show that it is suboptimal to idle servers when there are customers in both stages with the assumption that the service rewards are positive. Intuitively, idling servers for all customers will get zero reward, but serving at least one stage of them will get a positive reward.

Next, we consider a case where the number of servers large enough that all existing customers are served simultaneously. This is the case for a clearing model when the remaining number of customers is less than b, or the number of servers(b) is infinite for systems with positive arrival rates. Then, by considering each single customer separately, there are three stationary policies depending on the stage this customer is in, which are described below:

- policy π_0: we keep serving each customer until the service is completed.
- policy π_i for $i \in \{1, 2\}$: we keep each customer in the queue when s/he is in stage i, and start serving when the customer becomes stage $3 - i$, and repeat until s/he leaves the system.

If policy π_0 outperforms policy π_i for both $i \in \{1, 2\}$, then non-idling is optimal when we have sufficiently many servers. Let $\mathbf{r}^{\pi_i} = (r_1^{\pi_i}, r_2^{\pi_i})'$ denote the reward vector (a column vector), where $r_k^{\pi_i}$ is the reward of

serving a stage k customer under policy π_i for $i \in \{0, 1, 2\}$. The next result compares the reward vectors under each of the three policies.

Lemma 4.4. *For $i \in \{1, 2\}$ and $j = 3 - i$, $\mathbf{r}^{\pi_0} \geq \mathbf{r}^{\pi_i}$ if and only if*

$$p_{i0}R_i\Big[(1-\alpha q_{ii})(1-\alpha p_{jj}) - \alpha^2 q_{ij}p_{ji}\Big] \geq p_{j0}R_j\Big[\alpha q_{ij}(1-\alpha p_{ii}) - \alpha p_{ij}(1-\alpha q_{ii})\Big]. \tag{4.6}$$

If we assume $p_{j0}R_j > 0$, then (4.6) reduces to

$$\frac{p_{i0}R_i}{p_{j0}R_j} \geq \frac{\alpha q_{ij}(1-\alpha p_{ii}) - \alpha p_{ij}(1-\alpha q_{ii})}{(1-\alpha q_{ii})(1-\alpha p_{jj}) - \alpha^2 q_{ij}p_{ji}}.$$

Next, we consider the case when the resource is capacity-constrained, i.e., for a system where there are some customers who need to wait in the queue due to insufficient number of servers.

For $i = 1, 2$, let $\bar{\mathbf{R}} = (\bar{R}_1, \bar{R}_2)'$ be a column vector, where $\bar{R}_i$ is the expected total reward we can obtain from a complete and uninterrupted service of a customer who is admitted to service at stage i. Then, by definition we have $\bar{\mathbf{R}} = \mathbf{r}^{\pi_0}$. We define column vector $\bar{\mathbf{R}}^Q = \left(\bar{R}_1^Q, \bar{R}_2^Q\right)$, where $\bar{R}_i^Q$ denotes the expected reward obtained from keeping a stage i customer waiting for one unit of time in the queue and then completing the service of this customer without interruption. Then, $\bar{R}_i^Q = \alpha \sum_{j=1}^{2} q_{ij}\bar{R}_j$ for $i = 1, 2$, or in matrix form, we have,

$$\bar{\mathbf{R}}^Q = \alpha \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \mathbf{r}^{\pi_0}.$$

Assumption 4.2. $\bar{\mathbf{R}}^Q \leq \bar{\mathbf{R}}$.

Assumption 4.2 implies that the expected reward obtained by delaying the service for one time period is no larger than that from an immediate service of customers in either stage, given that the service is uninterrupted once started.

Assumption 4.3.

(i) Assume either one of the following cases hold for $i = 1$.

  (a) $p_{ij} \geq q_{ij}$ for $j = 1, 2$.

  (b) $p_{ij} < q_{ij}$ for $j = 1, 2$.

  (c) $p_{i1} \geq q_{i1}$, $p_{i2} < q_{i2}$ and $\bar{R}_i - \bar{R}_i^Q \geq \alpha(p_{i1} - q_{i1})\bar{R}_1$.

  (d) $p_{i1} < q_{i1}$, $p_{i2} \geq q_{i2}$ and $\bar{R}_i - \bar{R}_i^Q \geq \alpha(p_{i2} - q_{i2})\bar{R}_2$.

(ii) Assume either one of the above four cases hold for $i = 2$.

Proposition 4.1. *If Assumptions 4.2 and 4.3 hold, then* $T_{\mathbf{a}+\mathbf{e_i}}V_\alpha(\mathbf{x}) \geq T_{\mathbf{a}}V_\alpha(\mathbf{x})$ *given* $\mathbf{a}+\mathbf{e_i} \in A(\mathbf{x})$ *for both* $i \in \{1,2\}$*, and it is suboptimal to idle any servers when there are customers waiting in the queue.*

The conditions provided in Proposition 4.1 are sufficient but probably not necessary for idling to be suboptimal. For example, if we consider the case with sufficiently many servers, then Lemma 4.4 provides a necessary and sufficient condition under which non-idling is optimal. The following Lemma shows that Assumption 4.2 is equivalent to the condition provided in Lemma 4.4.

Lemma 4.5. *Assumption 4.2 is true if and only if* (4.6) *holds for both* $i \in \{1,2\}$*.*

Hence, Assumption 4.2 is a necessary and sufficient condition for $\mathbf{r}^{\pi_0} \geq \mathbf{r}^{\pi_i}$ for both $i = 1, 2$, which indicates non-idling is optimal for the case with infinite many servers under Assumption 4.2. In the remainder of this chapter, we assume that only non-idling policies are considered.

4.4.2 Optimality of State-independent priority policies

In practice, it is preferable to have policies that are relatively easy to use. One simple policy is to always give priority to the same stage of customers no matter what the state is. In this section, we provide conditions under which such a policy is optimal. It is trivial to show that if $R_i > R_{3-i}$ but the transition probabilities for both types are the same, then it is optimal to prioritize stage i customers.

The next result provides some sufficient conditions under which prioritizing stage 1 customers is optimal, which could be extended to the optimality of prioritizing stage 2 customers by symmetry.

Proposition 4.2. *Suppose* $R_1p_{10} \geq R_2p_{20}$*. Then,* $T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x}) \geq T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x})$ *given* $\mathbf{a}+\mathbf{e_1}, \mathbf{a}+\mathbf{e_2} \in A(\mathbf{x})$*, i.e., it is better to prioritize stage 1 customers, if one of the following holds:*

(i) $p_{1j}q_{2k} - q_{1j}p_{2k} \geq 0$*, and* $p_{1j} - q_{1j} + p_{10}q_{2j} - q_{10}p_{2j} \geq 0$ *for all* $j, k \in \{1,2\}$*.*

(ii) $p_{1j} \geq q_{1j}$ *and* $p_{2j} \leq q_{2j}$ *for both* $j \in \{1,2\}$*.*

Note that neither of the two sets of conditions in Proposition 4.2 implies the other, although both of them imply $p_{1j}q_{2k} - q_{1j}p_{2k} \geq 0$ and $p_{1j} - q_{1j} + q_{2j} - p_{2j} \geq 0$ for all $j, k \in \{1,2\}$. In addition, under either of these two sets of conditions, we can see that

$$\left(\bar{R}_1 - \bar{R}_1^G\right) - \left(\bar{R}_2 - \bar{R}_2^G\right) = p_{10}R_1 - p_{20}R_2 + \sum_{j=1}^{2}(p_{1j} - q_{1j} + q_{2j} - p_{2j})\bar{R}_j \geq 0.$$

That is, the expected reduced reward from one period of service delay for stage 1 customers is greater than that for stage 2 customers, and hence from a myopic point of view, we would like to prioritize stage 1 customers.

4.5 Main results for the average model

In this section, we extend the results from Section 4.4 by applying Lemma 4.2. We define an operator $H_{\mathbf{a}}$ on a bivariate function $w(\mathbf{x})$ as follows: for $\mathbf{x} \in S$, $H_{\mathbf{a}}w(\mathbf{x}) = -\infty$ if $\mathbf{a} \notin A(\mathbf{x})$ and

$$H_{\mathbf{a}}w(\mathbf{x}) = R(\mathbf{x}, \mathbf{a}) + \sum_{\mathbf{y} \in S} P_{\mathbf{a}}(\mathbf{x}, \mathbf{y})w(\mathbf{y}), \text{ if } \mathbf{a} \in A(\mathbf{x}).$$

Then, from Lemma 4.2, the optimality equations can be rewritten as

$$g + h(\mathbf{x}) = H_{\mathbf{a}}h(\mathbf{x}), \ \mathbf{x} \in S,$$

and $\mathbf{a}^*$ is an optimal action in state $\mathbf{x}$ if and only if $H_{\mathbf{a}^*}h(\mathbf{x}) = \max_{\mathbf{a} \in A} H_{\mathbf{a}}h(\mathbf{x})$

Proposition 4.3. *If Assumptions 4.2 and 4.3 hold for $\alpha = 1$, then, $H_{\mathbf{a}+\mathbf{e_i}}h(\mathbf{x}) \geq H_{\mathbf{a}}h(\mathbf{x})$ given $\mathbf{a} + \mathbf{e_i} \in A(\mathbf{x})$ for both $i \in \{1, 2\}$, and it is suboptimal to idle any servers when there are customers waiting in the queue.*

Proposition 4.4. *Suppose $R_1 p_{10} \geq R_2 p_{20}$. Then, $H_{\mathbf{a}+\mathbf{e_1}}h(\mathbf{x}) > H_{\mathbf{a}+\mathbf{e_2}}h(\mathbf{x})$ given $\mathbf{a} + \mathbf{e_1}, \mathbf{a} + \mathbf{e_2} \in A(\mathbf{x})$, i.e., it is better to prioritize stage 1 customers, if one of the following conditions hold:*

(i) $p_{1j}q_{2k} - q_{1j}p_{2k} \geq 0$, and $p_{1j} - q_{1j} + p_{10}q_{2j} - q_{10}p_{2j} \geq 0$ for all $j, k \in \{1, 2\}$.

(ii) $p_{1j} \geq q_{1j}$ and $p_{2j} \leq q_{2j}$ for both $j \in \{1, 2\}$.

4.6 Analysis of three special models

In this section, we will apply Propositions 4.1 to 4.4 to some special models and compare these results with existing literature. All results in this section are valid for both discounted and average models.

4.6.1 Special model I: no queueing

We first consider the case that customers in queue abandon with probability one at the end of each time period. Then, the model becomes a loss system, i.e., customers who are not admitted to service upon arrival or early discharged during service will be lost forever. In other words, we have $q_{10} = q_{20} = 1$, and hence $q_{ij} = 0$ for all $i, j \in \{1, 2\}$.

Corollary 4.1. *Idling is suboptimal for Model I.*

Corollary 4.1 follows from Propositions 4.1 and 4.3. We have $\bar{R}_i^Q = \sum_{j=1}^{2} q_{ij}\bar{R}_j = 0$ for both $i = 1, 2$ for all α, and hence Assumptions 4.2 holds. Besides, $q_{ij} = 0 \leq p_{ij}$ for all $i, j \in \{1, 2\}$, and hence Assumption 4.3 holds.

Corollary 4.2. *If $R_1 p_{10} \geq R_2 p_{20}$, $p_{1j} - p_{2j} \geq 0$ for all $j \in \{1, 2\}$, then it is optimal to prioritize stage 1 customers.*

Corollary 4.2 follows directly from part (i) of Propositions 4.2 and 4.4. The intuition behind Corollary 4.2 is that, if stage 1 customers have larger one-step reward in service, and the probability of this type remaining in the service in either stage is greater than stage 2 patients, then we would like to prioritize stage 1 customers.

This model is similar to our previous model in Chapter 3, which also considered a multi-server loss system with transitions between stages, with slightly different reward structure. In Chapter 3, we not only provided sufficient conditions for non-idling and conditions for prioritization of one stage, but also proved the threshold structure of the optimal policy. For Model I, we can also show that there exists a threshold-type optimal policy.

Proposition 4.5. *There exists a threshold $x^* \in [1, b+1]$ such that for any state $\mathbf{x} \geq \mathbf{e_1} + \mathbf{e_2}$ and $x_1 + x_2 = b + 1$, it is optimal to prioritize stage 1 customers if and only if $x_1 \leq x^*$.*

The proof is very similar to the proof of Proposition 3.3 in Chapter 3 since the structure of the optimality equations are similar. One can easily adapt the proof structure of Proposition 3.3 to get a complete proof of Proposition 4.5. Another model similar to Model I is presented in Ulukus et al. (2011), where the transitions in between customer stages $p_{i,3-i} = 0$ for $i = 1, 2$ while a termination cost is considered for existing customers but not for new arrivals.

4.6.2 Special model II: no transitions between stages

In this section, we consider the special case when there are no transitions between stages in service and in queue. In this model, for notational simplicity, we let μ_i and β_i denote the probability of a stage i customer leaving system while in service and in queue, respectively, and we assume $0 < \mu_i < 1$ and $0 < \beta_i < 1$ for $i = 1, 2$. In other words, we have $p_{i0} = \mu_i$, $q_{i0} = \beta_i$, $p_{ii} = 1 - \mu_i$, $q_{ii} = 1 - \beta_i$ and $p_{i,3-i} = q_{i,3-i} = 0$ for $i = 1, 2$.

Corollary 4.3. *Idling is suboptimal for Model II.*

We verify that Assumptions 4.2 and 4.3 hold to apply Propositions 4.1 and 4.3. Assumption 4.2 holds since $\sum_{j=1}^{2} q_{ij} \bar{R}_j = q_{ii} \bar{R}_i \leq \bar{R}_i$, and Assumption 4.3 holds since $q_{i,3-i} = p_{i,3-i} = 0$ for both $i = 1, 2$, and then it must fall into either (a) or (b) depending on the relation p_{ii} and q_{ii}.

Corollary 4.4. *If $R_1 \mu_1 \geq R_2 \mu_2$, $\mu_1 \leq \beta_1$ and $\mu_2 \geq \beta_2$, then it is optimal to prioritize stage 1 customers.*

This result follows directly from part (ii) of Propositions 4.2 and 4.4 by plugging the values of transition probabilities for Model II. This result is consistent with Proposition 5.1.1 of Jacobson (2010), who considered a continuous time single server system with Poisson arrivals of K classes of customers and i.i.d. exponentially distributed service times and patience for customers of the same class. The author applied truncation on the state space and used uniformization to prove the result. In our model, we extend their result to a multi-server system, without truncating the state space.

The next result provides another set of sufficient conditions for service prioritization of stage 1 customers for Model II when abandonment rates for all customers are smaller than their service rates.

Proposition 4.6. *Assume* $\mu_2 \geq \mu_1 \geq \beta_1 \geq \beta_2$ *and* $\beta_1 R_1 \geq (\beta_1 + \mu_2 - \mu_1)R_2$. *Then, for any* $\mathbf{x} \in S$ *and* $\mathbf{a}+\mathbf{e_1}, \mathbf{a}+\mathbf{e_2} \in A(\mathbf{x})$, $T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x}) \geq T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x})$ *and* $H_{\mathbf{a}+\mathbf{e_1}}h(\mathbf{x}) \geq H_{\mathbf{a}+\mathbf{e_2}}h(\mathbf{x})$, *i.e., it is optimal to prioritize type 1 customers.*

Proposition 4.7. *Assume* $\mu_1 R_1 \geq \mu_2 R_2$, $\frac{\beta_1}{\beta_2} \geq \frac{\mu_1}{\mu_2} \geq 1$ *and* $1 \leq \frac{1-\beta_1}{1-\mu_1} \leq \frac{1-\beta_2}{1-\mu_2}$. *Then, for any* $\mathbf{x} \in S$ *and* $\mathbf{a}+\mathbf{e_1}, \mathbf{a}+\mathbf{e_2} \in A(\mathbf{x})$, $T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x}) \geq T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x})$ *and* $H_{\mathbf{a}+\mathbf{e_1}}h(\mathbf{x}) \geq H_{\mathbf{a}+\mathbf{e_2}}h(\mathbf{x})$, *i.e., it is optimal to prioritize type 1 customers.*

When $\mu_1 = \mu_2 = \mu$, the conditions in Propositions 4.6 and 4.7 both reduce to $\mu \geq \beta_1 \geq \beta_2$ and $R_1 \geq R_2$. That is, when the service rates are the same and are greater than the abandonment rates, we prefer the customers with larger reward and abandonment rate. These conditions are slightly different from Theorem 3.3 of Down et al. (2011) for their reward model, which considered a single server queueing system with impatient customers that belong to two independent classes. However, an important difference between their reward model and Model II is that they assume abandonment could also happen for customers in service. To incorporate the difference, we redefine a model with abandonment in service as follows.

Model II(A): Assume type i customers in service will complete service with probability $\tilde{\mu}_i$ with a service reward $\tilde{R}_i$. If the service for a type i customer is not completed at the end of that period, this customer will abandon the system with probability β_i without any reward. Type i customers waiting in the queue abandon with probability β_i.

For Model II(A), the probability that type i customers in service will leave is $\tilde{\mu}_i + (1-\tilde{\mu}_i)\beta_i$, either due to service completion or abandonment. Then, we find that the MDP formulation of Model II(A) is the same as that of Model II with parameters $\mu_i = \tilde{\mu}_i + (1-\tilde{\mu}_i)\beta_i$ and $R_i = \frac{\tilde{\mu}_i}{\mu_i}\tilde{R}_i$. With this equivalence, the following result follows directly from Propositions 4.6 and 4.7.

Corollary 4.5. *For Model II(A), suppose* $\beta_1 \geq \beta_2$. *Assume either of the following two statements is true:*

(a) $\tilde{\mu}_1 + \beta_1(1-\tilde{\mu}_1) \leq \tilde{\mu}_2 + \beta_2(1-\tilde{\mu}_2)$ *and* $\frac{\beta_1\tilde{\mu}_1\tilde{R}_1}{\tilde{\mu}_1+(1-\tilde{\mu}_1)\beta_1} \geq \frac{(\beta_2+\tilde{\mu}_2(1-\beta_2)-\tilde{\mu}_1(1-\beta_1))\tilde{\mu}_2\tilde{R}_2}{\tilde{\mu}_2+(1-\tilde{\mu}_2)\beta_2}$

(b) $\tilde{\mu}_1 \leq \tilde{\mu}_2$, $\beta_1 \geq \beta_2$, $\tilde{\mu}_1\tilde{R}_1 \geq \tilde{\mu}_2\tilde{R}_2$ *and* $\tilde{\mu}_1 + \beta_1(1-\tilde{\mu}_1) \geq \tilde{\mu}_2 + \beta_2(1-\tilde{\mu}_2)$.

Then, it is optimal to prioritize stage 1 customers.

When $\tilde{\mu}_1 = \tilde{\mu}_2$, the conditions in Corollary 4.5(b) reduce to $\beta_1 \geq \beta_2$ and $\tilde{R}_1 \geq \tilde{R}_2$, which are consistent with Theorem 3.3 of Down et al. (2011). We extend their result to a multi-server system with different service rates under the discrete time setting in Model II(A).

4.6.3 Special model III: Stage changes in one direction only while in queue

The third special model we considered is that customers do not change their stages in service, and they could change from one stage to another while waiting in the queue. Suppose $p_{12} = p_{21} = 0$, $q_{12} = \beta_1, q_{11} = 1 - \beta_1$ and $q_{22} = \beta_2, q_{20} = 1 - \beta_2$. Stage 1 customers in queue will become stage 2, and stage 2 customers will renege from queue.

Corollary 4.6. *Idling is suboptimal if* $\bar{R}_1 \geq \bar{R}_2$.

The condition $\bar{R}_1 \geq \bar{R}_2$ is equivalent to

$$(1 - \alpha)(R_1\mu_1 - R_2\mu_2) \geq \alpha\mu_1\mu_2(R_2 - R_1).$$

As $\alpha \to 1$, the condition reduces to $R_2 \leq R_1$, and when α is close to 0, the condition requires $R_1\mu_1 \geq R_2\mu_2$. In general, the condition requires the customers in queue will change to a worse stage. When the discount factor is small, the stage with smaller one-step expected reward is considered as the worse stage, and when the discount factor is large the stage with smaller reward is considered as the worse stage.

Corollary 4.7. *If* $R_1\mu_1 \leq R_2\mu_2,\ \mu_1 \geq \beta_1$ *and* $\mu_2 \leq \beta_2$, *then it is optimal to prioritize stage 2 customers.*

Corollary 4.7 is consistent with Proposition 5.2.1 of Jacobson (2010), who considers a single server system with $K \geq 2$ stages. Our results can easily be extended to more than two stages, and we are considering a system with multiple servers.

4.7 Numerical study

In this section, we explore the structure of the optimal policy by means of a numerical study, and compare the performance of different index policies.

The MDP model has an infinite state space, and we will apply truncation on the state space in the numerical study. We truncate the state space in the way that $x_1 + x_2 \leq B$. The truncation error will be ignored in our analysis. Note, the truncation error would be significant for a busy system, i.e., when arrival rates are equal or larger than the service rate, and the reneging rate is very small. Hence, sin our analysis, we will avoid such situations by considering parameters so that the system is not overcrowded.

4.7.1 Switching curve

We start with $b = 1$ and $B = 20$ to look at the structures of the optimal policy. We consider the average model and use relative value iteration to obtain the optimal policy for each scenario.

For the base scenario, we set parameters as following:

(i) The arrival rate $\lambda = 0.15$ and probability that a new arrival belongs to stage 1 is 0.5, hence, $\lambda_1 = \lambda_2 = 0.075$.

(ii) For stage i customers ($i = 1, 2$), the probability of service completion is μ_i, and the probability of queue abandonment is β_i. The patient evolution probabilities are obtained as follows: we let $p_{i0} = \mu_i$ and $q_{i0} = \beta_i$, and $p_{ii} = 0.8(1 - \mu_i), p_{i,3-i} = 0.2(1 - \mu_i)$ and $q_{ii} = 0.8(1 - \beta_i), q_{i,3-i} = 0.2(1 - \beta_i)$.

(iii) $\mu_1 = 0.1, \mu_2 = 0.2$ and $\beta_1 = 0.15, \beta_2 = 0.05$.

(iv) $R_1 = 18$ and $R_2 = 10$.

In this scenario, stage 1 customers have larger service reward, larger abandonment rate in the queue, but smaller service rate, and we find that the optimal policy has a switching curve structure in Figure 4.2. By increasing B to 30, we get the same optimal policy and same optimal long-run average reward, and hence

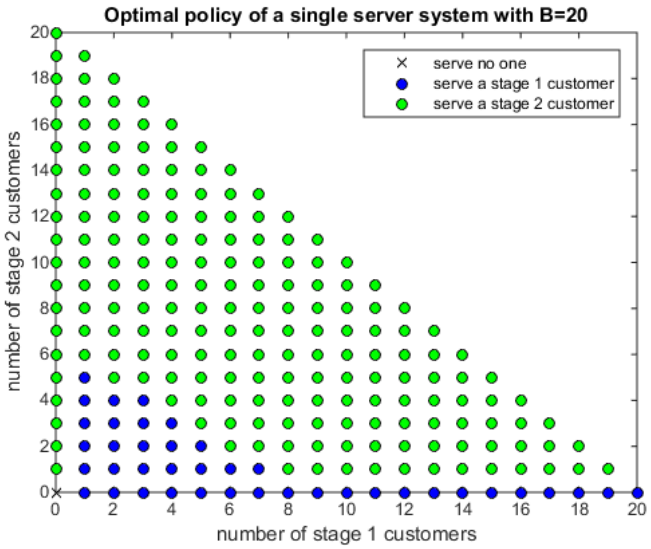

Figure 4.2: Optimal policy structure

we think it is alright to ignore the truncation error for this case.

Next we make changes to only one of the parameters of stage 1 customers (R_1, μ_1 or β_1) and see how the optimal policy changes.

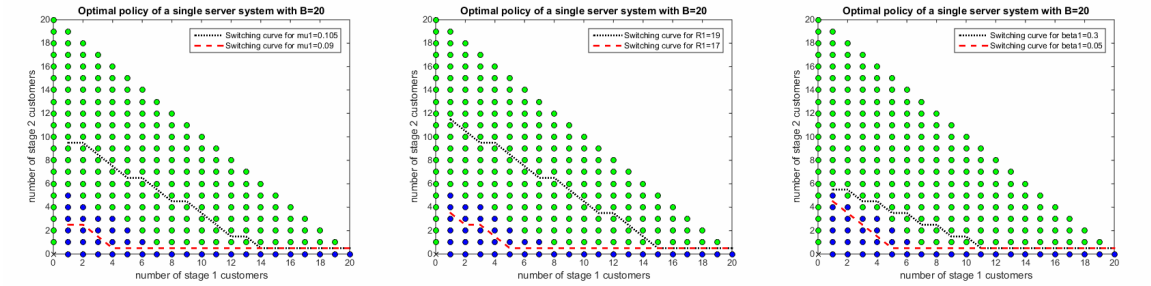

Figure 4.3: Changes of the switching curve in μ_1, R_1, and β_1

In the end of this section, we consider the optimal decision for a multi-server system. Let $b = 3$ and $B = 30$, we also increase the arrival rate by three times, i.e., $\lambda = 0.45$ in this scenario. All other parameters are the same as the base scenario in the single server case.

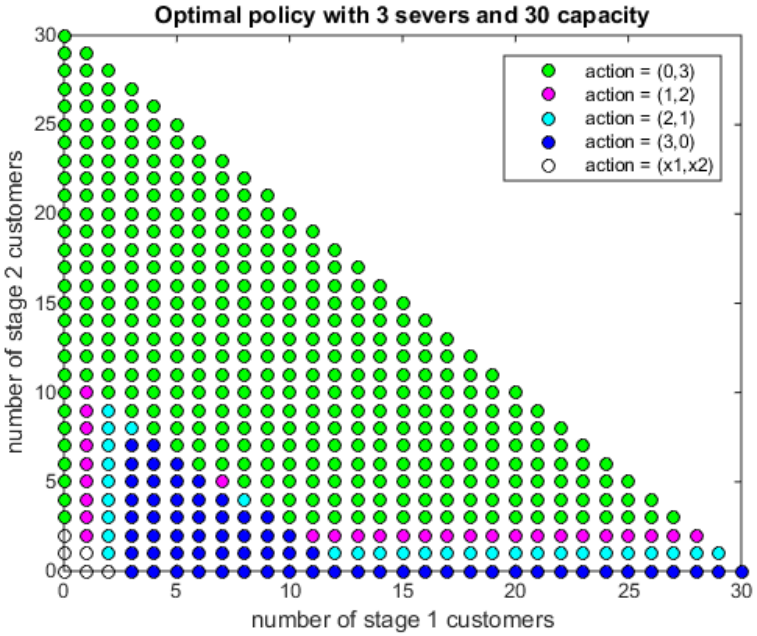

Figure 4.4: Optimal policy structure for a three-server system

4.7.2 Compare the performance of some priority index policies

In this section, we propose some simple priority index policies and compare their performance in a numerical study.

The most straightforward index policy would be the one prioritizes the stage with larger reward R_i (which we refer as the reward policy (R)), or a policy that prioritizes the stage with larger one-step expected

reward (which we refer as the one-step-reward policy (OSR)). Besides, we can also propose some index policies with structure similar to $c\mu$-rule (Cox and Smith (1961)), or $c\mu/\theta$-rule (Atar et al. (2010)). Let L_i denote the expected length of time required to serve a stage i customer. Then, the reward-rate policy (RR) prioritizes the stage with larger R_i/L_i, where L_i could be obtained by solving the following equations:

$$L_i = 1 + \sum_{j=1}^{2} p_{ij} L_j, \text{ for } i = 1, 2.$$

Matrix form: $L = (1,1)' + pL \to (I-p)L = (1,1') \to L = (I-p)^{-1}(1,1)'$.
Similarly, let L_i^Q denote the expected length of time that a stage i customer would wait in the queue before abandon, where L_i^Q could be obtained by solving the following equations:

$$L_i^Q = 1 + \sum_{j=1}^{2} q_{ij} L_j^Q, \text{ for } i = 1, 2.$$

Then, we consider a priority index $R_i L_i^Q / L_i$ that has a similar index structure as $c\mu/\theta$-rule proposed in Atar et al. (2010), which we refer to as reward-rate over abandonment-rate policy (RR/AR). The $c\mu/\theta$-rule is shown to be optimal in the overloaded system. However, intuitive speaking, we would like to prioritize customers with higher abandonment rates when they have the same service rates and service rewards. Hence, it makes more sense to consider a policy that prioritizes the type with larger $R_i/(L_i L_i^Q)$, which we refer to as reward-rate times abandonment-rate (RRAR).

At last, from the condition of non-idling and the sufficient conditions of service prioritization, we propose a policy with index $\bar{R}_i - \bar{R}_i^G$ ($\hat{R}_i - \hat{R}_i^G$ in the average model), which we refer to as expected delayed reward difference policy (EDRD). For this problem, possible index could be R_i, $R_i p_{i0}$ or $R_i p_{i0}/q_{i0}$. These indexes does not take into account the transitions between stages. Hence, we propose an index policy with priority index $\bar{R}_i - \bar{R}_i^Q$ under Assumption 4.2, i.e., we prioritize the stage that have larger one-step incremental survival probabilities.

To summarize, the policies and corresponding indices are shown in Table 4.1.

Table 4.1: Priority index policies and corresponding index

Policy Index	**Notation**
Reward (R)	R_i
One-step Reward (OSR)	$R_i p_{i0}$
Reward rate (RR)	R_i/L_i
Reward rate over abandonment rate (RR/AR)	$R_i L_i^Q / L_i$
Reward rate times abandonment rate (RRAR)	$R_i / L_i L_i^Q$
Expected delayed reward difference (EDRD)	$\bar{R}_i - \bar{R}_i^Q$

4.7.2.1 Single scenario comparison

We compare the policies for the scenarios given in Section 4.7.1. We first consider the base scenario, where we have the optimal reward $g^* = 1.3523$, and the expected rewards for always prioritize stage 1 (policy P1) and always prioritize stage 2 customers (policy P2) are $g_1 = 1.3523$ and $g_2 = 1.2595$. We notice that P1 is better than P2, and performs nearly optimal.

Next, we compute the respective indices for each policy given in Table 4.1, and discuss which policies performs better only for this scenario. We have

Table 4.2: Priority indices for the base scenario.

	R	Rp_{i0}	R_i/L_i	$R_iL_i^Q/L_i$	$R_i/(L_iL_i^Q)$	$\bar{R}_i - \bar{R}_i^Q$
stage 1	18	1.8	2.4	22.1124	0.2605	2.4731
stage 2	10	2	1.6364	18.7538	0.1428	0.1788

All the policies in Table 4.1 choose the right static policy except for policy OSR, which uses R_ip_{i0} as the priority index. We repeat these computations for all other scenarios in Section 4.7 with changed parameters.

Table 4.3: Computations of the priority indices for several scenarios

Scenario	g*	Stage	g1	R	Rmu	R/L	R*L^Q/L	R/(L^Q*L)	barR-barR^Q
r1=19	1.4021	1	1.4021	19	1.9	2.5333	23.3408	0.275	2.5951
		2	1.2858	10	2	1.6364	18.7538	0.1428	0.1375
r1=17	1.3029	1	1.3025	17	1.7	2.2667	20.8839	0.246	2.351
		2	1.2332	10	2	1.6364	18.7538	0.1428	0.2201
beta1=0.3	1.291	1	1.291	18	1.8	2.4	11.5443	0.4989	4.4922
		2	1.1734	10	2	1.6364	13.0495	0.2052	0.1788
beta1=0.05	1.4312	1	1.4306	18	1.8	2.4	48	0.12	1.1269
		2	1.3843	10	2	1.6364	32.7273	0.0818	0.1788
mu1=0.105	1.3835	1	1.3835	18	1.89	2.4579	22.6457	0.2668	2.5002
		2	1.274	10	2	1.6577	18.9979	0.1446	0.1696
mu1=0.9	1.288	1	1.2861	18	1.62	2.2849	21.0516	0.248	2.4151
		2	1.2296	10	2	1.5926	18.2522	0.139	0.1984

4.7.2.2 Random comparisons

The single scenario we presented in the previous section may not tell the whole story. Next, we will compare these policies with randomly generated parameters. We randomly generate 1000 sets of parameters, with all probabilities μ_i and β_i uniformly in $(0.1, 0.5)$, $p_{ii} \sim (1-\mu_i)U(0.8, 1)$ and $q_{ii} \sim (1-\beta_i)U(0.8, 1)$. We generate λ and *theta* uniformly in $(0, 1)$ so that $\lambda_1 = \theta]lambda$ and $\lambda_2 = \lambda(1-\theta)$. We fix $R_1 = 10$ and R_2

takes value in $\{10, 12, 14, 16, 18, 20\}$ for each set of the parameters. We generated a total of 6000 scenarios, and the following analysis will be conducted based on these 6000 scenarios.

We first look at how each static policy performs compared with the optimal policies. Let P1 denote the static policy that always prioritize stage 1 customers, and P2 denote the policy that always prioritize stage 2 customers. Let g_1 and g_2 denote the respective long-run average reward for policy P1 and P2, and g^* denote the long-run average reward under the optimal policy. We conduct 95% confidence intervals (C.I.) of the relative differences $(g^* - g_1)/g^*$, $(g^* - g_2)/g^*$ and $(g^* - \max(g_1, g_2))/g^*$, and the results are shown in Table 4.4 (all numbers are in percentage).

Table 4.4: Comparisons of the static policies and the optimal policy

	$(g^* - g_1)/g^*$	$(g^* - g_2)/g^*$	$(g^* - \max(g_1, g_2))/g^*$
95% C.I.	(6.6360, 7.2617)	(7.1521, 7.7919)	(0.0048, 0.0074)
Maximum	69.8460	72.1264	1.2222

We find that the best static policy performs nearly optimal. However, given the customers change stages in both service and queue, it is difficult to analytically compute and compare the long-run average rewards of P1 and P2. Hence, we focus on whether we could find some criteria that is easy to employ and determine which static policy should be chosen. The priority index policies provides in Table 4.1 are some of such criteria, and we next compare the performances of these policies. As before, we first provide the 95% C.I. of the relative differences in long-run average reward under proposed policy and the optimal policy, the maximum of the relative difference, and we also compute the percentage of scenarios that the proposed policy is consistent with the best static policy. More specifically, we conduct C.I.s and find maximum of the quantity $(g^* - g_\pi)/g^*$ where g_π denote the long-run average reward for policy π. Let N_π denote the number of scenarios that policy π is consistent with the best static policy, and we compute $N/6000$ as the percentage of scenarios that the proposed policy is consistent with the best static policy. The results are provided in Table 4.5 (all numbers are in percentage).

Table 4.5: Comparisons of the proposed index policies

Policy	Index	95% C.I. of $(g^* - g_\pi)/g^*$	Maximum of $(g^* - g_\pi)/g^*$	$N_\pi/6000$
R	R_i	(4.24, 4.73)	60.64	65.08
OSR	$R_i\mu_i$	**(0.08, 0.10)**	**6.35**	**93.05**
RR	R_i/L_i	(0.21, 0.28)	28.05	90.98
RR/AR	$R_iL_i^Q/L_i$	(2.49, 2.83)	55.89	69.97
RRAR	$R_i/(L_iL_i^Q)$	(1.22, 1.46)	48.10	84.58
EDRD	$\bar{R}_i - \bar{R}_i^Q$	(5.12, 5.69)	69.85	63.77

We found that in this study, the one-step-reward policy outperforms all the other index policies.

4.8 Conclusions

Customers are impatient and may change stages in many service systems. There are many parameters that could affect the prioritization decisions, such as service rates and rewards, abandonment rates, transition probabilities, etc. We aim to develop a model to see how the priority should be assigned under different settings of these parameters, and propose some priority index policies based on our model analysis.

We first formulate the system an MDP model. We find that our formulation does not guarantee the optimality of non-idling, which should be the case in reality for most service systems. We consider the case when there are ample servers and we find the necessary and sufficient condition under which non-idling should be optimal. Intuitive explanation for this condition is that delaying the service for one period of time will reduce the expected discounted reward obtained for every customer. We next consider the general case when server capacity is constrained. With an additional assumption on the transition probabilities (may not be necessary), we are able to provide a sufficient condition under which non-idling is optimal. Next, we compare the actions of serving a stage 1 customers versus the action of serving a stage 2 when a server is available, and find that the action of service a stage 1 is always better than serving a stage 2 when a server is available under certain conditions. Hence, these conditions are sufficient conditions under which it is optimal to always prioritize stage 1 customers.

We consider several special settings of the general model. The first setting is to assume no queueing for the system. Then, the model reduces to **Model I** that is very similar to the one in Chapter 3, with slightly different reward structure. However, utilizing the similar structure of the optimality equations, we conclude that the optimal policy for **Model I** is actually the same as the ICU model in Chapter 3. The second special parameter setting considered in **Model II** is the case when there is no transition between stages. We obtain

a sufficient condition under which always prioritizing one stage is optimal by applying our general model results. We also extend the result by considering reneging also happens for customers in service in **Model II(A)**, and obtain a result that is consistent with that in Down et al. (2011), but our **Model II(A)** is more general since we consider a multi-sever and non-identical service times. The third special parameter setting study the model in Chapter 5.2 of Jacobson (2010), and provide a similar result for a multi-server system.

We propose several index policies from our model analysis, and conduct a numerical study to compare the performances. In this study, we first find the switching curve structure of the optimal policy and observe how the switching curve changes with respect to different parameters. We conduct 6000 random parameter combinations and compare the performance of different index policies. We find that in our study, the priority index policy that prioritizes the stage with largest one-step expected reward performs the best among all index policies we proposed.

CHAPTER 5: CONCLUSIONS

Prioritization has been widely applied in many service systems to improve the overall system performances and to provide customized service to heterogeneous customers. In this dissertation, our main objective is to study how priority should be assigned in different systems when the cost is not of a simple linear form.

In Chapter 2, we compare several static policies in an M/G/1 queueing system with nonlinear waiting cost functions. The results we obtained are very general and could provide useful insights for the case when a simple policy is desired. Our first result provides a complete comparison of$F$, PF_1 and PF_2 when the exact cost function expressions for both types are known. Then, we provide sufficient conditions under which FCFS performs better and conditions under which the fixed priority policy performs better. These conditions could apply to the case when the cost function for one type is known and we have knowledge of the changing rate of the other cost function. We next show that when the cost functions are both convex it is sufficient to compare only F, PF_1 and PF_2. On the other hand, when the cost functions are both concave it is sufficient to only compare L, PL_1 and PL_2, and we also present results on the comparisons of these three policies. We find that the best static policy performs very well for most scenarios in our simulation study compared with the Generalized-$c\mu$ rule.

In Chapter 3, we investigate the prioritization decisions in ICUs. We assume patients health stages could change over time, and the priority decisions are made between different stages. Although we proved the optimal policy depends on the mix of the patients in the ICU, our numerical study shows that the difference between the best static policy and the optimal dynamic policy is very small. Hence, we propose a policy that determines which of the two stages should be prioritized based on the system load. Our proposed policy only needs information about the stage-dependent expected dying probabilities in ICU and in the general ward, and the stage-dependent expected length-of-stay, which could be estimated from data in practice, and hence is applicable.

In Chapter 4, we study the prioritization problem in a queueing system under the assumption that customers are impatient and they could change their stages over time. We model this problem as an MDP, and provide partial characterization of the optimal policy. We compare the performances of several index-based priority policies via a numerical study.

There are several possible extensions of the work on the models in Chapter 3 and 4. For example, we assume all patients are evolving with respect to the same pattern in the current models, and then an instant extension is to assume patients belong to different classes and each class associates with a transition scheme. The prioritization decisions are made among patient classes and stages. For the model in Chapter 4, we have demonstrated how this model could be applied to analyze the ICU model with readmissions. However, we did not differentiate between the readmitted patients and new arrivals, where many empirical studies have shown that readmission is associated with a longer length of stay and higher mortality. Hence, we could incorporate these differences in the future study. These models become very complicated and we should focus on either finding an index based priority assignment policy that performs well or applying fluid approximation to provide some insights.

APPENDIX A: PROOFS OF RESULTS IN CHAPTER 2

In this Appendix, we provide proofs of results in Chapter 2.

A.1 Proof of Results in Section 2.2

Proof of equivalence of Equations (2.1) ***and*** (2.2)***:*** The long-run average cost defined by (2.1) can be written as

$$\begin{aligned} C_\pi &= \sum_{i=1}^{2} \lim_{t\to\infty} \left(\frac{\sum_{k=1}^{n_i(t)} C_i(V_{i,k}^{\pi,x_0})}{n_i(t)} \right) \left(\frac{n_i(t)}{t} \right) \\ &= \sum_{i=1}^{2} \lim_{t\to\infty} \frac{\sum_{k=1}^{n_i(t)} C_i(V_{i,k}^{\pi,x_0})}{n_i(t)} \lim_{t\to\infty} \frac{n_i(t)}{t} = \sum_{i=1}^{2} \lambda p_i \lim_{n\to\infty} \frac{\sum_{k=1}^{n} C_i(V_{i,k}^{\pi,x_0})}{n}, \end{aligned} \tag{A.1}$$

which follows from the fact that $\{n_i(t), t \geq 0\}$ is a Poisson process with rate λp_i for $i = 1, 2$. In the following we will prove that for $i = 1, 2$ when $E\left[\left|C_i(W_i^\pi)\right|\right]$ is finite,

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} C_i(V_{i,k}^{\pi,x_0})}{n} = E\left[C_i(W_i^\pi)\right], \tag{A.2}$$

which shows that (A.1) (and hence (2.1)) is equivalent to (2.2).

In the remainder of this proof, we drop the superscripts π and x_0 for notational convenience, and let T_{ik}, S_{ik} and D_{ik} be the arrival time, service time and departure time of the kth type i customer, respectively, under policy π and initial state x_0. Then, $V_{ik} = D_{ik} - T_{ik} - S_{ik}$ is the queue-waiting time for this customer. Note that $\{V_{ik}, k = 1, 2, \ldots\}$ for each $i = 1, 2$ is a delayed regenerative process with nth regeneration happening at $N_{i,n}$ for $n = 0, 1, 2, \ldots$, where $N_{i,0} = 1$, and

$$N_{i,n} = \min\{k : k > N_{i,n-1}, V_{ik} = 0\}.$$

Note also that for each $i = 1, 2$, $\{C_i(V_{ik}), k = 1, 2, \ldots\}$ is a regenerative process with the same regeneration epoches as $\{V_{ik}, k = 1, 2, \ldots\}$. Then, by Theorem 13 of Chapter 2 and last paragraph of page 93 in (Wolff 1989), (A.2) holds if $\sum_{k=1}^{N_{i,1}-1} |C_i(V_{ik})| < \infty$ with probability one, $E[N_{i,2} - N_{i,1}] < \infty$, and $E\left[\sum_{k=N_{i,1}}^{N_{i,2}-1} |C_i(V_{ik})|\right] < \infty$. We next complete the proof by showing that these three conditions hold.

When $\rho < 1$, the system is stable, i.e., it will return to the empty state within finite time with probability one and also the expected time it takes to return to the empty state is finite (see, e.g., Theorem 7.11 in (Kulkarni 2009)). This implies that $N_{i,1} < \infty$ with probability one, $N_{i,2} - N_{i,1} < \infty$ with probability one, $V_{i,k} < \infty$ for any i and k with probability one and $E[N_{i,2} - N_{i,1}] < \infty$. At last, by Theorem B.5 (i) in (El-

Taha and Stidham 1999), $E\left[\sum_{k=N_{i,1}}^{N_{i,2}-1}|C_i(V_{ik})|\right] = E\left[|C_i(W_i)|\right]E\left[N_{i,2}-N_{i,1}\right]$ is finite under the assumption that $E\left[|C_i(W_i)|\right]$ is finite. □

A.2 Proof of results in Section 2.3

Proof of Lemma 2.3: We use sample path arguments to prove the stochastic inequalities. Let i be fixed to be either 1 or 2. Here type i and $3-i$ customers will be called priority and non-priority customers, respectively.

We index the customers by their arrival order to the system, and let s_j be the arriving time of customer j. Then, for customers l and j, where $j > l \geq 1$, we have $s_j > s_l$. Let t_j^π be the service starting time of customer j under policy π, then $t_j^\pi \geq s_j$. Let also V_j^π denote the waiting time of customer j under policy π, then $V_j^\pi = t_j^\pi - s_j$ for $j = 1, 2, \ldots$.

Under FCFS, we have $t_1^F < t_2^F < \cdots$ with probability one. Let j be the index of the first non-priority customer whose service starts when there are priority customers waiting, and k be the index of the first priority customer in the queue when j starts service under FCFS. Then, the customers indexed from j to $k-1$ are all non-priority customers. Note that $s_j < \cdots < s_{k-1} < s_k < t_j^F < \cdots < t_{k-1}^F < t_k^F$.

Consider a policy π that follows FCFS except that it serves customer k first, and then serves the non-priority customers $j, \ldots, k-1$. For the kth customer, who is a priority customer, $t_k^\pi = t_j^F < t_k^F$ and $V_k^\pi = t_k^\pi - s_k < t_k^F - s_k = V_k^F$. For $l = j, \ldots, k-1$, who are all non-priority customers, $t_l^\pi > t_l^F$ and $V_l^\pi = t_l^\pi - s_l > t_l^F - s_l = V_l^F$. For any $l \notin \{j, \ldots, k\}$, we have $V_l^\pi = V_l^F$.

If we keep changing the service order like this when there are non-priority customers starting service while priority customers are waiting in the queue, then we will eventually reach policy PF_i. This coupling argument then will yield $V_{i,n}^{PF_i} \leq_{st} V_{i,n}^F$ and $V_{3-i,n}^{PF_i} \geq_{st} V_{3-i,n}^F$ for $n \geq 1$. Since W_i^π is the steady-state waiting time for type i customers under policy π, then, as $n \to \infty$, $V_{i,n}^\pi \xrightarrow{d} W_i^\pi$ and $V_{3-i,n}^\pi \xrightarrow{d} W_{3-i}^\pi$, and hence, according to Theorem 1.A.3(d) in (Shaked and Shanthikumar 2007), we have $W_i^{PF_i} \leq_{st} W^F$ and $W_{3-i}^{PF_i} \geq_{st} W^F$.

□

Proof of Theorem 2.1: We prove this result by comparing the costs directly.

(a) For $i = 1, 2$, Equation (2.2) yields $C_F \leq C_{PF_i}$ if and only if

$$\begin{aligned} & p_i\Big(E\big[C_i(W^F)\big] - E\big[C_i(W_i^{PF_i})\big]\Big) \leq p_{3-i}\Big(E\big[C_{3-i}(W_{3-i}^{PF_i})\big] - E\big[C_{3-i}(W^F)\big]\Big) \\ \Leftrightarrow \quad & p_i E\big[C_i'(U_i^{PF_i})\big]\big(E[W^F] - E[W_i^{PF_i}]\big) \leq p_{3-i}E\big[C_{3-i}'(U_{3-i}^{PF_i})\big]\big(E[W_{3-i}^{PF_i}] - E[W^F]\big) \end{aligned}$$

based on Lemma 2.1. Since

$$\frac{p_{3-i}\big(E[W_{3-i}^{PF_i}]-E[W^F]\big)}{p_i\big(E[W^F]-E[W_i^{PF_i}]\big)}=\frac{p_{3-i}\left(\frac{1}{(1-\rho)(1-\rho_i)}-\frac{1}{1-\rho}\right)}{p_i\left(\frac{1}{1-\rho}-\frac{1}{1-\rho_i}\right)}=\frac{p_{3-i}\rho_i}{p_i\rho_{3-i}}=\frac{\tau_i}{\tau_{3-i}},$$

we have $C_F \le C_{PF_i}$ if and only if $a_i \le b_i$.

(b) Equation (2.2) yields $C_{PF_1} \le C_{PF_2}$ if and only

$$p_2\Big(E\big[C_2(W_2^{PF_1})\big]-E\big[C_2(W_2^{PF_2})\big]\Big)\le p_1\Big(E\big[C_1(W_1^{PF_2})\big]-E\big[C_1(W_1^{PF_1})\big]\Big). \tag{A.3}$$

We have,

$$\begin{aligned}
&p_2\left(E\big[C_2(W_2^{PF_1})\big]-E\big[C_2(W_2^{PF_2})\big]\right)\\
=&p_2\left(E\big[C_2(W_2^{PF_1})\big]-E\big[C_2(W^F)\big]+E\big[C_2(W^F)\big]-E\big[C_2(W_2^{PF_2})\big]\right)\\
=&p_2\left(\big(E[W_2^{PF_1}]-E[W^F]\big)E\big[C_2'(U_2^{PF_1})\big]+\big(E[W_2^F]-E[W_2^{PF_2}]\big)E\big[C_2'(U_2^{PF_2})\big]\right) \qquad \text{(A.4)}\\
=&p_2\left(\frac{\lambda\rho_1\bar{\xi}}{2(1-\rho_1)(1-\rho)}E\big[C_2'(U_2^{PF_1})\big]+\frac{\lambda\rho_1\bar{\xi}}{2(1-\rho_2)(1-\rho)}E\big[C_2'(U_2^{PF_2})\big]\right)\\
=&\frac{\tau_2 p_2\lambda\rho_1\bar{\xi}}{2(1-\rho_1)(1-\rho_2)(1-\rho)}\left[(1-\rho_2)\frac{E\big[C_2'(U_2^{PF_1})\big]}{\tau_2}+(1-\rho_1)\frac{E\big[C_2'(U_2^{PF_2})\big]}{\tau_2}\right],\\
=&\frac{\rho_1\rho_2\bar{\xi}}{2(1-\rho_1)(1-\rho_2)(1-\rho)}\big[(1-\rho_2)b_1+(1-\rho_1)a_2\big]. \qquad \text{(A.5)}
\end{aligned}$$

□

Proof of Corollary 2.2: (a) If $C_1'(t) \ge \tau_1 \max\{a_2, b_1\}$ for all $t \ge 0$, then for any non-negative random variable X, we have $E\left[C_1'(X)\right] \ge \tau_1 \max\{a_2, b_1\}$ when the expectation exists. Hence,

$$a_1=\frac{E\left[C_1'(U_1^{PF_1})\right]}{\tau_1}\ge\frac{\tau_1\max\{a_2,b_1\}}{\tau_1}\ge b_1,$$

which implies that $C_F \ge C_{PF_1}$ from Theorem 2.1(a). Similarly,

$$b_2=\frac{E\left[C_1'(U_1^{PF_2})\right]}{\tau_1}\ge\frac{\tau_1\max\{a_2,b_1\}}{\tau_1}\ge a_2,$$

which implies that $C_F \le C_{PF_2}$ from Theorem 2.1(a).

(b) If $C_1'(t) \leq \tau_1 \min\{a_2, b_1\}$ for all $t \geq 0$, then we have $E[C_1'(X)] \leq \tau_1 \min\{a_2, b_1\}$ for any non-negative random variable X when the expectation exists. Then, $a_1 \leq b_1$ and $b_2 \leq a_2$, which implies that $C_F \leq C_{PF_1}$ and $C_F \geq C_{PF_2}$ from Theorem 2.1(a).

(c) If $\tau_1 a_2 \leq C_1'(t) \leq \tau_1 b_1$ for all $t \geq 0$, then we have $\tau_1 a_2 \leq E[C_1'(X)] \leq \tau_1 b_1$ for any non-negative random variable X when the expectation exists. Then, $a_i \leq b_i$ for $i = 1, 2$, which implies that $C_F \leq C_{PF_1}$ and $C_F \leq C_{PF_2}$ from Theorem 2.1(a).

□

Proof of Corollary 2.3: (a) Since $C_1'(t) \geq \max\{\alpha, \beta\} C_2'(t)$ for all $t \geq 0$, then for any non-negative random variable X, we have $E[C_1'(X)] \geq \max\{\alpha, \beta\} E[C_2'(X)]$ when the expectations exist. Consequently, for $X = U_1^{PF_1}$ we have

$$E[C_1'(U_1^{PF_1})] \geq \beta E[C_2'(U_1^{PF_1})] = \left(\frac{\tau_1}{\tau_2}\right) E[C_2'(U_2^{PF_1})] \Leftrightarrow a_1 \geq b_1,$$

and hence by Theorem 2.1(a), $C_{PF_1} \leq C_F$. Similarly, for $X = U_1^{PF_2}$ we have

$$E[C_1'(U_1^{PF_2})] \geq \alpha E[C_2'(U_1^{PF_2})] = \left(\frac{\tau_1}{\tau_2}\right) E[C_2'(U_2^{PF_2})] \Leftrightarrow b_2 \geq a_2,$$

and hence by Theorem 2.1(a), $C_F \leq C_{PF_2}$.

(b) Similar to part (a), since $C_1'(t) \leq \min\{\alpha, \beta\} C_2'(t)$ for all $t \geq 0$, we have $a_1 \leq b_1$ and $b_2 \leq a_2$. Thus, by Theorem 2.1(a), we have $C_{PF_2} \leq C_F \leq C_{PF_1}$.

(c) Similar to part (a), since $\alpha C_2'(t) \leq C_1'(t) \leq \beta C_2'(t)$ for all $t \geq 0$, we have $a_1 \leq b_1$ and $a_2 \leq b_2$, which implies that $C_F \leq C_{PF_1}$ and $C_F \leq C_{PF_2}$ by Theorem 2.1(a).

□

A.3 Laplace-Stieltjies transforms (LSTs) of the busy period and waiting times under FCFS and fixed priority policies

For $i = 1, 2$, let $\tilde{S}_i(s)$ denote the LST of the service time distribution for type i customers, and $\tilde{S}(s) \equiv p_1 \tilde{S}_1(s) + p_2 \tilde{S}_2(s)$. We first cite the definition of a busy period from (Miller 1960) as following.

Definition A.1. ((Miller 1960)) The length of a busy period is the length of time between the arrival of an item at the empty queue and the first subsequent moment at which the queue is again empty.

The LST of the busy period of a single-class queue was introduced in Theorem 6 of (Takács 1955), which we also present in the next lemma.

Lemma A.1. *(Theorem 6 of (Takács 1955)) Let $B(s)$ denote the LST of the busy period distribution function, then $B(s)$ is the uniquely defined analytic solution of the equation*

$$B(s) = \tilde{S}\big(s + \lambda(1 - B(s))\big),$$

where $\lim_{s\to\infty} B(s) = 0$.

For a single-serve queue with two classes of customers, we define $B_j(s)$ as a busy period with only class j arrival. Then, we can adapt the result from Lemma A.1 that $B_j(s)$ is the unique solution to $B_j(s) = \tilde{S}_j\big(s + \lambda p_j(1 - B_j(s))\big)$ for $\mathrm{Re}\{s\} > 0$ and $\lim_{s\to\infty, s \text{ real}} B_j(s) = 0$.

We next obtain the following result from Equations (3.3), (3.8), (3.10), (4.1) and (4.2) of (Miller 1960).

Lemma A.2. *Let $\widetilde{W}^F$ and $\widetilde{W}_i^{PF_j}$ denote the respective LST of W^F and $W_i^{PF_j}$ for $i, j \in \{1, 2\}$. Then, for fixed $j \in \{1, 2\}$,*

$$\widetilde{W}_j^{PF_j}(s) = \frac{(1-\rho)s + \lambda p_{3-j}[1 - \tilde{S}_{3-j}(s)]}{s - \lambda p_j \left(1 - \tilde{S}_j(s)\right)},\ \widetilde{W}^F(s) = \frac{(1-\rho)s}{s - \lambda\left(1 - \tilde{S}(s)\right)},$$

and

$$\widetilde{W}_{3-j}^{PF_j}(s) = \widetilde{W}^F\big(\lambda p_j(1 - B_j(s)) + s\big) = \frac{1-\rho}{1 - \frac{\lambda\left(1-\tilde{S}\left(\lambda p_j(1-B_j(s))+s\right)\right)}{\lambda p_j(1-B_j(s))+s}}.$$

A.4 Proof of results in Section 2.4

Proof of Lemma 2.4. Assumption 2.1 holds for $C_i(t)$ in the form of (2.5) if and only if $E\left[\left(W^F\right)^l\right]$ and $E\left[\left(W_i^{PF_m}\right)^l\right]$ are finite for all $m \in \{1, 2\}$ and $l \le j(i)$.

In fact, with Lemma 2.3 and Theorem 1.A.3(a) of (Shaked and Shanthikumar 2007), we have, for any finite l,

$$E\left[\left(W_i^{PF_i}\right)^l\right] \le E\left[\left(W^F\right)^l\right] \le E\left[\left(W_i^{PF_{3-i}}\right)^l\right],$$

and thus we only need to prove that $E\left[\left(W_i^{PF_{3-i}}\right)^l\right]$ exists. We have,

$$E\left[\left(W_i^{PF_{3-i}}\right)^l\right] = \frac{d^l \widetilde{W}_i^{PF_{3-i}}(s)}{ds^l}\Big|_{s=0},$$

and from Lemma A.2, we have,

$$\widetilde{W}_i^{PF_{3-i}}(s) = \widetilde{W}^F\big(\lambda p_{3-i}(1 - B_{3-i}(s)) + s\big).$$

Using the Faa di Bruno's formula (see, e.g., Theorem 2 of (Roman 1980)), we have, $\frac{d^l \widetilde{W}_i^{PF_{3-i}}(s)}{ds^l}|_{s=0}$ will be finite if $\frac{d^n \widetilde{W}^F(s)}{ds^n}|_{s=0}$ and $\frac{d^n B_{3-i}(s)}{ds^n}|_{s=0}$ are finite for all $n \leq l$, which is true if the nth moment of W^F and the nth moment of the busy period are finite.

When $\rho < 1$, we can obtain the nth moments of W^F from (Gross et al. 2008) (page 238) as

$$E\left[W^F\right]^n = \frac{\lambda}{1-\rho} \sum_{j=1}^{n} E\left[W^F\right]^{n-j} \frac{E\left[S\right]^{j+1}}{j+1},$$

where $E\left[S\right]^{j+1}$ is $(j+1)$st moment of service times, and hence $E\left[W^F\right]^n$ is finite if $\rho < 1$ and the first $(n+1)$ moments of service times for all customers are finite. Besides, from Theorem 1 of (Ghahramani and Wolff 1989), the nth moment of the busy period is finite if and only if the nth moment of the service times is finite. Thus, $E\left[\left(W_i^{PF_{3-i}}\right)^l\right]$ is finite if $\rho < 1$ and the first $(l+1)$ moments of service times are finite.

□

Proof of Equations (2.8) ***and*** (2.9)***:*** For some $i, k, m \in \{1, 2\}$, we have

$$E\big[C_i'(U_k^{PF_m})\big] = 2k_i E\left[U_k^{PF_m}\right] + h_i = k_i \left(\frac{E\big[(W^F)^2\big] - E\big[(W_k^{PF_m})^2\big]}{E[W^F] - E[W_k^{PF_m}]} \right) + h_i,$$

from Equations (2.6) and (2.7). The expected waiting times have been given in Lemma 2.2, and the second moments can be obtained from (Gross et al. 2008) and (Miller 1960):

$$E\big[(W^F)^2\big] = \frac{\lambda\bar{\zeta}}{3(1-\rho)} + \frac{\lambda^2\bar{\xi}^2}{2(1-\rho)^2},\ E\big[(W_k^{PF_k})^2\big] = \frac{\lambda\bar{\zeta}}{3(1-\rho_k)} + \frac{\lambda^2 p_k \xi_k \bar{\xi}}{2(1-\rho_k)^2},$$

and

$$E\big[(W_{3-k}^{PF_k})^2\big] = \frac{\lambda\bar{\zeta}}{3(1-\rho_k)^2(1-\rho)} + \frac{\lambda^2\bar{\xi}^2}{2(1-\rho_k)^2(1-\rho)^2} + \frac{\lambda^2 p_k \xi_k \bar{\xi}}{2(1-\rho_k)^3(1-\rho)}.$$

Then,

$$E[W^F] - E[W_k^{PF_k}] = \frac{\lambda^2 p_{3-k}\tau_{3-k}\bar{\xi}}{2(1-\rho_k)(1-\rho)},\ E[W_{3-k}^{PF_k}] - E[W^F] = \frac{\lambda^2 p_k \tau_k \bar{\xi}}{2(1-\rho_k)(1-\rho)},$$

and

$$
\begin{aligned}
&E\left[\left(W^F\right)^2\right]-E\left[\left(W_k^{PF_k}\right)^2\right]\\
=&\frac{\rho_{3-k}\lambda\bar{\zeta}}{3(1-\rho_k)(1-\rho)}+\frac{\lambda^2\bar{\xi}}{2(1-\rho_k)(1-\rho)}\left[\frac{(1-\rho_k)(p_1\xi_1+p_2\xi_2)}{1-\rho}-\frac{p_k\xi_k(1-\rho)}{1-\rho_k}\right]\\
=&\frac{\lambda^2 p_{3-k}\tau_{3-k}\bar{\zeta}}{3(1-\rho_k)(1-\rho)}+\frac{\lambda^2 p_{3-k}\bar{\xi}}{2(1-\rho_k)(1-\rho)}\left[\frac{p_k\xi_k\lambda\tau_{3-k}(2-\rho-\rho_k)}{(1-\rho)(1-\rho_k)}+\frac{\xi_{3-k}(1-\rho_k)}{1-\rho}\right],\\
&E\left[\left(W_{3-k}^{PF_k}\right)^2\right]-E\left[\left(W^F\right)^2\right]\\
=&\left[\frac{\bar{\zeta}}{3}+\frac{\lambda\bar{\xi}^2}{2(1-\rho)}\right]\frac{\lambda\rho_k(2-\rho_k)}{(1-\rho)(1-\rho_k)^2}+\frac{\lambda^2 p_k\xi_k\bar{\xi}}{2(1-\rho_k)^3(1-\rho)}\\
=&\left[\frac{2\bar{\zeta}(2-\rho_k)}{3\bar{\xi}(1-\rho_k)}+\frac{\lambda\bar{\xi}(2-\rho_k)}{(1-\rho)(1-\rho_k)}\right]\frac{\lambda\rho_k\bar{\xi}}{2(1-\rho_k)(1-\rho)}+\frac{\lambda\rho_k\xi_k\bar{\xi}}{2\tau_k(1-\rho_k)^3(1-\rho)}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
E\big[C_i'(U_k^{PF_k})\big] &= k_i\left[\frac{2\bar{\zeta}}{3\bar{\xi}}+\frac{p_k\xi_k\lambda(2-\rho-\rho_k)}{(1-\rho)(1-\rho_k)}+\frac{\xi_{3-k}(1-\rho_k)}{\tau_{3-k}(1-\rho)}\right]+h_i\\
&= k_i\left[\frac{2\bar{\zeta}}{3\bar{\xi}}+\lambda p_k\xi_k\left(\frac{1}{1-\rho}+\frac{1}{1-\rho_k}\right)+\lambda p_{3-k}\xi_{3-k}\left(\frac{1}{1-\rho}+\frac{1}{\rho_{3-k}}\right)\right]+h_i\\
&= k_i\left[\frac{2\bar{\zeta}}{3\bar{\xi}}+\frac{\lambda\bar{\xi}}{1-\rho}+\frac{\lambda p_k\xi_k}{1-\rho_k}+\frac{\xi_{3-k}}{\tau_{3-k}}\right]+h_i, \qquad \text{(A.6)}
\end{aligned}
$$

and

$$
E\big[C_i'(U_{3-k}^{PF_k})\big] = k_i\left[\frac{2\bar{\zeta}}{3\bar{\xi}}\left(1+\frac{1}{1-\rho_k}\right)+\frac{\lambda\bar{\xi}}{1-\rho}\left(1+\frac{1}{1-\rho_k}\right)+\frac{\xi_k}{\tau_k(1-\rho_k)^2}\right]+h_i. \qquad \text{(A.7)}
$$

□

Proof of Proposition 2.1. The expressions for A and B and the characterization of the optimal policy follow from (2.11). We next prove that $A<B$. Let $G_i(\lambda)=\frac{2\bar{\zeta}}{3\bar{\xi}}+\frac{\lambda\bar{\xi}}{1-\lambda\bar{\tau}}+\frac{\lambda p_i\xi_i}{1-\lambda p_i\tau_i}$ and $X_i=\frac{2\bar{\zeta}}{3\bar{\xi}}+\frac{\lambda\bar{\xi}}{1-\lambda\bar{\tau}}+\frac{\xi_i}{\tau_i}$ for $i=1,2$. Then, we have

$$
A=\frac{X_1+\frac{\lambda p_2\xi_2}{1-\rho_2}}{X_2+\frac{G_2(\lambda)+\lambda p_2\xi_2}{1-\rho_2}},\ B=\frac{X_1+\frac{G_1(\lambda)+\lambda p_1\xi_1}{1-\rho_1}}{X_2+\frac{\lambda p_1\xi_1}{1-\rho_1}}.
$$

Note that for $i=1,2$,

$$
\frac{G_i(\lambda)+\lambda p_i\xi_i}{1-\rho_i}>\frac{G_i(\lambda)}{1-\rho_i}>\frac{\lambda\bar{\xi}}{(1-\rho_i)(1-\rho)}>\frac{\lambda p_{3-i}\xi_{3-i}}{1-\rho_{3-i}},
$$

where the last inequality follows from the fact that $\bar{\xi} > p_{3-i}\xi_{3-i}$ and $(1-\rho_i)(1-\rho) < 1-\rho < 1-\rho_{3-i}$. Hence, we have

$$A < \frac{X_1 + \frac{\lambda p_2 \xi_2}{1-\rho_2}}{X_2 + \frac{\lambda p_1 \xi_1}{1-\rho_1}} < B.$$

□

Proof of Proposition 2.2. From the expression of A, we have,

$$\begin{aligned}\frac{\partial A}{\partial \lambda} &= \frac{G_2'(\lambda)\left(\frac{2-\rho_2}{1-\rho_2}G_2(\lambda) + \frac{\xi_2}{\tau_2}\right) - \left(G_2(\lambda) + \frac{\xi_1}{\tau_1}\right)\left(\frac{p_2\tau_2}{(1-\rho_2)^2}G_2(\lambda) + \frac{2-\rho_2}{1-\rho_2}G_2'(\lambda)\right)}{\left(\frac{2-\rho_2}{1-\rho_2}G_2(\lambda) + \frac{\xi_2}{\tau_2}\right)^2} \\ &= \frac{G_2'(\lambda)\left(\frac{\xi_2}{\tau_2} - \frac{(2-\rho_2)\xi_1}{(1-\rho_2)\tau_1}\right) - \left(G_2(\lambda) + \frac{\xi_1}{\tau_1}\right)\frac{p_2\tau_2}{(1-\rho_2)^2}G_2(\lambda)}{\left(\frac{2-\rho_2}{1-\rho_2}G_2(\lambda) + \frac{\xi_2}{\tau_2}\right)^2} < 0\end{aligned}$$

if and only if

$$G_2'(\lambda)\left(\frac{\xi_2}{\tau_2} - \frac{(2-\rho_2)\xi_1}{(1-\rho_2)\tau_1}\right) - \left(G_2(\lambda) + \frac{\xi_1}{\tau_1}\right)\frac{p_2\tau_2}{(1-\rho_2)^2}G_2(\lambda) < 0. \tag{A.8}$$

Note for $i = 1, 2$,

$$G_i'(\lambda) = \frac{\bar{\xi}}{(1-\lambda\bar{\tau})^2} + \frac{p_i\xi_i}{(1-\lambda p_i\tau_i)^2} > 0.$$

Then, (A.8) is equivalent to

$$\frac{\xi_2}{\tau_2} - \frac{(2-\rho_2)\xi_1}{(1-\rho_2)\tau_1} < \frac{\frac{p_2\tau_2}{(1-\rho_2)^2}\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_2\xi_2}{1-\rho_2}\right)\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_2\xi_2}{1-\rho_2} + \frac{\xi_1}{\tau_1}\right)}{\frac{\bar{\xi}}{(1-\rho)^2} + \frac{p_2\xi_2}{(1-\rho_2)^2}}.$$

Similarly,

$$\begin{aligned}\frac{\partial B}{\partial \lambda} &= \frac{\left(\frac{p_1\tau_1}{(1-\rho_1)^2}G_1(\lambda) + \frac{2-\rho_1}{1-\rho_1}G_1'(\lambda)\right)\left(G_1(\lambda) + \frac{\xi_2}{\tau_2}\right) - G_1'(\lambda)\left(\frac{2-\rho_1}{1-\rho_1}G_1(\lambda) + \frac{\xi_1}{\tau_1}\right)}{\left(G_1(\lambda) + \frac{\xi_2}{\tau_2}\right)^2} \\ &= \frac{\frac{p_1\tau_1}{(1-\rho_1)^2}G_1(\lambda)\left(G_1(\lambda) + \frac{\xi_2}{\tau_2}\right) + G_1'(\lambda)\left(\frac{(2-\rho_1)\xi_2}{(1-\rho_1)\tau_2} - \frac{\xi_1}{\tau_1}\right)}{\left(G_1(\lambda) + \frac{\xi_2}{\tau_2}\right)^2} > 0\end{aligned}$$

if and only if

$$\frac{\xi_1}{\tau_1} - \frac{(2-\rho_1)\xi_2}{(1-\rho_1)\tau_2} < \frac{\frac{p_1\tau_1}{(1-\rho_1)^2}\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_1\xi_1}{1-\rho_1}\right)\left(\frac{2\bar{\zeta}}{3\bar{\xi}} + \frac{\lambda\bar{\xi}}{1-\rho} + \frac{\lambda p_1\xi_1}{1-\rho_1} + \frac{\xi_2}{\tau_2}\right)}{\frac{\bar{\xi}}{(1-\rho)^2} + \frac{p_1\xi_1}{(1-\rho_1)^2}}.$$

□

Proof of Proposition 2.3. When service times are i.i.d., for notational convenience we drop the subscript from all parameters related to the service time distribution, i.e., τ_i, ξ_i and ζ_i. Then,

$$A = \frac{\frac{2\zeta}{3\xi} + \frac{\lambda\xi}{1-\rho} + \frac{\xi}{\tau(1-\rho_2)}}{\frac{2-\rho_2}{1-\rho_2}\left(\frac{2\zeta}{3\xi} + \frac{\lambda\xi}{1-\rho}\right) + \frac{\xi}{\tau(1-\rho_2)^2}},\ B = \frac{\frac{2-\rho_1}{1-\rho_1}\left(\frac{2\zeta}{3\xi} + \frac{\lambda\xi}{1-\rho}\right) + \frac{\xi}{\tau(1-\rho_1)^2}}{\frac{2\zeta}{3\xi} + \frac{\lambda\xi}{1-\rho} + \frac{\xi}{\tau(1-\rho_1)}}.$$

Let $M \equiv \frac{2\zeta}{3\xi} + \frac{\lambda\xi}{1-\rho}$, which is positive and not changing with respect to p_i for $i = 1, 2$, then we have

$$\frac{\partial A}{\partial p_2} = \frac{-\tau M^2 - \frac{\rho_2\xi}{(1-\rho_2)}M}{\frac{(1-\rho_2)^2}{\lambda}\left[\frac{2-\rho_2}{1-\rho_2}M + \frac{\xi}{\tau(1-\rho_2)^2}\right]^2} < 0.$$

Similarly, computing the partial derivative of B with respect to p_1, we have

$$\frac{\partial B}{\partial p_1} = \frac{\tau M^2 + \frac{\rho_1\xi}{(1-\rho_1)}M}{\frac{(1-\rho_1)^2}{\lambda}\left[M + \frac{\xi}{\tau(1-\rho_1)}\right]^2} > 0.$$

□

Proof of Proposition 2.4. When service times are exponential, $\zeta_i = 6\tau_i^3$ and $\xi_i = 2\tau_i^2$ for $i = 1, 2$, and hence we have,

$$A_{exp} = \frac{\frac{p_1\tau_1^3 + p_2\tau_2^3}{p_1\tau_1^2 + p_2\tau_2^2} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_2\tau_2^2}{1-\rho_2} + \tau_1}{\frac{2-\rho_2}{1-\rho_2}\left(\frac{p_1\tau_1^3 + p_2\tau_2^3}{p_1\tau_1^2 + p_2\tau_2^2} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_2\tau_2^2}{1-\rho_2}\right) + \tau_2},$$

$$B_{exp} = \frac{\frac{2-\rho_1}{1-\rho_1}\left(\frac{p_1\tau_1^3 + p_2\tau_2^3}{p_1\tau_1^2 + p_2\tau_2^2} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_1\tau_1^2}{1-\rho_1}\right) + \tau_1}{\frac{p_1\tau_1^3 + p_2\tau_2^3}{p_1\tau_1^2 + p_2\tau_2^2} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_1\tau_1^2}{1-\rho_1} + \tau_2}.$$

When service times are deterministic, $\zeta_i = \tau_i^3$ and $\xi_i = \tau_i^2$ for $i = 1, 2$, and then we have

$$A_{det} = \frac{\frac{2\left(p_1\tau_1^3 + p_2\tau_2^3\right)}{3\left(p_1\tau_1^2 + p_2\tau_2^2\right)} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_2\tau_2^2}{1-\rho_2} + \tau_1}{\frac{2-\rho_2}{1-\rho_2}\left(\frac{2\left(p_1\tau_1^3 + p_2\tau_2^3\right)}{3\left(p_1\tau_1^2 + p_2\tau_2^2\right)} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_2\tau_2^2}{1-\rho_2}\right) + \tau_2},$$

$$B_{det} = \frac{\frac{2-\rho_1}{1-\rho_1}\left(\frac{2\left(p_1\tau_1^3 + p_2\tau_2^3\right)}{3\left(p_1\tau_1^2 + p_2\tau_2^2\right)} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_1\tau_1^2}{1-\rho_1}\right) + \tau_1}{\frac{2\left(p_1\tau_1^3 + p_2\tau_2^3\right)}{3\left(p_1\tau_1^2 + p_2\tau_2^2\right)} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_1\tau_1^2}{1-\rho_1} + \tau_2}.$$

(a) For notational simplicity, for $i = 1, 2$, we let

$$M_{exp}^{(i)} \equiv \frac{p_1\tau_1^3 + p_2\tau_2^3}{p_1\tau_1^2 + p_2\tau_2^2} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_i\tau_i^2}{1-\rho_i},$$

and

$$M_{det}^{(i)} \equiv \frac{2\left(p_1\tau_1^3 + p_2\tau_2^3\right)}{3\left(p_1\tau_1^2 + p_2\tau_2^2\right)} + \frac{\lambda\left(p_1\tau_1^2 + p_2\tau_2^2\right)}{1-\rho} + \frac{\lambda p_i\tau_i^2}{1-\rho_i},$$

where $M_{exp}^{(i)} > M_{det}^{(i)}$. Taking the difference of A_{exp} and A_{det}, we have

$$A_{exp} - A_{det} = \frac{\left(\tau_2 - \left(\frac{2-\rho_2}{1-\rho_2}\right)\tau_1\right)\left(M_{exp}^{(2)} - M_{det}^{(2)}\right)}{\left(\left(\frac{2-\rho_2}{1-\rho_2}\right)M_{exp}^{(2)} + \tau_2\right)\left(\left(\frac{2-\rho_2}{1-\rho_2}\right)M_{det}^{(2)} + \tau_2\right)}.$$

Hence, $A_{exp} \le A_{det}$ if and only if $\tau_2 \le \frac{2-\rho_2}{1-\rho_2}\tau_1$.

(b) Taking the difference of B_{exp} and B_{det}, we have

$$B_{exp} - B_{det} = \frac{\left(\left(\frac{2-\rho_1}{1-\rho_1}\right)\tau_2 - \tau_1\right)\left(M_{exp}^{(1)} - M_{det}^{(1)}\right)}{\left(M_{det}^{(1)} + \tau_2\right)\left(M_{exp}^{(1)} + \tau_2\right)}.$$

Hence, $B_{exp} \ge B_{det}$ if and only if $\tau_1 \le \frac{2-\rho_1}{1-\rho_1}\tau_2$.

□

Proof of Proposition 2.5. From Corollary 2.3 and Equations (A.6) and (A.7), α and β are computed as:

$$\begin{aligned}
\alpha &= \left(\frac{\tau_1}{\tau_2}\right)\frac{k_2\left[\frac{2\bar{\zeta}}{3\xi} + \frac{\lambda\bar{\xi}(2-\rho-\rho_2)}{(1-\rho)(1-\rho_2)} + \frac{\lambda p_1\xi_1(1-\rho)}{\rho_1(1-\rho_2)}\right] + h_2}{k_2\left[\frac{2\bar{\zeta}(2-\rho_2)}{3\bar{\xi}(1-\rho_2)} + \frac{\lambda\bar{\xi}(2-\rho_2)}{(1-\rho)(1-\rho_2)} + \frac{\lambda p_2\xi_2}{\rho_2(1-\rho_2)^2}\right] + h_2}, \\
\beta &= \left(\frac{\tau_1}{\tau_2}\right)\frac{k_2\left[\frac{2\bar{\zeta}(2-\rho_1)}{3\bar{\xi}(1-\rho_1)} + \frac{\lambda\bar{\xi}(2-\rho_1)}{(1-\rho)(1-\rho_1)} + \frac{\lambda p_1\xi_1}{\rho_1(1-\rho_1)^2}\right] + h_2}{k_2\left[\frac{2\bar{\zeta}}{3\xi} + \frac{\lambda\bar{\xi}(2-\rho-\rho_1)}{(1-\rho)(1-\rho_1)} + \frac{\lambda p_2\xi_2(1-\rho)}{\rho_2(1-\rho_1)}\right] + h_2}.
\end{aligned}$$

Let $A_d[B_d]$ and $A_n[B_n]$ denote the denominator and numerator of $A[B]$, as given in Proposition 2.1, respectively. Then,

$$\alpha = \left(\frac{\tau_1}{\tau_2}\right)\left(\frac{k_2A_n + h_2}{k_2A_d + h_2}\right),\ \beta = \left(\frac{\tau_1}{\tau_2}\right)\left(\frac{k_2B_n + h_2}{k_2B_d + h_2}\right).$$

(a) We have $A < B$ and hence $A_nB_d - B_nA_d < 0$. Besides,

$$B_d + A_n - A_d - B_n = -\frac{G_1(\lambda)}{1-\rho_1} - \frac{G_2(\lambda)}{1-\rho_2} < 0.$$

Hence,

$$\alpha - \beta = \left(\frac{\tau_1}{\tau_2}\right)\left(\frac{k_2^2(A_nB_d - B_nA_d) + h_2k_2(B_d + A_n - A_d - B_n)}{(k_2A_d + h_2)(k_2B_d + h_2)}\right) < 0.$$

(b) From Proposition 2.1, if (2.12) holds, then

$$\frac{\partial A}{\partial \lambda} = \frac{A_n'A_d - A_d'A_n}{A_d^2} < 0,$$

where A_d' and A_n' denote the partial derivatives of A_d and A_n with respect to λ, respectively. Then, $A_n'A_d - A_d'A_n < 0$. Besides, the difference $A_d - A_n = \frac{G_2(\lambda)}{1-\rho_2} + \frac{\xi_2}{\tau_2} - \frac{\xi_1}{\tau_1}$ increases in λ since $G_2(\lambda)$ increases and $1 - \rho_2$ decreases in λ (see the proof of Proposition 2.1). Then, $A_d' > A_n'$.

$$\frac{\partial \alpha}{\partial \lambda} = \left(\frac{k_2\tau_1}{\tau_2}\right)\left(\frac{k_2(A_n'A_d - A_d'A_n) + h_2(A_n' - A_d')}{(k_2A_d + h_2)^2}\right) < 0.$$

Thus, α decreases as λ increases.

Similarly, if (2.13) holds, then $B_n'B_d - B_d'B_n > 0$ from Proposition 2.1, where B_d' and B_n' denote the partial derivatives of B_d and B_n with respect to λ, respectively. Besides, the difference $B_n - B_d = \frac{G_1(\lambda)}{1-\rho_1} + \frac{\xi_1}{\tau_1} - \frac{\xi_2}{\tau_2}$ increases in λ since $G_1(\lambda)$ increases and $1-\rho_1$ decreases in λ (see the proof of Proposition 2.1). Then, $B_n' > B_d'$.

$$\frac{\partial \beta}{\partial \lambda} = \left(\frac{k_2\tau_1}{\tau_2}\right)\left(\frac{k_2(B_n'B_d - B_d'B_n) + h_2(B_n' - B_d')}{(k_2B_d + h_2)^2}\right) > 0.$$

Thus, β increases as λ increases.

When $h_2 = 0$, $\alpha = A\tau_1/\tau_2$ and $\beta = B\tau_1/\tau_2$, and hence from Proposition 2.2, (2.12) and (2.13) become necessary in this case.

As $\lambda \to 1/\bar{\tau}$, then, $\rho \to 1$, and we have,

$$\lim_{\lambda \to 1/\bar{\tau}} \alpha = \frac{\tau_1}{\tau_2} \cdot \frac{1-\rho_2}{2-\rho_2}, \quad \lim_{\lambda \to 1/\bar{\tau}} \beta = \frac{\tau_1}{\tau_2} \cdot \frac{2-\rho_1}{1-\rho_1}.$$

(c) Let $\bar{A}_n[\bar{B}_n]$ and $\bar{A}_d[\bar{B}_d]$ denote the numerator and denominator of $A[B]$, as given in the proof of Proposition 2.3, then

$$\alpha = \frac{k_2\bar{A}_n + h_2}{k_2\bar{A}_d + h_2}, \ \beta = \frac{k_2\bar{B}_n + h_2}{k_2\bar{B}_d + h_2}.$$

From Proposition 2.3, A and B both increase in p_1, then $\bar{A}_n'\bar{A}_d - \bar{A}_d'\bar{A}_n > 0$ and $\bar{B}_n'\bar{B}_d - \bar{B}_d'\bar{B}_n > 0$, where $A_n',\ A_d',\ B_n',\ B_d'$ denote the partial derivatives of each quantity with respect to p_1.

Besides, the difference $\bar{A}_d - \bar{A}_n = \frac{1}{1-\rho_2}\left(\frac{2\zeta}{3\xi} + \frac{\lambda\xi}{1-\rho} + \frac{\lambda p_2 \xi}{1-\rho_2}\right)$ increases in p_2, and hence decreases in p_1, which implies that $\bar{A}'_d < \bar{A}'_n$. Then, we have

$$\frac{\partial \alpha}{\partial p_1} = \frac{k_2^2(\bar{A}'_n\bar{A}_d - \bar{A}'_d\bar{A}_n) + k_2h_2(\bar{A}'_n - \bar{A}'_d)}{(k_2\bar{A}_d + h_2)^2} > 0.$$

Similarly, the difference $\bar{B}_n - \bar{B}_d = \frac{1}{1-\rho_1}\left(\frac{2\zeta}{3\xi} + \frac{\lambda\xi}{1-\rho} + \frac{\lambda p_1 \xi}{1-\rho_1}\right)$ increases in p_1, and hence $\bar{B}'_n > \bar{B}'_d$. Then, we have $\frac{\partial \beta}{\partial p_1} > 0$.

If $\lambda \to 1/\bar{\tau}$ and $p_1 \to 0$, then $\rho, \rho_2 \to 1$ and $\rho_1 \to 0$, and we have

$$\lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 0}} \alpha = 0, \quad \text{and} \quad \lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 0}} \beta = \frac{2\tau_1}{\tau_2}.$$

If $\lambda \to 1/\bar{\tau}$ and $p_1 \to 0$, then $\rho, \rho_1 \to 1$ and $\rho_2 \to 0$, and we have

$$\lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 1}} \alpha = \frac{\tau_1}{2\tau_2}, \text{ and } \lim_{\substack{\lambda\to 1/\bar{\tau}\\ p_1\to 0}} \beta = \infty.$$

□

Proof of Example 2.2(ii). Let $f(t) = C'_1(t) - \beta C'_2(t) = h_1\alpha_1 e^{\alpha_1 t} - \beta(2k_2t + h_2)$ for $t \geq 0$. We need to find conditions so that $f(t) \geq 0$ for all $t \geq 0$. First we find that $f(t)$ is convex in t for $t \geq 0$ by second derivative test, so there is a global minimum for $t \geq 0$. We solve for $f'(t) = 0$ and we have a stationary point $t^* = \frac{1}{\alpha_1}\ln\left(\frac{2\beta k_2}{h_1\alpha_1^2}\right)$. We have the following two cases:

Case 1: If $t^* \leq 0$, i.e., $h_1 \geq \frac{2\beta k_2}{\alpha_1^2}$, then the minimum happens at $t = 0$, and hence we need $f(0) = h_1\alpha_1 - \beta h_2 \geq 0$, which is equivalent to $h_1 \geq \frac{h_2\beta}{\alpha_1}$. Thus, in this case,

$$h_1 \geq \max\left\{\frac{h_2\beta}{\alpha_1}, \frac{2k_2\beta}{\alpha_1^2}\right\}. \tag{A.9}$$

Case 2: If $t^* > 0$, i.e., $h_1 < \frac{2\beta k_2}{\alpha_1^2}$, the minimum happens at t^* and hence we need $f(t^*) = \frac{2\beta k_2}{\alpha_1} - \frac{2\beta k_2}{\alpha_1}\ln\left(\frac{2\beta k_2}{h_1\alpha_1^2}\right) - \beta h_2 \geq 0$, which is equivalent to

$$\frac{2\beta k_2}{\alpha_1}\left(1 - \ln\left(\frac{2\beta k_2}{h_1\alpha_1^2}\right)\right) \geq \beta h_2 \Leftrightarrow 1 - \ln\left(\frac{2\beta k_2}{h_1\alpha_1^2}\right) \geq \frac{h_2\alpha_1}{2k_2}$$
$$\Leftrightarrow \ln\left(\frac{2\beta k_2}{h_1\alpha_1^2}\right) \leq 1 - \frac{h_2\alpha_1}{2k_2} \Leftrightarrow \frac{2\beta k_2}{h_1\alpha_1^2} \leq e^{\left(1-\frac{h_2\alpha_1}{2k_2}\right)} \Leftrightarrow h_1 \geq \frac{2\beta k_2}{\alpha_1^2}e^{\left(1-\frac{h_2\alpha_1}{2k_2}\right)}$$

Thus, in this case,

$$\frac{2k_2\beta}{\alpha_1^2}e^{\left(\frac{h_2\alpha_1}{2k_2}-1\right)} \le h_1 < \frac{2k_2\beta}{\alpha_1^2}. \tag{A.10}$$

Furthermore, we notice that if $h_2 \ge \frac{2k_2}{\alpha_1}$, then (A.9) reduces to $h_1 \ge \frac{h_2\beta}{\alpha_1}$, and (A.10) could not hold since the lower bound is larger than the upper bound; and if $h_2 \ge \frac{2k_2}{\alpha_1}$, (A.9) and (A.10) together could reduce to $h_1 \ge \frac{2k_2\beta}{\alpha_1^2}e^{\left(\frac{h_2\alpha_1}{2k_2}-1\right)}$. Hence, $f(t) \ge 0$ if (2.14) holds, which then implies PF_1 is better than F and PF_2 by Corollary 2.3(a).

□

A.5 Proof of results in Section 2.5

Derivation of Equations (2.17) ***and*** (2.18)**.** From Theorem 2.1,

From Lemma A.2, we have for $i = 1, 2$,

$$\begin{aligned}
&\widetilde{W}^F(s) - \widetilde{W}_i^{PF_i}(s) = \frac{(1-\rho)s}{s-\lambda+\lambda\tilde{S}(s)} - \frac{(1-\rho)s + \lambda p_{3-i}\left(1-\tilde{S}_{3-i}(s)\right)}{s-\lambda p_i + \lambda p_i\tilde{S}_i(s)} \\
&= \frac{(1-\rho)s\lambda p_{3-i}\left(1-\tilde{S}_{3-i}(s)\right) - \lambda p_{3-i}\left(1-\tilde{S}_{3-i}(s)\right)\left(s-\lambda+\lambda\tilde{S}(s)\right)}{\left(s-\lambda+\lambda\tilde{S}(s)\right)\left(s-\lambda p_i+\lambda p_i\tilde{S}_i(s)\right)} \\
&= \frac{\lambda p_{3-i}\left(1-\tilde{S}_{3-i}(s)\right)\left[(1-\rho)s - \left(s-\lambda+\lambda\tilde{S}(s)\right)\right]}{\left(s-\lambda+\lambda\tilde{S}(s)\right)\left(s-\lambda p_i+\lambda p_i\tilde{S}_i(s)\right)} \\
&= \frac{\lambda p_{3-i}\left(1-\tilde{S}_{3-i}(s)\right)\left[-\rho s + \lambda\left(1-\tilde{S}(s)\right)\right]}{\left(s-\lambda+\lambda\tilde{S}(s)\right)\left(s-\lambda p_i+\lambda p_i\tilde{S}_i(s)\right)}, \\
&\widetilde{W}_i^{PF_{3-i}}(s) - \widetilde{W}^F(s) = \widetilde{W}^F(s') - \widetilde{W}^F(s) = \frac{1-\rho}{1-\frac{\lambda-\lambda\tilde{S}(s')}{s'}} - \frac{1-\rho}{1-\frac{\lambda-\lambda\tilde{S}(s)}{s}} \\
&= \frac{(1-\rho)\left(\frac{\lambda-\lambda\tilde{S}(s')}{s'} - \frac{\lambda-\lambda\tilde{S}(s)}{s}\right)}{\left(1-\frac{\lambda-\lambda\tilde{S}(s')}{s'}\right)\left(1-\frac{\lambda-\lambda\tilde{S}(s)}{s}\right)} = \frac{(1-\rho)\lambda\left[s\left(1-\tilde{S}(s')\right) - s'\left(1-\tilde{S}(s)\right)\right]}{\left(s'-\lambda+\lambda\tilde{S}(s')\right)\left(s-\lambda+\lambda\tilde{S}(s)\right)},
\end{aligned}$$

where $s' = f_{3-j}(s) = \lambda p_{3-i}(1 - B_{3-i}(s)) + s$. Then,

$$\begin{aligned}
\widetilde{U}_i^{PF_i}(s) &= \frac{\widetilde{W}_i^{PF_i}(s) - \widetilde{W}^F(s)}{s\left(E\left[W^F\right] - E\left[W_i^{PF_i}\right]\right)} \\
&= \left(\frac{2(1-\rho_i)(1-\rho)}{\lambda^2\bar{\xi}p_{3-i}\tau_{3-i}s}\right)\left(-\frac{\lambda p_{3-i}\left(1-\tilde{S}_{3-i}(s)\right)\left[-\rho s + \lambda\left(1-\tilde{S}(s)\right)\right]}{\left(s-\lambda+\lambda\tilde{S}(s)\right)\left(s-\lambda p_i+\lambda p_i\tilde{S}_i(s)\right)}\right) \\
&= -\frac{2(1-\rho_i)(1-\rho)\left(1-\tilde{S}_{3-i}(s)\right)\left[-\rho s + \lambda\left(1-\tilde{S}(s)\right)\right]}{\lambda\bar{\xi}\tau_{3-i}s\left(s-\lambda+\lambda\tilde{S}(s)\right)\left(s-\lambda p_i+\lambda p_i\tilde{S}_i(s)\right)}, \\
\widetilde{U}_i^{PF_{3-i}}(s) &= \frac{\widetilde{W}^F(s) - \widetilde{W}_i^{PF_{3-i}}(s)}{s\left(E\left[W_i^{PF_{3-i}}\right] - E\left[W^F\right]\right)} \\
&= \left(\frac{2(1-\rho_{3-i})(1-\rho)}{\lambda^2\bar{\xi}p_{3-i}\tau_{3-i}s}\right)\left(-\frac{(1-\rho)\lambda\left[s\left(1-\tilde{S}(s')\right) - s'\left(1-\tilde{S}(s)\right)\right]}{\left(s'-\lambda+\lambda\tilde{S}(s')\right)\left(s-\lambda+\lambda\tilde{S}(s)\right)}\right) \\
&= -\frac{2(1-\rho_{3-i})(1-\rho)^2\left[s\left(1-\tilde{S}\left(f_{3-j}(s)\right)\right) - f_{3-j}(s)\left(1-\tilde{S}(s)\right)\right]}{\lambda\bar{\xi}p_{3-i}\tau_{3-i}s\left[f_{3-j}(s)-\lambda+\lambda\tilde{S}\left(f_{3-j}(s)\right)\right]\left[s-\lambda+\lambda\tilde{S}(s)\right]}.
\end{aligned}$$

Plugging into (2.16), we have

$$\begin{aligned}
E\left[C_i'(U_j^{PF_j})\right] &= \frac{2k_i(1-\rho_j)(1-\rho)\left(1-\tilde{S}_{3-j}(-h_i)\right)\left[\rho h_i + \lambda\left(1-\tilde{S}(-h_i)\right)\right]}{\lambda\bar{\xi}\tau_{3-j}\left(-h_i-\lambda+\lambda\tilde{S}(-h_i)\right)\left(-h_i-\lambda p_j+\lambda p_j\tilde{S}_j(-h_i)\right)}, \\
E\left[C_i'(U_j^{PF_{3-j}})\right] &= \frac{2k_i(1-\rho_{3-j})(1-\rho)^2\left[-h_i\left(1-\tilde{S}\left(f_{3-j}(-h_i)\right)\right) - f_{3-j}(-h_i)\left(1-\tilde{S}(-h_i)\right)\right]}{\lambda\bar{\xi}p_{3-j}\tau_{3-j}\left[-h_i-\lambda+\lambda\tilde{S}(-h_i)\right]\left[f_{3-j}(-h_i)-\lambda+\lambda\tilde{S}\left(f_{3-j}(-h_i)\right)\right]}.
\end{aligned}$$

□

A.6 Proof of results in Section 2.6

Proof of Proposition 2.8: Let W^π denote the steady-state waiting time under policy $\pi \in \{F, PF_1, PF_2\}$. By conditioning on the customer type, we have

$$E\left[C_2(W^\pi)\right] = p_1 E\left[C_2(W_1^\pi)\right] + p_2 E\left[C_2(W_2^\pi)\right]. \tag{A.11}$$

According to Theorem 2 in (Vasicek 1977), if $C_2(\cdot)$ is convex,

$$E[C_2(W^F)] \le E[C_2(W^{PF_i})] \text{ for } i = 1, 2. \tag{A.12}$$

From (A.11) and (A.12), we have,

$$p_i E\left[C_2(W^F)\right] + p_{3-i} E\left[C_2(W^F)\right] \leq p_i E\left[C_2(W_i^{PF_i})\right] + p_{3-i} E\left[C_2(W_{3-i}^{PF_i})\right]$$
$$\Leftrightarrow \frac{E\left[C_2(W^F)\right] - E\left[C_2(W_i^{PF_i})\right]}{p_{3-i}} \leq \frac{E\left[C_2(W_{3-i}^{PF_i})\right] - E\left[C_2(W^F)\right]}{p_i} \text{ for } i = 1, 2$$

Since $C_2(\cdot)$ is non-decreasing and $W_i^{PF_i} \leq_{st} W^F \leq_{st} W_{3-i}^{PF_i}$ from Lemma 2.3, we have

$$E\left[C_2(W^F)\right] - E\left[C_2(W_i^{PF_i})\right] \geq 0,\ E\left[C_2(W_{3-i}^{PF_i})\right] - E\left[C_2(W^F)\right] \geq 0,\ \text{for } i = 1, 2.$$

Then,

$$\left(\frac{p_i}{p_{3-i}}\right)\left(\frac{E\left[C_2(W^F)\right] - E\left[C_2(W_i^{PF_i})\right]}{E\left[C_2(W_{3-i}^{PF_i})\right] - E\left[C_2(W^F)\right]}\right) \leq 1 \text{ for } i = 1, 2, \tag{A.13}$$

Hence,

$$\alpha = \left(\frac{\tau_1}{\tau_2}\right)\left(\frac{E[C_2'(U_2^{PF_2})]}{E[C_2'(U_1^{PF_2})]}\right) = \left(\frac{p_2}{p_1}\right)\left(\frac{E[C_2(W^F)] - E[C_2(W_2^{PF_2})]}{E[C_2(W_1^{PF_2})] - E[C_2(W^F)]}\right) \leq 1,$$

and

$$\beta = \left(\frac{\tau_1}{\tau_2}\right)\left(\frac{E[C_2'(U_2^{PF_1})]}{E[C_2'(U_1^{PF_1})]}\right) = \left(\frac{p_2}{p_1}\right)\left(\frac{E[C_2(W_2^{PF_1})] - E[C_2(W^F)]}{E[C_2(W^F)] - E[C_2(W_1^{PF_1})]}\right) \geq 1.$$

□

A.7 Proof of results in Section 2.7

Proof of Proposition 2.9: Let W^π denote the steady-state waiting time under policy $\pi \in \{L, PL_1, PL_2\}$. Then, (A.11) still holds for policy $\pi \in \{L, PL_1, PL_2\}$.

If $C_2(\cdot)$ is concave, then $-C_2$ is convex, then according to Theorem 2 in (Vasicek 1977), for $i = 1, 2$,

$$E[-C_2(W^L)] \leq E[-C_2(W^{PL_i})] \Leftrightarrow E[C_2(W^L)] \geq E[C_2(W^{PL_i})] \tag{A.14}$$

Lrom (A.11) and (A.14), we have,

$$\begin{aligned} & p_i E\left[C_2(W^L)\right] + p_{3-i} E\left[C_2(W^L)\right] \geq p_i E\left[C_2(W_i^{PL_i})\right] + p_{3-i} E\left[C_2(W_{3-i}^{PL_i})\right] \\ \Leftrightarrow & \frac{E\left[C_2(W^L)\right] - E\left[C_2(W_i^{PL_i})\right]}{p_{3-i}} \geq \frac{E\left[C_2(W_{3-i}^{PL_i})\right] - E\left[C_2(W^L)\right]}{p_i} \text{ for } i = 1, 2 \end{aligned} \tag{A.15}$$

Since $C_2(\cdot)$ is non-decreasing and $W_i^{PL_i} \leq_{st} W^L \leq_{st} W_{3-i}^{PL_i}$ from Lemma 2.3, we have

$$E\left[C_2(W^L)\right] - E\left[C_2(W_i^{PL_i})\right] \geq 0,\ E\left[C_2(W_{3-i}^{PL_i})\right] - E\left[C_2(W^L)\right] \geq 0,\ \text{for } i = 1, 2.$$

Then,

$$\left(\frac{p_i}{p_{3-i}}\right)\left(\frac{E\left[C_2(W^L)\right] - E\left[C_2(W_i^{PL_i})\right]}{E\left[C_2(W_{3-i}^{PL_i})\right] - E\left[C_2(W^L)\right]}\right) \geq 1 \text{ for } i = 1, 2,$$

Hence,

$$\alpha^L = \left(\frac{\tau_1}{\tau_2}\right)\left(\frac{E[C_2'(U_2^{PL_2})]}{E[C_2'(U_1^{PL_2})]}\right) = \left(\frac{p_2}{p_1}\right)\left(\frac{E[C_2(W^L)] - E[C_2(W_2^{PL_2})]}{E[C_2(W_1^{PL_2})] - E[C_2(W^L)]}\right) \geq 1,$$

and

$$\beta^L = \left(\frac{\tau_1}{\tau_2}\right)\left(\frac{E[C_2'(U_2^{PL_1})]}{E[C_2'(U_1^{PL_1})]}\right) = \left(\frac{p_2}{p_1}\right)\left(\frac{E[C_2(W_2^{PL_1})] - E[C_2(W^L)]}{E[C_2(W^L)] - E[C_2(W_1^{PL_1})]}\right) \leq 1.$$

□

Proof of Lemma 2.7: We consider policy PL_i for fixed $i \in \{1, 2\}$, under which we refer type i customers as "priority customers" and type $3-i$ customers as "non-priority customers". We first establish the expression for $\widetilde{W}_i^{PL_i}$.

The waiting time of a priority customer consists of two components, the remaining service time of the current customer R, and time required to serve all subsequent priority arrivals which enters service before the priority customer we considered.

Let F_R denote the cdf of R, and $\tilde{R}(s)$ denote the LST of F_R, then((Wishart 1960)),

$$\tilde{R}(s) = 1 - \rho + \frac{\lambda}{s}\big(1 - \tilde{S}(s)\big).$$

Let N_R denote the number of priority customers arriving during $R(N_R = O$ when $R = O)$. Let $T_k, k = 1, \ldots, N_R$ denote the time until the number of priority customers in the queue decreased from k to

$k-1$, starting with a priority customer enters service. Then, $T_k, k = 1, 2, \ldots, N_R$ are i.i.d., and T_1 is defined the same as the busy period of the M/G/1 queue with only type i arrivals, which has LST $B_i(s)$.

We have $W_i^{PL_i} = R + \sum_{k=1}^{N_R} T_k$, then,

$$\begin{aligned}
\widetilde{W}_i^{PL_i} =& E\left[e^{-sR-\sum_{k=1}^{N_R} sT_k}\right] \\
=& \int_0^\infty \sum_{n=0}^{\infty} E\left[e^{-sr-\sum_{k=1}^{n} sT_k}\Big| N_R = n, R = r\right] P\big[N_R = n\big|R = r\big] dF_R(t) \\
=& \int_0^\infty \sum_{n=0}^{\infty} e^{-sr} B_i(s)^n e^{-\lambda p_i r} \frac{(\lambda p_i r)^n}{n!} dF_R(t) = \int_0^\infty e^{-sr-\lambda p_i r(1-B_i(s))} dF_R(t) \\
=& E\left[e^{\left(-s-\lambda p_i(1-B_i(s))\right)R}\right] = \tilde{R}\big(s + \lambda p_i - \lambda p_i B_i(s)\big)
\end{aligned}$$

Plugging the expression of $\tilde{R}$, we have,

$$\widetilde{W}_i^{PL_i} = 1 - \rho + \frac{\lambda\big[1 - \tilde{S}(s + \lambda p_i - \lambda p_i B_i(s))\big]}{s + \lambda p_i - \lambda p_i B_i(s)}$$

Next, we establish $\widetilde{W}_{3-i}^{PL_i}$. The waiting time of a non-priority customer consists of two parts, denoted by W^* and W^{**}, where W^* is the time required to serve all priority customers already in the system at the arrival of the non-priority item and W^{**} is the time required to serve all subsequent customers who arrive after the non-priority customer but enters service before due to the queueing policy. The service order of customers served in W^* and W^{**} will not affect the waiting time of the non-priority customer. We notice that $W^* = W_i^{PF_i}$ since both waiting times are the time required to serve all priority customers (with different service order) in the system observed by a Poisson arrival. Then, the LST of W^* is given by((Miller 1960))

$$\widetilde{W^*} = \widetilde{W}_i^{PF_i} = \frac{(1-\rho) + \lambda p_{3-i}\left[1 - \tilde{S}_{3-i}(s)\right]}{s - \lambda p_i\left[1 - \tilde{S}_i(s)\right]}.$$

Let N denote the number of customers arrive during W^*, and T_k' denote the time until the number of customers decreased from k to $k-1$ for $k = 1, \ldots, N$ starting from a customer just entering the service. Then, $W^{**} = \sum_{k=1}^{N} T_k'$. By definition, we have $T_k',\ k = 1, \ldots, N$ are independent and have the same distribution as T_1', which is the busy period of the M/G/1 queue with LST denoted by $B(s)$.

Then,

$$\begin{aligned}
E\left[e^{-sW^{**}}\Big|W^* = w\right] = \sum_{n=0}^{\infty} E\left[e^{-\sum_{k=1}^{n} sT_k'}\right] e^{-\lambda w} \frac{(\lambda w)^n}{n!} \\
= \sum_{n=0}^{\infty} [B(s)]^n e^{-\lambda w} \frac{(\lambda w)^n}{n!} = e^{-\lambda w(1-B(s))}.
\end{aligned}$$

Hence,

$$\widetilde{W}_{3-i}^{PL_i} = E\left[e^{-s(W^*+W^{**})}\right] = \int_0^\infty E\left[e^{sW^*}e^{-sW^{**}}\middle|W^*=w\right]dF_{W^*}(w)$$
$$= \int_0^\infty e^{sw}e^{-\lambda w(1-B(s))}dF_{W^*}(w) = \int_0^\infty e^{-w[s+\lambda(1-B(s))]}dF_{W^*}(r) = \widetilde{W}^*\left(s+\lambda-\lambda B(s)\right),$$

where $F_{W^*}(\cdot)$ is the cdf of W^*. Plugging the expression of $\widetilde{W}^*$, we have

$$\widetilde{W}_{3-i}^{PL_i} = \frac{(1-\rho)(s+\lambda-\lambda B(s)) + \lambda p_{3-i}\left[1-\tilde{S}_{3-i}(s+\lambda-\lambda B(s))\right]}{s+\lambda-\lambda B(s)-\lambda p_i\left[1-\tilde{S}_i(s+\lambda-\lambda B(s))\right]}.$$

□

APPENDIX B: PROOFS OF RESULTS IN CHAPTER 3

B.1 Proofs of the Analytical Results in Section 3.4

B.1.1 Proof of Proposition 3.1

We first compute the cost functions under policy $\bar{\pi}_1$ and $\bar{\pi}_2$. When $b = 1$, the state space for the MDP is $\mathcal{S} = \{(0,0),(0,1),(1,0),(1,1),(2,0),(0,2)\}$.

Under both policy $\bar{\pi}_1$ and $\bar{\pi}_2$, we do not discharge the patient if there is only one patient who needs ICU care, i.e., $a(x_1,x_2) = (0,0)$ for $(x_1,x_2) \in \{(0,0),(0,1),(1,0)\}$. We discharge either one of the two patients when there are two patients of the same stage, i.e., $a(2,0) = (1,0)$ and $a(0,2) = (0,1)$. The only difference between these two policies is that in state $(1,1)$, we discharge a stage 1 patient under policy $\bar{\pi}_1$ and we discharge a stage 2 patient under policy $\bar{\pi}_2$.

Under either $\bar{\pi}_1$ or $\bar{\pi}_2$, the process can be modeled as a DTMC. Let X_n be the system state at the end of nth decision epoch. X_n takes values in $S_D = \{0,1,2\}$, where state 0 means the ICU is empty and state $i \in \{1,2\}$ means the ICU is occupied by a stage i patient.

Let P^1_{ij} denote the probability that the DTMC jumps from state i to j under policy $\bar{\pi}_1$ for $i,j \in \{0,1,2\}$. Then,

$$P^1_{00} = \bar{\lambda} + \lambda_1 q_1 + \lambda_2 p_2,\ P^1_{01} = \lambda_1 r_1 + \lambda_2 q_2,\ P^1_{02} = \lambda_1 p_1 + \lambda_2 r_2,$$
$$P^1_{10} = (\bar{\lambda} + \lambda_1) q_1 + \lambda_2 p_2,\ P^1_{11} = (\bar{\lambda} + \lambda_1) r_1 + \lambda_2 q_2,\ P^1_{12} = (\bar{\lambda} + \lambda_1) p_1 + \lambda_2 r_2,$$
$$P^1_{20} = p_2,\ P^1_{21} = q_2,\ P^1_{22} = r_2.$$

Let u_i denote the long-run average probability that the DTMC is in state i for $i = 0,1,2$, then u_i can be obtained by solving the balance equations and normalization equation given by

$$\begin{aligned}
(\lambda_1(1-q_1) + \lambda_2(1-p_2))u_0 &= ((\bar{\lambda} + \lambda_1)q_1 + \lambda_2 p_2)u_1 + p_2 u_2, \\
(q_2 + p_2)u_2 &= (\lambda_1 p_1 + \lambda_2 r_2)u_0 + ((\bar{\lambda} + \lambda_1)p_1 + \lambda_2 r_2)u_1, \\
u_0 + u_1 + u_2 &= 1.
\end{aligned}$$

Solving above equations, we get:

$$\begin{aligned} u_0 &= \frac{\bar{\lambda}a + \lambda_1 a + \lambda_2 p_2}{D^{\bar{\pi}_1}}, \\ u_1 &= \frac{\lambda_1(p_2 + q_2 - a) + \lambda_2 q_2}{D^{\bar{\pi}_1}}, \\ u_2 &= \frac{\bar{\lambda}(\lambda_1 p_1 + \lambda_2(p_1 + q_1 - a)) + (\lambda_1 + \lambda_2)(\lambda_1 p_1 + \lambda_2 r_2)}{D^{\bar{\pi}_1}}. \end{aligned}$$

where $a = q_1 q_2 + q_1 p_2 + p_1 p_2$ and

$$\begin{aligned} D^{\bar{\pi}_1} &= \bar{\lambda}\Big[a\bar{\lambda} + \lambda_1(p_1 + p_2 + q_2 + a) + \lambda_2(p_1 + q_1 + p_2 + q_2)\Big] \\ &\quad + (\lambda_1 + \lambda_2)\Big[\lambda_1(p_1 + p_2 + q_2) + \lambda_2\Big]. \end{aligned}$$

Let m_i denote the expected cost that incurs in state i for $i = 0, 1, 2$. Then,

$$m_0 = \lambda_1 q_1,\ m_1 = \bar{\lambda} q_1 + \lambda_1(q_1 + \phi_1^G) + \lambda_2 \phi_1^G,\ m_2 = \lambda_1 \phi_1^G + \lambda_2 \phi_2^G.$$

The long-run average cost under this policy is $J^{\bar{\pi}_1} = u_0 m_0 + u_1 m_1 + u_2 m_2$. We can then show that

$$\begin{aligned} J^{\bar{\pi}_1} D^{\bar{\pi}_1} &= (\bar{\lambda}a + \lambda_1 a + \lambda_2 p_2)\lambda_1 q_1 + \Big[\lambda_1(p_2 + q_2 - a) + \lambda_2 q_2\Big]\Big[\bar{\lambda} q_1 + \lambda_1(q_1 + \phi_1^G) + \lambda_2 \phi_1^G\Big] \\ &\quad + \Big[\bar{\lambda}(\lambda_1 p_1 + \lambda_2(p_1 + q_1 - a)) + (\lambda_1 + \lambda_2)(\lambda_1 p_1 + \lambda_2 r_2)\Big]\Big[\lambda_1 \phi_1^G + \lambda_2 \phi_2^G\Big]. \end{aligned} \tag{B.1}$$

Similarly, let P_{ij}^2 denote the probability that the DTMC jumps from state i to j under policy $\bar{\pi}_2$ for $i, j \in \{0, 1, 2\}$. Then,

$$P_{00}^2 = \bar{\lambda} + \lambda_1 q_1 + \lambda_2 p_2,\ P_{01}^2 = \lambda_1 r_1 + \lambda_2 q_2,\ P_{02}^2 = \lambda_1 p_1 + \lambda_2 r_2,$$

$$P_{10}^2 = q_1,\ P_{11}^2 = r_1,\ P_{12}^2 = p_1,$$

$$P_{20}^2 = (\bar{\lambda} + \lambda_2)p_2 + \lambda_1 q_1,\ P_{21}^2 = (\bar{\lambda} + \lambda_2)q_2 + \lambda_1 r_1,\ P_{22}^2 = (\bar{\lambda} + \lambda_2)r_2 + \lambda_1 p_1.$$

Let u_i' denote the long-run average probability that the DTMC is in state i for $i = 0, 1, 2$, then u_i' can be obtained by solving the balance equations and normalize equation given by

$$\begin{aligned} (\lambda_1(1 - q_1) + \lambda_2(1 - p_2))u_0' &= q_1 u_1' + ((\bar{\lambda} + \lambda_2)p_2 + \lambda_1 q_1)u_2', \\ (q_1 + p_1)u_1' &= (\lambda_1 r_1 + \lambda_2 q_2)u_0' + ((\bar{\lambda} + \lambda_2)q_2 + \lambda_1 r_1)u_2', \\ u_0' + u_1' + u_2' &= 1. \end{aligned}$$

Solving the above equations, we have

$$\begin{aligned} u_0' &= \frac{\bar{\lambda}a + \lambda_1 q_1 + \lambda_2 a}{D^{\bar{\pi}_2}}, \\ u_1' &= \frac{\bar{\lambda}(\lambda_1(p_2+q_2-a)+\lambda_2 q_2) + (\lambda_1+\lambda_2)(\lambda_1 r_1 + \lambda_2 q_2)}{D^{\bar{\pi}_2}}, \\ u_2' &= \frac{\lambda_1 p_1 + \lambda_2(p_1+q_1-a)}{D^{\bar{\pi}_2}}. \end{aligned}$$

where

$$\begin{aligned} D^{\bar{\pi}_2} &= \bar{\lambda}\Big[a\bar{\lambda} + (p_1+q_1+p_2+q_2)\lambda_1 + (p_1+q_1+q_2+a)\lambda_2\Big] \\ &\quad +(\lambda_1+\lambda_2)\Big[\lambda_1 + (p_1+q_1+q_2)\lambda_2\Big]. \end{aligned}$$

Let m_i' denote the expected cost that incurs in state i for $i = 0, 1, 2$. Then,

$$m_0' = \lambda_1 q_1,\ m_1' = q_1 + \lambda_1\phi_1^G + \lambda_2\phi_2^G,\ m_2' = \lambda_1(q_1+\phi_2^G) + \lambda_2\phi_2^G.$$

The long-run average cost under this policy is $J^{\bar{\pi}_2} = u_0'm_0' + u_1'm_1' + u_2'm_2'$. We can then show that

$$\begin{aligned} J^{\bar{\pi}_2}D^{\bar{\pi}_2} &= (\bar{\lambda}a + \lambda_1 q_1 + \lambda_2 a)\lambda_1 q_1 + \Big[\lambda_1 p_1 + \lambda_2(p_1+q_1-a)\Big]\Big[\lambda_1(q_1+\phi_2^G) + \lambda_2\phi_2^G\Big] \\ &\quad +\Big[\bar{\lambda}(\lambda_1(p_2+q_2-a)+\lambda_2 q_2) + (\lambda_1+\lambda_2)(\lambda_1 r_1 + \lambda_2 q_2)\Big]\Big[q_1 + \lambda_1\phi_1^G + \lambda_2\phi_2^G\Big]. \qquad \text{(B.2)} \end{aligned}$$

Next, using (B.1) and (B.2).

Take the difference of $J^{\bar{\pi}_1}$ and $J^{\bar{\pi}_2}$, after some algebra, we can show that

$$\begin{aligned} & J^{\bar{\pi}_1} - J^{\bar{\pi}_2} \\ =& \frac{M}{D^{\bar{\pi}_1}D^{\bar{\pi}_2}(1+\beta_1^G+\beta_1^G\beta_2^G)}\Big\{(1-\lambda)p_1p_2\big[\lambda\big[p_1(\beta_1^G-\beta_1) + p_2\beta_1^G(\beta_2-\beta_2^G)\big] \\ & \qquad + (\beta_1^G-\beta_1) + \beta_1\beta_1^G(\beta_2-\beta_2^G)\big]\Big\} \\ =& \frac{Mp_1p_2\beta_1\beta_1^G}{D^{\bar{\pi}_1}D^{\bar{\pi}_2}(1+\beta_1^G+\beta_1^G\beta_2^G)}\left\{\lambda\left[\frac{\beta_1^G-\beta_1}{p_2\beta_1\beta_1^G} + \frac{\beta_2-\beta_2^G}{q_1} + (1-\lambda)\left[\frac{\beta_1^G-\beta_1}{\beta_1\beta_1^G} + (\beta_2-\beta_2^G)\right]\right]\right\}, \end{aligned}$$

where $\lambda = \lambda_1 + \lambda_2$ and

$$M = \lambda_1^2 p_1 + \lambda_2^2 q_2 + \bar{\lambda}\lambda_1\lambda_2(p_1+q_1+p_2+q_2-2a) + \lambda_1\lambda_2(\lambda_1+\lambda_2)(1-a).$$

Since $a = p_1p_2 + q_1p_2 + q_1q_2 = (p_1 + q_1)(p_2 + q_2) - p_1q_2 \leq (p_1 + q_1)(p_2 + q_2)$, we have $1 - a \geq 0$ and

$$\begin{aligned} p_1 + q_1 + p_2 + q_2 - 2a &\geq p_1 + q_1 + p_2 + q_2 - 2(p_1 + q_1)(p_2 + q_2) \\ &= (p_1 + q_1)(1 - p_2 - q_2) + (p_2 + q_2)(1 - p_1 - q_1) \geq 0, \end{aligned}$$

and thus $M \geq 0$. Then, we have $J^{\bar{\pi}_1} - J^{\bar{\pi}_2} \leq 0$ if and only if

$$\lambda \left[\frac{\beta_1^G - \beta_1}{p_2\beta_1\beta_1^G} - \frac{\beta_2^G - \beta_2}{q_1} \right] + (1 - \lambda) \left[\frac{\beta_1^G - \beta_1}{\beta_1\beta_1^G} - (\beta_2^G - \beta_2) \right] \leq 0. \tag{B.3}$$

Next, we find that

$$\frac{\beta_1^G - \beta_1}{\beta_1^G\beta_1} - (\beta_2^G - \beta_2) = \frac{(\phi_1^G - \phi_2^G) - (\phi_1 - \phi_2)}{(\phi_1 - \phi_2)(\phi_1^G - \phi_2^G)}$$

and

$$\frac{\beta_1^G - \beta_1}{p_2\beta_1^G\beta_1} - \frac{\beta_2^G - \beta_2}{q_1} = \frac{L_2(\phi_1^G - \phi_1) - L_1(\phi_2^G - \phi_2)}{(\phi_1 - \phi_2)(\phi_1^G - \phi_2^G)},$$

since

$$\begin{aligned} &\frac{\beta_1^G - \beta_1}{\beta_1^G\beta_1} - (\beta_2^G - \beta_2) = \frac{1}{\beta_1} - \frac{1}{\beta_1^G} - (\beta_2^G - \beta_2) = \left(\frac{1}{\beta_1} + \beta_2 + 1 \right) - \left(\frac{1}{\beta_1^G} + \beta_2^G + 1 \right) \\ = \quad &\frac{1 + \beta_1 + \beta_1\beta_2}{\beta_1} - \frac{1 + \beta_1^G + \beta_1^G\beta_2^G}{\beta_1^G} = \frac{1}{\phi_1 - \phi_2} - \frac{1}{\phi_1^G - \phi_2^G} = \frac{(\phi_1^G - \phi_2^G) - (\phi_1 - \phi_2)}{(\phi_1 - \phi_2)(\phi_1^G - \phi_2^G)} \end{aligned}$$

and

$$
\begin{aligned}
&\frac{\beta_1^G-\beta_1}{p_2\beta_1^G\beta_1}-\frac{\beta_2^G-\beta_2}{q_1}=\frac{1}{q_1p_2}\left(\frac{q_1}{\beta_1}-\frac{q_1}{\beta_1^G}-p_2(\beta_2^G-\beta_2)\right)\\
=&\frac{1}{q_1p_2}\left(\frac{q_1}{\beta_1}+p_2\beta_2-\frac{q_1}{\beta_1^G}-p_2\beta_2^G\right)=\frac{1}{q_1p_2}\left(p_1+q_2-\frac{q_1}{\beta_1^G}-p_2\beta_2^G\right)\\
=&\frac{1}{q_1p_2}\left(p_1+q_2-\frac{q_1+p_2\beta_1^G\beta_2^G}{\beta_1^G}\right)\\
=&\frac{1+\beta_1^G+\beta_1^G\beta_2^G}{q_1p_2\beta_1^G}\left(\frac{(p_1+q_2)\beta_1^G}{1+\beta_1^G+\beta_1^G\beta_2^G}-\frac{q_1+p_2\beta_1^G\beta_2^G}{1+\beta_1^G+\beta_1^G\beta_2^G}\right)\\
=&\frac{1}{q_1p_2(\phi_1^G-\phi_2^G)}\big((p_1+q_2)(\phi_1^G-\phi_2^G)-q_1(1-\phi_1^G)-p_2\phi_2^G\big)\\
=&\frac{1}{q_1p_2(\phi_1^G-\phi_2^G)}\big((p_1+q_1+q_2)\phi_1^G-(p_1+p_2+q_2)\phi_2^G-q_1\big)\\
=&\frac{p_1p_2+q_1p_2+q_1q_2}{q_1p_2(\phi_1^G-\phi_2^G)}\left(\frac{(p_1+q_1+q_2)\phi_1^G}{p_1p_2+q_1p_2+q_1q_2}-\frac{(p_1+p_2+q_2)\phi_2^G}{p_1p_2+q_1p_2+q_1q_2}-\frac{q_1}{p_1p_2+q_1p_2+q_1q_2}\right)\\
&\big(\text{Since } q_1=(p_1+q_1+q_2)\phi_1-(p_1+p_2+q_2)\phi_2\big)\\
=&\frac{\frac{p_1+q_1+q_2}{p_1p_2+q_1p_2+q_1q_2}\phi_1^G-\frac{p_1+p_2+q_2}{p_1p_2+q_1p_2+q_1q_2}\phi_2^G-\frac{(p_1+q_1+q_2)\phi_1-(p_1+p_2+q_2)\phi_2}{p_1p_2+q_1p_2+q_1q_2}}{(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)}\\
=&\frac{L_2\phi_1^G-L_1\phi_2^G-L_2\phi_1+L_1\phi_2}{(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)}\\
=&\frac{L_2(\phi_1^G-\phi_1)-L_1(\phi_2^G-\phi_2)}{(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)}.
\end{aligned}
$$

Then, since $(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)\geq 0$, (B.3) is equivalent to (3.5).

B.1.2 Proof of Corollary 3.1

From Proposition 3.1, we have

(a) If $\phi_2^G-\phi_2\geq\phi_1^G-\phi_1$ and $\frac{\phi_2^G-\phi_2}{L_2}\geq\frac{\phi_1^G-\phi_1}{L_1}$, then (3.5) holds for any λ, and thus $J^{\bar\pi_1}\leq J^{\bar\pi_2}$.

Similarly, if $\phi_1^G-\phi_1\geq\phi_2^G-\phi_2$ and $\frac{\phi_1^G-\phi_1}{L_1}\geq\frac{\phi_2^G-\phi_2}{L_2}$, then (3.5) holds in the opposite direction for any λ, and thus, $J^{\bar\pi_2}\leq J^{\bar\pi_1}$.

(b) If $\phi_1^G-\phi_1\geq\phi_2^G-\phi_2$ and $\frac{\phi_1^G-\phi_1}{L_1}\leq\frac{\phi_2^G-\phi_2}{L_2}$, then $J^{\bar\pi_1}\geq J^{\bar\pi_2}$, i.e., (3.5) holds in the opposite direction if

$$\lambda<\frac{(\phi_1^G-\phi_1)-(\phi_2^G-\phi_2)}{(\phi_1^G-\phi_1)-(\phi_2^G-\phi_2)+\big[L_1(\phi_2^G-\phi_2)-L_2(\phi_1^G-\phi_1)\big]}.$$

Similarly, if $\phi_2^G-\phi_2\geq\phi_1^G-\phi_1$ and $\frac{\phi_2^G-\phi_2}{L_2}\leq\frac{\phi_1^G-\phi_1}{L_1}$, then $J^{\bar\pi_1}\leq J^{\bar\pi_2}$, i.e., (3.5) holds if and only if

$$\lambda\leq\frac{(\phi_2^G-\phi_2)-(\phi_1^G-\phi_1)}{(\phi_2^G-\phi_2)-(\phi_1^G-\phi_1)+\big[L_2(\phi_1^G-\phi_1)-L_1(\phi_2^G-\phi_2)\big]}.$$

B.2 Proofs of the Results in Section 3.5

We first define the operators D_1, D_2, Δ, D_{11}, D_{22}, and D_{12}. In the following, $w : \mathcal{S} \to \mathbb{R}$ is a function from the state space $\mathcal{S}$ to the set of real numbers. For some of the results, $w(\cdot)$ is restricted to be defined over a subset of the state space $\mathcal{S}$.

Definition B.1. The first difference operators D_1 and D_2 are defined as

$$D_1 w(x_1, x_2) = w(x_1 + 1, x_2) - w(x_1, x_2),\ D_2 w(x_1, x_2) = w(x_1, x_2 + 1) - w(x_1, x_2),$$
$$\text{for } x_1 + x_2 \leq b.$$

Definition B.2. The operator Δ is defined as

$$\Delta w(x_1, x_2) = w(x_1 + 1, x_2) - w(x_1, x_2 + 1) = D_1 w(x_1, x_2) - D_2 w(x_1, x_2),$$
$$\text{for } x_1 + x_2 \leq b.$$

Definition B.3. The second difference operators $D_{11} = D_1 D_1, D_{22} = D_2 D_2$ and $D_{12} = D_1 D_2 = D_2 D_1$ are defined as

$$\begin{aligned}
D_{11} w(x_1, x_2) =& D_1 w(x_1 + 1, x_2) - D_1 w(x_1, x_2) = w(x_1 + 2, x_2) - 2w(x_1 + 1, x_2) + w(x_1, x_2),\\
D_{22} w(x_1, x_2) =& D_2 w(x_1, x_2 + 1) - D_2 w(x_1, x_2) = w(x_1, x_2 + 2) - 2w(x_1, x_2 + 1) + w(x_1, x_2),\\
D_{12} w(x_1, x_2) =& D_1 w(x_1, x_2 + 1) - D_1 w(x_1, x_2) = D_2 w(x_1 + 1, x_2) - D_2 w(x_1, x_2)\\
=& w(x_1 + 1, x_2 + 1) - w(x_1, x_2 + 1) - w(x_1 + 1, x_2) + w(x_1, x_2) \text{ for } x_1 + x_2 \leq b - 1.
\end{aligned}$$

B.2.1 Proof of Proposition 3.2

Lemma B.1. *For $x_1 + x_2 \leq b$, we have $D_1 v_\alpha(x_1, x_2) \leq c_1^G,\ D_2 v_\alpha(x_1, x_2) \leq c_2^G$.*

Proof of Lemma B.1: Let $v_\alpha^\pi(x_1, x_2)$ denote the total expected discounted cost under policy π starting from state (x_1, x_2). Define a policy π_1, under which starting from state $(x_1 + 1, x_2)$ we initially discharge a stage 1 patient and use the action that is optimal for state (x_1, x_2), and then use the optimal policy thereafter. Then,

$$v_\alpha^{\pi_1}(x_1 + 1, x_2) = c_1^G + v_\alpha(x_1, x_2) \geq v_\alpha(x_1 + 1, x_2) \Rightarrow v_\alpha(x_1 + 1, x_2) - v_\alpha(x_1, x_2) \leq c_1^G.$$

Similarly, define policy π_2 as the policy under which starting from state (x_1, x_2+1) we initially discharge a stage 2 patient and use the action that is optimal for state (x_1, x_2), and then use the optimal policy thereafter. Then,

$$v_\alpha^{\pi_2}(x_1, x_2+1) = c_2^G + v_\alpha(x_1, x_2) \geq v_\alpha(x_1, x_2+1) \Rightarrow v_\alpha(x_1, x_2+1) - v_\alpha(x_1, x_2) \leq c_2^G.$$

□

Lemma B.2. *Suppose that (3.10) holds and for any $w : \mathcal{S} \to \mathbb{R}$, we have (i) $D_1 w(i,j) \leq c_1^G$ for $i+j \leq b$, and (ii) $D_2 w(i,j) \leq c_2^G$ for $i+j \leq b$. Then, for $x+y \leq b-1$, $D_1 \Gamma w(x,y) \leq \frac{c_1^G}{\alpha} - q_1$, and $D_2 \Gamma w(x,y) \leq \frac{c_2^G}{\alpha}$.*

Proof of Lemma B.2: Let $x+y \leq b-1$. Thus, $\Gamma w(x+1,y)$ and $\Gamma w(x,y+1)$ are well defined according to Definition 3.1.

We first rewrite $\Gamma w(x+1,y)$ by conditioning on how x stage 1 jobs and y stage 2 jobs evolve. Patients evolve independently, Therefore, if x stage 1 jobs and y stage 2 jobs evolve to i stage 1 patients and j stage 2 patients (which happens with probability $P(i,j|x,y)$), the extra stage 1 patient jumps to stage 0 with probability q_1, remains stage 1 with probability r_1, and jumps to stage 2 with probability p_1. Thus, we can write

$$\Gamma w(x+1,y) = \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y) \left[q_1 w(i,j) + r_1 w(i+1,j) + p_1 w(i,j+1) \right].$$

Then,

$$\begin{aligned} D_1 \Gamma w(x,y) &= \Gamma w(x+1,y) - \Gamma w(x,y), \\ &= \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y) \left[r_1 D_1 w(i,j) + p_1 D_2 w(i,j) \right] \\ &\leq r_1 c_1^G + p_1 c_2^G \leq \frac{c_1^G}{\alpha} - q_1, \end{aligned}$$

where again the first inequality follows from the lemma assumptions (i) and (ii), and the second inequality follows from Condition (3.10), which is another lemma assumption.

Similarly, we can write $\Gamma w(x,y+1)$

$$\Gamma w(x,y+1) = \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y) \left[p_2 w(i,j) + q_2 w(i+1,j) + r_2 w(i,j+1) \right],$$

and then

$$\begin{aligned} D_2\Gamma w(x,y) &= \Gamma w(x,y+1) - \Gamma w(x,y), \\ &= \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y)\left[q_2 D_1 w(i,j) + r_2 D_2 w(i,j)\right] \\ &\leq q_2 c_1^G + r_2 c_2^G \leq \frac{c_2^G}{\alpha}, \end{aligned}$$

where the first inequality follows from the lemma assumptions (i) and (ii), and the second inequality follows from Condition (3.10), which is another lemma assumption.

□

Proof of Proposition 3.2: From Lemmas B.1 and B.2, we have for any $x + y \leq b - 1$,

$$D_1\Gamma v_\alpha(x,y) \leq \frac{c_1^G}{\alpha} - q_1, \tag{B.4}$$

$$D_2\Gamma v_\alpha(x,y) \leq \frac{c_2^G}{\alpha}. \tag{B.5}$$

Then, for any $x_1 \geq a_1 \geq 1$, $x_2 \geq a_2 \geq 0$ and $x_1 - (a_1 - 1) + x_2 - a_2 \leq b$,

$$\begin{aligned} & V_\alpha(x_1, x_2, a_1 - 1, a_2) - V_\alpha(x_1, x_2, a_1, a_2) \\ =& - c_1^G + \alpha\left[q_1 + \Gamma v_\alpha(x_1 - a_1 + 1, x_2) - \Gamma v_\alpha(x_1 - a_1, x_2 - a_2)\right] \\ =& \alpha\left[-\frac{c_1^G}{\alpha} + q_1 + D_1\Gamma v(x_1 - a_1, x_2 - a_2)\right] \leq 0, \end{aligned} \tag{B.6}$$

where the inequality follows from (B.4). We can then conclude that $V_\alpha(x_1, x_2, a_1 - 1, a_2) \leq V_\alpha(x_1, x_2, a_1, a_2)$, which implies that decreasing the number of stage 1 discharges does not increase the expected cost.

Similarly, for any $x_1 \geq a_1 \geq 0$, $x_2 \geq a_2 \geq 1$ and $x_1 - a_1 + x_2 - (a_2 - 1) \leq b$,

$$\begin{aligned} & V_\alpha(x_1, x_2, a_1, a_2 - 1) - V_\alpha(x_1, x_2, a_1, a_2) \\ =& - c_2^G + \alpha\left[\Gamma v_\alpha(x_1 - a_1, x_2 - a_2 + 1) - \Gamma v_\alpha(x_1 - a_1, x_2 - a_2)\right] \\ =& \alpha\left[-\frac{c_2^G}{\alpha} + D_2\Gamma v(x_1 - a_1, x_2 - a_2)\right] \leq 0, \end{aligned} \tag{B.7}$$

where the inequality follows from (B.5). We can then conclude that $V_\alpha(x_1, x_2, a_1, a_2 - 1) \leq V_\alpha(x_1, x_2, a_1, a_2)$, which implies that decreasing the number of stage 2 discharges does not increase the expected cost, either. Then, the result follows.

□

B.2.2 Proof of Proposition 3.3

Lemma B.3. *If (3.10) holds, then $c_i \le c_i^G$ for $i = 1, 2$, where c_i^G is given in* (3.7) *and c_i is given in* (3.16).

Proof of Lemma B.3: Let $f_i(0) = c_i^G$ for $i = 1, 2$, and for $n \ge 0$ we define

$$f_1(n+1) = \alpha(q_1 + r_1 f_1(n) + p_1 f_2(n)),\ f_2(n+1) = \alpha(q_2 f_1(n) + r_2 f_2(n)). \tag{B.8}$$

Next, by induction, we show that $f_i(n+1) \le f_i(n)$ for all $n \ge 0$: From (3.10) and (B.8), we have $f_i(1) \le f_i(0)$.

Suppose $f_i(k+1) \le f_i(k)$ for $i = 1, 2$ and for some $k \ge 0$, then,

$$f_1(k+2) = \alpha(q_1 + r_1 f_1(k+1) + p_1 f_2(k+1)) \le \alpha(q_1 + r_1 f_1(k) + p_1 f_2(k)) = f_1(k+1),$$

and

$$f_2(k+2) = \alpha(q_2 f_1(k+1) + r_2 f_2(k+1)) \le \alpha(q_2 f_1(k) + r_2 f_2(k)) = f_2(k+1).$$

Hence by induction we can conclude that $f_i(n+1) \le f_i(n)$ for any $n \ge 0$. Then, $\{f_i(n), n \ge 0\}$ is a decreasing sequence and $f_i(n) \ge 0$ is bounded below, so $\lim_{n\to\infty} f_i(n)$ exists. Let f_i denote the limit for $i = 1, 2$.

Letting $n \to \infty$ in (B.8), we have

$$f_1 = \alpha(q_1 + r_1 f_1 + p_1 f_2),\ f_2 = \alpha(q_2 f_1 + r_2 f_2).$$

Solving for f_i, we find $f_i = c_i$. Since the sequence $\{f_i(n), n \ge 0\}$ is decreasing, we have $c_i^G = f_i(0) \ge \lim_{n\to\infty} f_i(n) = f_i = c_i$.

□

Definition B.4. Let $\mathcal{V}_b$ be a set of functions such that if $w \in \mathcal{V}_b$, then

$$c_1 \le D_1 w(i,j) \le c_1^G,\ c_2 \le D_2 w(i,j) \le c_2^G \text{ for all } i + j \le b. \tag{B.9}$$

Lemma B.4. *If (3.10) holds and $w \in \mathcal{V}_b$, then*

(a) for $x_1 + x_2 \le b - 1$,

$$\frac{c_1}{\alpha} - q_1 \le D_1 \Gamma w(x_1, x_2) \le \frac{c_1^G}{\alpha} - q_1,$$

$$\frac{c_2}{\alpha} \le D_2 \Gamma w(x_1, x_2) \le \frac{c_2^G}{\alpha}.$$

(b) $Tw \in \mathcal{V}_b$.

Proof of Lemma B.4: (a) If $w \in \mathcal{V}_b$, then $D_1 w(x_1, x_2) \le c_1^G$ and $D_2 w(x_1, x_2) \le c_2^G$ for $x_1 + x_2 \le b$, and from Lemma B.2, we have for $x_1 + x_2 \le b - 1$,

$$D_1 \Gamma w(x_1, x_2) \le \frac{c_1^G}{\alpha} - q_1, \; D_2 \Gamma w(x_1, x_2) \le \frac{c_2^G}{\alpha}.$$

If $w \in \mathcal{V}_b$, then for $x_1 + x_2 \le b - 1$,

$$\begin{aligned} D_1 \Gamma w(x_1, x_2) &= \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i, j | x_1, x_2) \left[r_1 D_1 w(i, j) + p_1 D_2 w(i, j) \right] \\ &\ge r_1 c_1 + p_1 c_2 = \frac{c_1}{\alpha} - q_1, \\ D_2 \Gamma w(x_1, x_2) &= \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i, j | x_1, x_2) \left[q_2 D_1 w(i, j) + r_2 D_2 w(i, j) \right] \\ &\ge q_2 c_1 + r_2 c_2 = \frac{c_2}{\alpha}, \end{aligned}$$

where the inequalities follow from the fact that $i + j \le x_1 + x_2 + 1 \le b$ and then $D_1 w(i, j) \ge c_1$ and $D_2 w(i, j) \ge c_2$ from $w \in \mathcal{V}_b$. Thus, part (a) follows.

(b) We need to consider four different cases from Definition 3.2.

Case 1: If $x_1 + x_2 \le b - 1$, we have,

$$D_1 T w(x_1, x_2) = \alpha \left[q_1 + D_1 \Gamma w(x_1, x_2) \right], \; D_2 T w(x_1, x_2) = \alpha D_2 \Gamma w(x_1, x_2).$$

Then, part (a) implies that

$$c_1 \le D_1 T w(x_1, x_2) \le c_1^G, \; c_2 \le D_2 T w(x_1, x_2) \le c_2^G \text{ for } x_1 + x_2 \le b. \tag{B.10}$$

Case 2: If $x_1 + x_2 = b$ and $x_1 > 0, x_2 > 0$, we have

$$D_1 T w(x_1, x_2) = \min \left\{ c_1^G, \; c_2^G + T w(x_1 + 1, x_2 - 1) - T w(x_1, x_2) \right\} \le c_1^G.$$

Furthermore, since $D_1 T w(x_1, x_2 - 1) \ge c_1$ and $D_2 T w(x_1, x_2 - 1) \le c_2^G$ from (B.10),

$$\begin{aligned} &c_2^G + T w(x_1 + 1, x_2 - 1) - T w(x_1, x_2) \\ &\qquad = c_2^G + D_1 T w(x_1, x_2 - 1) - D_2 T w(x_1, x_2 - 1) \ge c_1, \end{aligned}$$

and $c_1^G \ge c_1$, then we can conclude that $D_1 T w(x_1, x_2) \ge c_1$.

Similarly,

$$D_2Tw(x_1,x_2) \quad = \quad \min\left\{c_1^G + Tw(x_1-1,x_2+1) - Tw(x_1,x_2),\ c_2^G\right\} \le c_2^G$$

Furthermore, since $D_1Tw(x_1-1,x_2) \le c_1^G$ and $D_2Tw(x_1-1,x_2) \ge c_2$ from (B.10),

$$c_1^G + Tw(x_1-1,x_2+1) - Tw(x_1,x_2)$$
$$= c_1^G - D_1Tw(x_1-1,x_2) + D_2Tw(x_1-1,x_2) \ge c_2,$$

and $c_2^G \ge c_2$, we can conclude that $D_2Tw(x_1,x_2) \ge c_2$.

Case 3: If $x_1 = b$ and $x_2 = 0$, we have $D_2Tw(x_1,x_2) = c_2^G$, and thus, $c_2 \le D_2Tw(x_1,x_2) \le c_2^G$. $D_1Tw(x_1,x_2)$ has the same expression as in case 2, and hence $c_1 \le D_1Tw(x_1,x_2) \le c_1^G$.

Case 4: If $x_1 = 0$ and $x_2 = b$, we have $D_1Tw(x_1,x_2) = c_1^G$, and hence $c_1 \le D_1Tw(x_1,x_2) \le c_1^G$. $D_2Tw(x_1,x_2)$ has the same expression as in case 2, and hence $c_2 \le D_2Tw(x_1,x_2) \le c_2^G$.

The above four cases cover all the possibilities for $x_1 + x_2 \le b$. Hence, $Tw \in \mathcal{V}_b$.

□

Definition B.5. Let $\mathcal{V}$ be the set of functions such that if $w \in \mathcal{V}$, then, $w \in \mathcal{V}_b$ and w satisfies the following three conditions:

Condition 1: $D_{11}w(x_1,x_2) \ge 0,\ D_{22}w(x_1,x_2) \ge 0$ and $D_{12}w(x_1,x_2) \ge 0$ for $x_1 + x_2 \le b-1$,

Condition 2:

$$D_{11}w(x_1,x_2)D_{22}w(y_1,y_2) + D_{11}w(y_1,y_2)D_{22}w(x_1,x_2) - 2D_{12}w(x_1,x_2)D_{12}w(y_1,y_2) \ge 0,$$
$$\text{for } x_1 + x_2 \le b-1 \text{ and } y_1 + y_2 \le b-1,$$

Condition 3: $D_{11}w(x_1,x_2) + D_{22}w(x_1,x_2) - 2D_{12}w(x_1,x_2) \ge 0$ for $x_1 + x_2 \le b-1$.

Lemma B.5. *Suppose that (3.10) holds, and $w \in \mathcal{V}$. Then,*

(i) for $x_1 + x_2 \le b-2$ and $y_1 + y_2 \le b-2$,

$$D_{11}\Gamma w(x_1,x_2) \ge 0,\ D_{22}\Gamma w(x_1,x_2) \ge 0,\ \textit{and}\ D_{12}\Gamma w(x_1,x_2) \ge 0, \tag{B.11}$$

$$D_{11}\Gamma w(x_1,x_2)D_{22}\Gamma w(y_1,y_2) + D_{11}\Gamma w(y_1,y_2)D_{22}\Gamma w(x_1,x_2)$$
$$- 2D_{12}\Gamma w(x_1,x_2)D_{12}\Gamma w(y_1,y_2) \ge 0, \tag{B.12}$$

$$D_{11}\Gamma w(x_1,x_2) + D_{22}\Gamma w(x_1,x_2) - 2D_{12}\Gamma w(x_1,x_2) \ge 0. \tag{B.13}$$

(ii) $Tw \in \mathcal{V}$.

Proof of Lemma B.5(i): **Establishing** (B.11)**,** (B.12)**,** (B.13)**.**

Proof of (B.11):

Using the fact that patients' health conditions change independently of each other, we have for $x_1+x_2 \leq b-2$,

$$\begin{aligned}
\Gamma w(x_1+2,x_2) = & \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)\Big[q_1^2 w(i,j)+2q_1r_1w(i+1,j) \\
& +2q_1p_1w(i,j+1)+r_1^2w(i+2,j)+2r_1p_1w(i+1,j+1)+p_1^2w(i,j+2)\Big] \\
\Gamma w(x_1+1,x_2) = & \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)\left[q_1w(i,j)+r_1w(i+1,j)+p_1w(i,j+1)\right] \\
\Gamma w(x_1,x_2) = & \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)w(i,j).
\end{aligned}$$

Then, we have

$$\begin{aligned}
D_{11}\Gamma w(x_1,x_2) &= \Gamma w(x_1+2,x_2)-2\Gamma w(x_1+1,x_2)+\Gamma w(x_1,x_2) \\
&= \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)g_{11}(i,j),
\end{aligned}$$

where

$$\begin{aligned}
g_{11}(i,j) &= \Big[(r_1+p_1)^2w(i,j)-2r_1(r_1+p_1)w(i+1,j)-2p_1(r_1+p_1)w(i,j+1) \\
&\quad +r_1^2w(i+2,j)+2r_1p_1w(i+1,j+1)+p_1^2w(i,j+2)\Big] \\
&= r_1^2D_{11}w(i,j)+2p_1r_1D_{12}w(i,j)+p_1^2D_{22}w(i,j).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
D_{22}\Gamma w(x_1,x_2) &= \Gamma w(x_1,x_2+2)-2\Gamma w(x_1,x_2+1)+\Gamma w(x_1,x_2) \\
&= \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)g_{22}(i,j) \\
D_{12}\Gamma w(x_1,x_2) &= \Gamma w(x_1+1,x_2+1)-\Gamma w(x_1+1,x_2)-\Gamma w(x_1,x_2+1)+\Gamma w(x_1,x_2) \\
&= \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)g_{12}(i,j),
\end{aligned}$$

where

$$\begin{aligned} g_{22}(i,j) &= q_2^2 D_{11}w(i,j) + 2q_2r_2D_{12}w(i,j) + r_2^2 D_{22}w(i,j), \\ g_{12}(i,j) &= r_1q_2D_{11}w(i,j) + (r_1r_2 + p_1q_2)D_{12}w(i,j) + p_1r_2D_{22}w(i,j). \end{aligned}$$

Since $D_{11}w(i,j) \geq 0,\ D_{12}w(i,j) \geq 0,\ D_{22}w(i,j) \geq 0$ for $i+j \leq b-1$, we can conclude that $g_{11}(i,j) \geq 0,\ g_{22}(i,j) \geq 0,\ g_{12}(i,j) \geq 0$ for all i,j such that $i+j \leq x_1+x_2+1 \leq b-1$, and consequently,

$$D_{11}\Gamma w(x_1,x_2) \geq 0, D_{22}\Gamma w(x_1,x_2) \geq 0, D_{12}\Gamma w(x_1,x_2) \geq 0.$$

Proof of (B.12):

Conditioning on the event that state (x_1,x_2) transitions to (i_1,i_2) and state (y_1,y_2) transitions to (j_1,j_2), we can write, for $x_1+x_2 \leq b-2$ and $y_1+y_2 \leq b-2$,

$$\begin{aligned} &D_{11}\Gamma w(x_1,x_2)D_{22}\Gamma w(y_1,y_2) + D_{11}\Gamma w(y_1,y_2)D_{22}\Gamma w(x_1,x_2) - 2D_{12}\Gamma w(x_1,x_2)D_{12}\Gamma w(y_1,y_2) \\ &= \sum_{i_1=0}^{x_1+x_2+1} \sum_{i_2=0}^{x_1+x_2+1-i_1} \sum_{j_1=0}^{y_1+y_2+1} \sum_{j_2=0}^{y_1+y_2+1-j_1} P(i_1,i_2|x_1,x_2)P(j_1,j_2|y_1,y_2)g(i_1,i_2,j_1,j_2), \end{aligned}$$

where

$$
\begin{aligned}
&g(i_1,i_2,j_1,j_2) = g_{11}(i_1,i_2)g_{22}(j_1,j_2) + g_{11}(j_1,j_2)g_{22}(i_1,i_2) - 2g_{12}(i_1,i_2)g_{12}(j_1,j_2)\\
=&\Big[r_1^2 D_{11}w(i_1,i_2) + 2p_1r_1D_{12}w(i_1,i_2) + p_1^2D_{22}w(i_1,i_2)\Big]\\
&\qquad\times\Big[q_2^2D_{11}w(j_1,j_2) + 2q_2r_2D_{12}w(j_1,j_2) + r_2^2D_{22}w(j_1,j_2)\Big]\\
+&\Big[q_2^2D_{11}w(i_1,i_2) + 2q_2r_2D_{12}w(i_1,i_2) + r_2^2D_{22}w(i_1,i_2)\Big]\\
&\qquad\times\Big[r_1^2D_{11}w(j_1,j_2) + 2p_1r_1D_{12}w(j_1,j_2) + p_1^2D_{22}w(j_1,j_2)\Big]\\
-2&\Big[r_1q_2D_{11}w(i_1,i_2) + (r_1r_2+p_1q_2)D_{12}w(i_1,i_2) + p_1r_2D_{22}w(i_1,i_2)\Big]\\
&\qquad\times\Big[r_1q_2D_{11}w(j_1,j_2) + (r_1r_2+p_1q_2)D_{12}w(j_1,j_2) + p_1r_2D_{22}w(j_1,j_2)\Big]\\
=&D_{11}w(i_1,i_2)\Big\{\big[r_1^2q_2^2D_{11}w(j_1,j_2) + 2r_1^2q_2r_2D_{12}w(j_1,j_2) + r_1^2r_2^2D_{22}w(j_1,j_2)\big]\\
&+\big[q_2^2r_1^2D_{11}w(j_1,j_2) + 2q_2^2r_1p_1D_{12}w(j_1,j_2) + q_2^2p_1^2D_{22}w(j_1,j_2)\big]\\
&-\big[2r_1q_2r_1q_2D_{11}w(j_1,j_2) + 2r_1q_2(r_1r_2+p_1q_2)D_{12}w(j_1,j_2) + 2r_1q_2p_1r_2D_{22}w(j_1,j_2)\big]\Big\}\\
+&D_{22}w(i_1,i_2)\Big\{\big[p_1^2q_2^2D_{11}w(j_1,j_2) + 2p_1^2q_2r_2D_{12}w(j_1,j_2) + p_1^2r_2^2D_{22}w(j_1,j_2)\big]\\
&+\big[r_2^2r_1^2D_{11}w(j_1,j_2) + 2r_2^2r_1p_1D_{12}w(j_1,j_2) + r_2^2p_1^2D_{22}w(j_1,j_2)\big]\\
&-\big[2p_1r_2r_1q_2D_{11}w(j_1,j_2) + 2p_1r_2(r_1r_2+p_1q_2)D_{12}w(j_1,j_2) + 2p_1r_2p_1r_2D_{22}w(j_1,j_2)\big]\Big\}\\
+&D_{12}w(i_1,i_2)\Big\{\big[2r_1p_1q_2^2D_{11}w(j_1,j_2) + 4r_1p_1q_2r_2D_{12}w(j_1,j_2) + 2r_1p_1r_2^2D_{22}w(j_1,j_2)\big]\\
&+\big[2q_2r_2r_1^2D_{11}w(j_1,j_2) + 4q_2r_2r_1p_1D_{12}w(j_1,j_2) + 2q_2r_2p_1^2D_{22}w(j_1,j_2)\big]\\
&-\big[2(r_1r_2+p_1q_2)r_1q_2D_{11}w(j_1,j_2) + 2(r_1r_2+p_1q_2)^2D_{12}w(j_1,j_2)\\
&\qquad\qquad+2(r_1r_2+p_1q_2)p_1r_2D_{22}w(j_1,j_2)\big]\Big\}\\
=&(r_1r_2-p_1q_2)^2\Big[D_{11}w(i_1,i_2)D_{22}w(j_1,j_2) + D_{11}w(j_1,j_2)D_{22}w(i_1,i_2)\\
&\qquad\qquad-2D_{12}w(i_1,i_2)D_{12}w(j_1,j_2)\Big].
\end{aligned}
$$

Since $D_{11}w(i_1,i_2)D_{22}w(j_1,j_2) + D_{11}w(j_1,j_2)D_{22}w(i_1,i_2) - 2D_{12}w(i_1,i_2)D_{12}w(j_1,j_2) \geq 0$ for any $i_1+i_2 \leq b-1$ and $j_1+j_2 \leq b-1$, we can conclude that $g_{11}(i_1,i_2,j_1,j_2) \geq 0$ and consequently,

$$
\begin{aligned}
&D_{11}\Gamma w(x_1,x_2)D_{22}\Gamma w(y_1,y_2) + D_{11}\Gamma w(y_1,y_2)D_{22}\Gamma w(x_1,x_2)\\
&\qquad\qquad-2D_{12}\Gamma w(x_1,x_2)D_{12}\Gamma w(y_1,y_2) \geq 0
\end{aligned}
$$

Proof of (B.13):

Conditioning on the event that state (x_1, x_2) transitions to (i_1, i_2), we can write, for $x_1 + x_2 \leq b - 2$,

$$D_{11}\Gamma w(x_1, x_2) + D_{22}\Gamma w(x_1, x_2) - 2D_{12}\Gamma w(x_1, x_2)$$
$$= \sum_{i_1=0}^{x_1+x_2+1} \sum_{i_2=0}^{x_1+x_2+1-i_1} P(i_1, i_2|x_1, x_2) f(i_1, i_2),$$

where, for $i_1 + i_2 \leq x_1 + x_2 + 1 \leq b - 1$,

$$\begin{aligned} & f(i_1, i_2) = g_{11}(i_1, i_2) + g_{22}(i_1, i_2) - 2g_{12}(i_1, i_2) \\ = \; & (r_1 - q_2)^2 D_{11} w(i_1, i_2) + 2(r_1 - q_2)(p_1 - r_2) D_{12} w(i_1, i_2) + (p_1 - r_2)^2 D_{22} w(i_1, i_2). \end{aligned}$$

Since Condition 2 holds for w, we have for $i_1 + i_2 \leq b - 1$,

$$\begin{aligned} & D_{11}w(i_1, i_2) D_{22}w(i_1, i_2) + D_{11}w(i_1, i_2) D_{22}w(i_1, i_2) - 2D_{12}w(i_1, i_2) D_{12}w(i_1, i_2) \geq 0 \\ & \Rightarrow D_{12}w(i_1, i_2) \leq \sqrt{D_{11}w(i_1, i_2) D_{22}w(i_1, i_2)}. \end{aligned}$$

Then, if $(r_1 - q_2)(p_1 - r_2) < 0$, we have

$$\begin{aligned} f(i_1, i_2) \geq & (r_1 - q_2)^2 D_{11} w(i_1, i_2) \\ & + (p_1 - r_2)^2 D_{22} w(i_1, i_2) + 2(r_1 - q_2)(p_1 - r_2) \sqrt{D_{11}w(i_1, i_2) D_{22}w(i_1, i_2)} \\ = & \left[(r_1 - q_2) \sqrt{D_{11}w(i_1, i_2)} + (p_1 - r_2) \sqrt{D_{22}w(i_1, i_2)} \right]^2 \geq 0. \end{aligned}$$

On the other hand, if $(r_1 - q_2)(p_1 - r_2) \geq 0$, then since $D_{11}w(i_1, i_2), D_{22}w(i_1, i_2), D_{12}w(i_1, i_2)$ for $i_1 + i_2 \leq b - 1$ are nonnegative (from Condition 1), we have

$$f(i_1, i_2) = (r_1 - q_2)^2 D_{11} w(i_1, i_2) + 2(r_1 - q_2)(p_1 - r_2) D_{12} w(i_1, i_2) + (p_1 - r_2)^2 D_{22} w(i_1, i_2) \geq 0.$$

Thus, $f(i_1, i_2) \geq 0$ for all $i_1 + i_2 \leq b - 1$, and we can conclude that

$$D_{11}\Gamma w(x_1, x_2) + D_{22}\Gamma w(x_1, x_2) - 2D_{12}\Gamma w(x_1, x_2) \geq 0$$

Proof of Lemma B.5(ii):

We have already established in Lemma B.4 that $Tw \in \mathcal{V}_b$, then, we only need to show that Conditions 1 to 3 given in Definition B.5 hold for Tw. We first establish the expressions for $D_{11}Tw(x_1, x_2)$, $D_{12}Tw(x_1, x_2)$, and $D_{22}Tw(x_1, x_2)$ for the four different cases descried in Definition 3.2.

Case 1: When $x_1 + x_2 + 2 \leq b$, $Tw(i,j) = \alpha\left[q_1 i + \Gamma w(i,j)\right]$ for all $(i,j) \in \{(x_1,x_2),\ (x_1+1,x_2),\ (x_1,x_2+1),\ (x_1+2,x_2),\ (x_1+1,x_2+1),\ (x_1,x_2+2)\}$. Hence,

$$\begin{aligned} D_{11}Tw(x_1,x_2) &= Tw(x_1+2,x_2) - 2Tw(x_1+1,x_2) + Tw(x_1,x_2), \\ &= \alpha\left[\Gamma w(x_1+2,x_2) - 2\Gamma w(x_1+1,x_2) + \Gamma w(x_1,x_2)\right] = \alpha D_{11}\Gamma w(x_1,x_2). \end{aligned}$$

Similarly, $D_{22}Tw(x_1,x_2) = \alpha D_{22}\Gamma w(x_1,x_2)$ and $D_{12}Tw(x_1,x_2) = \alpha D_{12}\Gamma w(x_1,x_2)$.

Case 2: For $x_1 + x_2 = b - 1$ and $x_1 > 0, x_2 > 0$, then $Tw(i,j) = \alpha\left[q_1 i + \Gamma w(i,j)\right]$ for all $(i,j) \in \{(x_1,x_2),\ (x_1+1,x_2),\ (x_1,x_2+1)\}$. Besides,

$$\begin{aligned} Tw(x_1+2,x_2) &= \min\Big\{c_1^G + \alpha\left[q_1(x_1+1) + \Gamma w(x_1+1,x_2)\right], \\ &\qquad c_2^G + \alpha\left[q_1(x_1+2) + \Gamma w(x_1+2,x_2-1)\right]\Big\}, \\ Tw(x_1+1,x_2+1) &= \min\Big\{c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1,x_2+1)\right], \\ &\qquad c_2^G + \alpha\left[q_1(x_1+1) + \Gamma w(x_1+1,x_2)\right]\Big\}, \\ Tw(x_1,x_2+2) &= \min\Big\{c_1^G + \alpha\left[q_1(x_1-1) + \Gamma w(x_1-1,x_2+2)\right], \\ &\qquad c_2^G + \alpha\left[q_1 x_1 + \Gamma w(x_1,x_2+1)\right]\Big\}. \end{aligned}$$

Then we have,

$$\begin{aligned} D_{11}Tw(x_1,x_2) = \min\Big\{&c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right], \\ &c_2^G - \alpha D_2\Gamma w(x_1,x_2) + \alpha\left[\Delta\Gamma w(x_1+1,x_2-1) - \Delta\Gamma w(x_1,x_2)\right]\Big\}, \\ D_{12}Tw(x_1,x_2) = \min\Big\{&c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right],\ \ c_2^G - \alpha D_2\Gamma w(x_1,x_2)\Big\}, \\ D_{22}Tw(x_1,x_2) = \min\Big\{&c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right] + \alpha\left[\Delta\Gamma w(x_1,x_2) - \Delta\Gamma w(x_1-1,x_2+1)\right], \\ &c_2^G - \alpha D_2\Gamma w(x_1,x_2)\Big\}. \end{aligned}$$

Case 3: For $x_1 = b - 1$ and $x_2 = 0$, the only difference from Case 2 is the expression for $Tw(x_1+2,x_2)$. Then, $D_{12}Tw(x_1,x_2)$ and $D_{22}Tw(x_1,x_2)$ are the same. From Definition 3.2, $Tw(x_1+2,x_2) = c_1^G + \alpha[q_1(x_1+1) + \Gamma w(x_1+1,x_2)]$. Then,

$$D_{11}Tw(x_1,x_2) = c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right].$$

Case 4: For $x_1 = 0$ and $x_2 = b-1$, the only difference from Case 2 is the expression for $Tw(x_1, x_2+2)$. Then, $D_{11}Tw(x_1,x_2)$ and $D_{12}Tw(x_1,x_2)$ are the same. From Definition 3.2, $Tw(x_1,x_2+2) = c_2^G + \alpha[q_1 x_1 + \Gamma w(x_1, x_2+1)]$. Then,

$$D_{22}Tw(x_1,x_2) = c_2^G - \alpha D_2 \Gamma w(x_1,x_2).$$

Proof of Condition 1:

For $x_1 + x_2 \le b-2$, $D_{11}Tw(x_1,x_2) \ge 0,\ D_{22}Tw(x_1,x_2) \ge 0$ and $D_{12}Tw(x_1,x_2) \ge 0$ from (B.11) as we have established earlier in the proof of this lemma.

For $x_1 + x_2 = b-1$ and $w \in \mathcal{V}$, from Definition B.5 we have $w \in \mathcal{V}_b$, and using Lemma B.4(a), we have

$$c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right] \ge 0,\ c_2^G - \alpha D_2 \Gamma w(x_1,x_2) \ge 0.$$

Then, we can conclude $D_{11}Tw(x_1,x_2) \ge 0$ for Case 3, $D_{22}Tw(x_1,x_2) \ge 0$ for Case 4, and $D_{12}Tw(x_1,x_2) \ge 0$ for Cases 2, 3 and 4.

Furthermore, for any $i + j \le b-2$, we have

$$\begin{aligned}
&\Delta\Gamma w(i+1,j) - \Delta\Gamma w(i,j+1) \\
&= \Gamma w(i+2,j) - \Gamma w(i+1,j+1) - \Gamma w(i+1,j+1) + \Gamma w(i,j+2) \\
&= D_{11}\Gamma w(i,j) + D_{22}\Gamma w(i,j) - 2D_{12}\Gamma w(i,j) \ge 0,
\end{aligned}$$

where the inequality follows from (B.13) as we have established earlier in the proof of this lemma.

It then follows that for $x_1 + x_2 = b-1$ and $x_1 > 0$(Cases 2 and 3), $\Delta\Gamma w(x_1,x_2) - \Delta\Gamma w(x_1-1,x_2+1) \ge 0$ and thus $D_{22}Tw(x_1,x_2) \ge 0$, and for $x_1 + x_2 = b-1$ and $x_2 > 0$ (Cases 2 and 4), $\Delta\Gamma w(x_1+1,x_2-1) - \Delta\Gamma w(x_1,x_2) \ge 0$ and thus $D_{11}Tw(x_1,x_2) \ge 0$.

Hence, $D_{11}T(x_1,x_2) \ge 0,\ D_{22}T(x_1,x_2) \ge 0,\ D_{12}T(x_1,x_2) \ge 0$ for all four cases.

Proof of Condition 3: (We prove this condition first because it will be used in the proof of Condition 2). For $x_1 + x_2 \le b-2$,

$$\begin{aligned}
&D_{11}Tw(x_1,x_2) + D_{22}Tw(x_1,x_2) - 2D_{12}Tw(x_1,x_2) \\
&= \alpha\big[D_{11}\Gamma w(x_1,x_2) + D_{22}\Gamma w(x_1,x_2) - 2D_{12}\Gamma w(x_1,x_2)\big] \ge 0,
\end{aligned}$$

where the inequality follows from (B.13) as we have established earlier in the proof of this lemma.

For $x_1 + x_2 = b - 1$, $\Delta\Gamma w(x_1 + 1, x_2 - 1) - \Delta\Gamma w(x_1, x_2) \geq 0$ if $x_2 > 0$ and $\Delta\Gamma w(x_1, x_2) - \Delta\Gamma w(x_1 - 1, x_2 + 1) \geq 0$ for $x_1 > 0$ (as we have already established in the proof of Condition 1). From the expressions of $D_{11}Tw(x_1, x_2)$, $D_{22}Tw(x_1, x_2)$ and $D_{12}Tw(x_1, x_2)$, we have

$$D_{11}Tw(x_1, x_2) \geq D_{12}Tw(x_1, x_2),\ D_{22}Tw(x_1, x_2) \geq D_{12}Tw(x_1, x_2),$$

Then for $x_1 + x_2 = b - 1$, $D_{11}Tw(x_1, x_2) + D_{22}Tw(x_1, x_2) - 2D_{12}Tw(x_1, x_2) \geq 0$.

Hence, Condition 3 holds for all $x_1 + x_2 \leq b - 1$.

Proof of Condition 2:

If $x_1 + x_2 \leq b - 2$ and $y_1 + y_2 \leq b - 2$, then from Case 1,

$$\begin{aligned}
&D_{11}Tw(x_1, x_2)D_{22}Tw(y_1, y_2) + D_{11}Tw(y_1, y_2)D_{22}Tw(x_1, x_2) - 2D_{12}Tw(x_1, x_2)D_{12}Tw(y_1, y_2)\\
=&\alpha^2\big[D_{11}\Gamma w(x_1, x_2)D_{22}\Gamma w(y_1, y_2) + D_{11}\Gamma w(y_1, y_2)D_{22}\Gamma w(x_1, x_2) - 2D_{12}\Gamma w(x_1, x_2)D_{12}\Gamma w(y_1, y_2)\big]\\
\geq&0,
\end{aligned}$$

where the inequality follows from (B.12) as we have established earlier in the proof of this lemma.

If $x_1 + x_2 = b - 1$, then $D_{11}Tw(x_1, x_2) \geq D_{12}Tw(x_1, x_2)$ and $D_{22}Tw(x_1, x_2) \geq D_{12}Tw(x_1, x_2)$ from the proof of Condition 3. Then for any $y_1 + y_2 \leq b - 1$,

$$\begin{aligned}
&D_{11}Tw(x_1, x_2)D_{22}Tw(y_1, y_2) + D_{11}Tw(y_1, y_2)D_{22}Tw(x_1, x_2) - 2D_{12}Tw(x_1, x_2)D_{12}Tw(y_1, y_2)\\
\geq&D_{12}Tw(x_1, x_2)D_{22}Tw(y_1, y_2) + D_{11}Tw(y_1, y_2)D_{12}Tw(x_1, x_2) - 2D_{12}Tw(x_1, x_2)D_{12}Tw(y_1, y_2)\\
=&D_{12}Tw(x_1, x_2)\big[D_{22}Tw(y_1, y_2) + D_{11}Tw(y_1, y_2) - 2D_{12}Tw(y_1, y_2)\big] \geq 0,
\end{aligned}$$

where the last inequality holds since we have proved that $Tw(y_1, y_2)$ satisfies Condition 3 for all $y_1 + y_2 \leq b - 1$.

If $y_1 + y_2 = b - 1$, then for any $x_1 + x_2 \leq b - 1$ we can similarly show

$$\begin{aligned}
&D_{11}Tw(x_1, x_2)D_{22}Tw(y_1, y_2) + D_{11}Tw(y_1, y_2)D_{22}Tw(x_1, x_2)\\
&\qquad - 2D_{12}Tw(x_1, x_2)D_{12}Tw(y_1, y_2) \geq 0.
\end{aligned}$$

Hence, Condition 2 holds for all $x_1 + x_2 \leq b - 1$ and $y_1 + y_2 \leq b - 1$.

Thus, $Tw \in \mathcal{V}$.

□

Lemma B.6. *Suppose that (3.10) holds. Then,*

(a) the optimal value function $v_\alpha \in \mathcal{V}$, and

(b) for $x_1 + x_2 \leq b-1$ and $x_1 > 0$, $\Delta\Gamma v_\alpha(x_1, x_2) \geq \Delta\Gamma v_\alpha(x_1 - 1, x_2 + 1)$.

***Proof of Lemma B.6**:* (a) The proof is based on Theorem 11.5 of (Porteus 2002), which requires the existence of three conditions for the optimality of structured policies.

(i) Completeness.

Define the distance of two functions u, v in $\mathcal{V}$ by

$$\rho(u, v) := \sup_{s \in \mathcal{S}} |u(s) - v(s)| \text{ for } u, v \in \mathcal{V}.$$

We need to show that $(\rho, \mathcal{V})$ is complete (see, e.g., (Porteus 2002) for a definition). Specifically, we need to show that for any Cauchy sequence $\{v_n, n \geq 0\}$ in $\mathcal{V}$, there must exist $v \in \mathcal{V}$ such that $\lim_{n\to\infty} \rho(v_n, v) = 0$.

First we show that any Cauchy sequence in $\mathcal{V}$ is convergent. Let $V = R^{(b+2)\times(b+2)}$ be a $b+2$ by $b+2$ dimensional real vector space. Then, $\mathcal{V} \subset V$, and hence for any Cauchy sequence $\{v_n, n = 1, 2, \ldots\}$ in $\mathcal{V}$, is also a Cauchy sequence in V. It is known that V is complete, thus, $\{v_n, n = 1, 2, \ldots\}$ has a limit in V, i.e., there exists $v \in V$ such that $\lim_{n\to\infty} \rho(v_n, v) = 0$.

Next we show that the limit $v \in \mathcal{V}$. Now, for any sequence of value functions $\{v_n : v_n \in \mathcal{V}, n = 1, 2, \ldots\}$, we know that

$$c_1 \leq D_1 v_n(x_1, x_2) \leq c_1^G,\ c_2 \leq D_2 v_n(x_1, x_2) \leq c_2^G \text{ for } x_1 + x_2 \leq b,$$

$$D_{11} v_n(x_1, x_2) \geq 0,\ D_{22} v_n(x_1, x_2) \geq 0, \text{ and } D_{12} v_n(x_1, x_2) \geq 0 \text{ for } x_1 + x_2 \leq b-1,$$

$$D_{11} v_n(x_1, x_2) D_{22} v_n(y_1, y_2) + D_{11} v_n(y_1, y_2) D_{22} v_n(x_1, x_2) - 2 D_{12} v_n(x_1, x_2) D_{12} v_n(y_1, y_2) \geq 0, \text{ for } x_1 + x_2 \leq b-1 \text{ and } y_1 + y_2 \leq b-1,$$

$$D_{11} v_n(x_1, x_2) + D_{22} v_n(x_1, x_2) - 2 D_{12} v_n(x_1, x_2) \geq 0 \text{ for } x_1 + x_2 \leq b-1.$$

Suppose that $v \notin \mathcal{V}$. Then, it must be the case that at least one of the inequalities above, which define the set $\mathcal{V}$ does not hold for sufficiently large n.This is a contradiction to the fact that $v_n \in \mathcal{V}$. Thus, $(\rho, \mathcal{V})$ is complete.

(ii) Attainment. For any $w \in \mathcal{V}$, we must show that there exists a decision rule that can attain the minimum. Define a decision rule δ, which in state $(x_1, x_2) \in S$, does not discharge any patient if $x_1 + x_2 \leq b$, discharges a stage 1 patient if $x_1 = b+1, x_2 = 0$, discharges a stage 2 patient at state $x_1 =$

$0, x_2 = b+1$, and otherwise discharges a stage 1 patient if $\Gamma w(x_1-1,x_2) - \Gamma w(x_1,x_2-1) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1$, and discharges a stage 2 otherwise. Then, according to the optimality equations (3.11), δ is optimal.

(iii) Preservation. This follows immediately from Lemma B.5.

Then, from Theorem 11.5 of (Porteus 2002), we can conclude that the optimal value function $v_\alpha \in \mathcal{V}$.

(b) It follows from Lemma B.5 (i) that $\Gamma v_\alpha(x_1,x_2)$ satisfies (B.13), i.e., for any $i+j \leq b-2$,

$$D_{11}\Gamma v_\alpha(i,j) + D_{22}\Gamma v_\alpha(i,j) - 2D_{12}\Gamma v_\alpha(i,j) \geq 0.$$

Then, for any $x_1 + x_2 \leq b-1$ and $x_1 > 0$,

$$\begin{aligned} & \Delta\Gamma v_\alpha(x_1,x_2) - \Delta\Gamma v_\alpha(x_1-1,x_2+1) \\ = \;& \Gamma v_\alpha(x_1+1,x_2) - \Gamma v_\alpha(x_1,x_2+1) - \Gamma v_\alpha(x_1,x_2+1) + \Gamma v_\alpha(x_1-1,x_2+2) \\ = \;& D_{11}\Gamma v_\alpha(x_1-1,x_2) + D_{22}\Gamma v_\alpha(x_1-1,x_2) - 2D_{12}\Gamma v_\alpha(x_1-1,x_2) \geq 0, \end{aligned}$$

□

Proof of Proposition 3.3: Define function $\delta_n(i) = \Delta\Gamma v_\alpha(i, n-i)$ for $1 \leq i \leq n \leq b-1$. Then, from Lemma B.6 (b), we have $\delta_n(i) \geq \delta_n(i-1)$. Thus, for fixed n, $\delta_n(i)$ is non-decreasing in i for $1 \leq i \leq n$.

From the optimality equations (3.11) and Definition 3.2, for $x_1 + x_2 = b+1$ and $x_1 > 0, x_2 > 0$, $a^*(x_1,x_2) = (1,0)$ if and only if

$$c_1^G + \alpha\left[q_1(x_1-1) + \Gamma v_\alpha(x_1-1,x_2)\right] < c_2^G + \alpha\left[q_1 x_1 + \Gamma v_\alpha(x_1,x_2-1)\right],$$

which can equivalently be written as

$$\delta_{b-1}(x_1-1) > \frac{c_1^G - c_2^G}{\alpha} - q_1. \tag{B.14}$$

First, suppose that there exists $x_1 \in [1,b]$ such that $\delta_{b-1}(x_1-1) > \frac{c_1^G - c_2^G}{\alpha} - q_1$, and let

$$x_\alpha^* = \min\left\{x_1 : 1 \leq x_1 \leq b \text{ and } \delta_{b-1}(x_1-1) > \frac{c_1^G - c_2^G}{\alpha} - q_1\right\}.$$

Then, since $\delta_n(i)$ is non-decreasing in i, we have $\delta_{b-1}(x_1-1) > \frac{c_1^G - c_2^G}{\alpha} - q_1$ for all $x_1 \in [x^*, b]$, and $\delta_{b-1}(x_1-1) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1$ for all $x_1 \in [1, x^*)$. Thus, we have $a_\alpha^*(x_1,x_2) = (1,0)$ for $x_1 + x_2 = b+1$ and $x_1 \geq x^*$, and $a_\alpha^*(x_1,x_2) = (0,1)$ for $x_1 + x_2 = b+1$ and $x_1 < x^*$.

Now suppose that $\delta_{b-1}(x_1-1) < \frac{c_1^G-c_2^G}{\alpha} - q_1$ for all $x_1+x_2 = b+1$ and $x_1 > 0, x_2 > 0$, then let $x_\alpha^* = b+1$ and the result follows.

□

B.2.3 Proof of Proposition 3.4

Definition B.6. Define an operator F on w by

$$Fw(i,j) \quad = \quad (r_1-q_2)D_1w(i,j) + (p_1-r_2)D_2w(i,j).$$

Definition B.7. Let $\mathcal{V}_1$ be a set of functions such that if $w \in \mathcal{V}_1$, then $w \in \mathcal{V}$ and

$$\text{Condition 4: } \Delta w(x_1,x_2) = w(x_1+1,x_2) - w(x_1,x_2+1) \geq c_1^G - c_2^G, \qquad \text{for } x_1+x_2 \leq b,$$

$$\text{Condition 5: } Fw(x_1,x_2) > \frac{c_1^G-c_2^G}{\alpha} - q_2, \qquad \text{for } x_1+x_2 \leq b.$$

Lemma B.7. *If (3.10) and (3.17) hold, then for any function $w \in \mathcal{V}_1$,*

(a) $\Delta\Gamma w(x_1,x_2) > \frac{c_1^G-c_2^G}{\alpha} - q_1$, *for all* $x_1+x_2 \leq b-1$.

(b) $Tw \in \mathcal{V}_1$.

(c) $v_\alpha \in \mathcal{V}_1$.

Proof of Lemma B.7: From (3.4), we can show that (3.17) is equivalent to

$$c_1 - c_2 > c_1^G - c_2^G, \; q_1 \leq p_2. \tag{B.15}$$

(a) By conditioning on how x_1 stage 1 patients and x_2 stage 2 patients evolve, we have for $x_1+x_2 \leq b-1$,

$$\Delta\Gamma w(x_1,x_2) = \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)Fw(i,j) > \frac{c_1^G-c_2^G}{\alpha} - q_1,$$

where the inequality follows from Condition 5 for $w \in \mathcal{V}_1$ and $i+j \leq x_1+x_1+1 \leq b$.

(b) To show $Tw \in \mathcal{V}_1$, we need to show Conditions 4 and 5 hold for Tw since we have already shown $Tw \in \mathcal{V}$ for $w \in \mathcal{V}$.

We first establish the expression of $\Delta[Tw(i,j)]$ for the four difference cases described in Definition 3.2.

Case 1: If $x_1+x_2 \leq b-1$, we have,

$$\Delta[Tw(i,j)] = \alpha[q_1 + \Delta\Gamma w(x_1,x_2)].$$

Case 2: If $x_1 + x_2 = b$ and $x_1 > 0, x_2 > 0$, then from part (a) we have

$$\Delta\Gamma w(x_1, x_2 - 1) = \Gamma w(x_1 + 1, x_2 - 1) - \Gamma w(x_1, x_2) > \frac{c_1^G - c_2^G}{\alpha} - q_1,$$

which is equivalent to

$$c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right] < c_2^G + \alpha\left[q_1(x_1 + 1) + \Gamma w(x_1 + 1, x_2 - 1)\right].$$

Then,

$$\begin{aligned} Tw(x_1 + 1, x_2) &= \min\Big\{c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right], \\ &\qquad c_2^G + \alpha\left[q_1(x_1 + 1) + \Gamma w(x_1 + 1, x_2 - 1)\right]\Big\} \\ &= c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right] = c_1^G + Tw(x_1, x_2), \end{aligned}$$

and thus $D_1\left[Tw(x_1, x_2)\right] = Tw(x_1 + 1, x_2) - Tw(x_1, x_2) = c_1^G$.

Similarly, from part (a) we have

$$\Delta\Gamma w(x_1 - 1, x_2) = \Gamma w(x_1, x_2) - \Gamma w(x_1 - 1, x_2 + 1) > \frac{c_1^G - c_2^G}{\alpha} - q_1,$$

which is equivalent to

$$c_1^G + \alpha\left[q_1(x_1 - 1) + \Gamma w(x_1 - 1, x_2 + 1)\right] < c_2^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right].$$

Then,

$$\begin{aligned} Tw(x_1, x_2 + 1) &= \min\Big\{c_1^G + \alpha\left[q_1(x_1 - 1) + \Gamma w(x_1 - 1, x_2 + 1)\right], \\ &\qquad c_2^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right]\Big\} \\ &= c_1^G + \alpha\left[q_1(x_1 - 1) + \Gamma w(x_1 - 1, x_2 + 1)\right] \\ &= c_1^G + Tw(x_1 - 1, x_2 + 1), \end{aligned}$$

and thus $D_2\left[Tw(x_1, x_2)\right] = Tw(x_1, x_2 + 1) - Tw(x_1, x_2) = c_1^G - \Delta\left[Tw(x_1 - 1, x_2)\right]$. Hence,

$$\Delta\left[Tw(x_1, x_2)\right] = D_1\left[Tw(x_1, x_2)\right] - D_2\left[Tw(x_1, x_2)\right] = \Delta\left[Tw(x_1 - 1, x_2)\right]. \tag{B.16}$$

Case 3: If $x_1 = b$ and $x_2 = 0$, then $Tw(x_1, x_2 + 1)$ and $Tw(x_1, x_2)$ have the same expressions as in Case 2 and thus $D_2 [Tw(x_1, x_2)] = c_1^G - \Delta [Tw(x_1 - 1, x_2)]$, and

$$Tw(x_1 + 1, x_2) = c_1^G + \alpha [q_1 x_1 + \Gamma w(x_1, x_2)] = c_1^G - \Delta [Tw(x_1, x_2)],$$

and thus $D_1 [Tw(x_1, x_2)] = c_1^G$. Then,

$$\Delta [Tw(x_1, x_2)] = \Delta [Tw(x_1 - 1, x_2)].$$

Case 4: If $x_1 = 0$ and $x_2 = b$, then $Tw(x_1 + 1, x_2)$ and $Tw(x_1, x_2)$ have the same expressions as in Case 2 and thus $D_1 [Tw(x_1, x_2)] = c_1^G$, and

$$Tw(x_1, x_2 + 1) = c_2^G + \alpha [q_1(x_1 + 1) + \Gamma w(x_1, x_2)],$$

and thus $D_2 [Tw(x_1, x_2)] = c_2^G$. Then,

$$\Delta [Tw(x_1, x_2)] = c_1^G - c_2^G.$$

Proof of Condition 4:

Case 1: If $x_1 + x_2 \leq b - 1$, we have $\Delta Tw(i, j) = \alpha[q_1 + \Delta\Gamma w(x_1, x_2)] > c_1^G - c_2^G$, where the inequality follows from part (a).

Case 2 and 3: If $x_1 + x_2 = b$ and $x_1 > 0$, we have,

$$\Delta [Tw(x_1, x_2)] = \Delta [Tw(x_1 - 1, x_2)] > c_1^G - c_2^G,$$

where the inequality follows from Case 1.

Case 4: If $x_1 = 0$ and $x_2 = b$, then $\Delta Tw(x_1, x_2) = c_1^G - c_2^G$.

Thus, $\Delta Tw(x_1, x_2) \geq c_1^G - c_2^G$ for all $x_1 + x_2 \leq b$.

Proof of Condition 5:

(i) If $r_1 \geq q_2$, then for $i + j \leq b$,

$$\begin{aligned}
F[Tw(i,j)] &= (r_1 - q_2)D_1[Tw(i,j)] + (p_1 - r_2)D_2[Tw(i,j)] \\
&= (r_1 - q_2)\Delta[Tw(i,j)] + (p_2 - q_1)D_2[Tw(i,j)] \\
&\geq (r_1 - q_2)(c_1^G - c_2^G) + (p_2 - q_1)c_2 \\
&= (c_1^G - c_2^G) - (p_1 + q_1 + q_2)(c_1^G - c_2^G) + (p_2 - q_1)c_2 \\
&> (c_1^G - c_2^G) - (p_1 + q_1 + q_2)(c_1 - c_2) + (p_2 - q_1)c_2 \\
&= (c_1^G - c_2^G) - (c_1 - c_2) + (r_1 - q_2)c_1 + (p_1 - r_2)c_2 \\
&= \frac{c_1^G - c_2^G}{\alpha} + (c_1^G - c_2^G)\left[1 - \frac{1}{\alpha}\right] - (c_1 - c_2) + \frac{c_1 - c_2}{\alpha} - q_1 \\
&= \frac{c_1^G - c_2^G}{\alpha} - q_1 + \left[1 - \frac{1}{\alpha}\right]\left[(c_1^G - c_2^G) - (c_1 - c_2)\right] \geq \frac{c_1^G - c_2^G}{\alpha} - q_1,
\end{aligned}$$

where the first inequality follows from $\Delta Tw(i,j) \geq c_1^G - c_2^G$ which we established earlier, and $D_2Tw(i,j) \geq c_2$ since $Tw \in V_b$, and the second inequality follows from (B.15). Hence, $F[Tw(i,j)] > \frac{c_1^G - c_2^G}{\alpha} - q_1$.

(ii) If $r_1 < q_2$, since $q_1 \leq p_2$ from (B.17), we have $p_1 > r_2$ using the fact that $p_1+q_1+r_1 = p_2+q_2+r_2 = 1$. Then, $p_1 q_2 > r_1 r_2$. Next we consider four different cases as before.

Case 1: If $x_1 + x_2 \leq b - 1$,

$$\begin{aligned}
F[Tw(x_1,x_2)] &= (r_1 - q_2)D_1Tw(x_1,x_2) + (p_1 - r_2)D_2Tw(x_1,x_2) \\
&= \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)H(i,j),
\end{aligned}$$

where

$$\begin{aligned}
H(i,j) = (r_1 - q_2)\alpha\left[q_1 + r_1 D_1 w(i,j) + p_1 D_2 w(i,j)\right] \\
+ (p_1 - r_2)\alpha\left[q_2 D_1 w(i,j) + r_2 D_2 w(i,j)\right].
\end{aligned}$$

We have,

$$\begin{aligned}
\frac{H(i,j)}{\alpha} =& (r_1 - q_2)\left[q_1 + r_1 D_1 w(i,j) + p_1 D_2 w(i,j)\right] \\
&\qquad + (p_1 - r_2)\left[q_2 D_1 w(i,j) + r_2 D_2 w(i,j)\right] \\
=& -q_1(r_2+q_2) + (r_1+r_2)\left[q_1 + (r_1-q_2)D_1 w(i,j) + (p_1-r_2)D_2 w(i,j)\right] \\
&+ (p_1 q_2 - r_1 r_2)\left[D_1 w(i,j) - D_2 w(i,j)\right] \\
=& -q_1(r_2+q_2) + (r_1+r_2)[q_1 + Fw(i,j)] + (p_1 q_2 - r_1 r_2)\Delta w(i,j) \\
\geq& -q_1(r_2+q_2) + (r_1+r_2)\left(\frac{c_1^G - c_2^G}{\alpha}\right) + (p_1 q_2 - r_1 r_2)(c_1^G - c_2^G),
\end{aligned}$$

where the inequality follows from Conditions 4 and 5 for $w \in \mathcal{V}_1$. Then,

$$\begin{aligned}
& \alpha H(i,j) - \left[c_1^G - c_2^G - \alpha q_1\right] \\
\geq\ & \alpha q_1\Big[(1-\alpha(r_2+q_2)\Big] + \Big[\alpha^2(r_1+r_2) + \alpha^2(p_1 q_2 - r_1 r_2) - 1\Big](c_1^G - c_2^G) \\
=\ & \alpha q_1(1 - \alpha r_2 - \alpha q_2) - \left[1 - \alpha(r_1+r_2) + \alpha^2 r_1 r_2 - \alpha^2 p_1 q_2\right](c_1^G - c_2^G) \\
=\ & \alpha q_1(1 - \alpha r_2 - \alpha q_2) - \left[(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2\right](c_1^G - c_2^G) \\
=\ & \left[(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2\right]\left[\frac{\alpha q_1(1-\alpha r_2 - \alpha q_2)}{(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2} - (c_1^G - c_2^G)\right] \\
=\ & \left[(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2\right]\left[(c_1 - c_2) - (c_1^G - c_2^G)\right] > 0,
\end{aligned}$$

since $(1-\alpha r_1)(1-\alpha r_2)) - \alpha^2 p_1 q_2 = (1-\alpha+\alpha q_1+\alpha p_1)(1-\alpha+\alpha p_2+\alpha q_2) - \alpha^2 p_1 q_2 > 0$ and $c_1 - c_2 > c_1^G - c_2^G$. Hence, $H(i,j) > \frac{c_1^G - c_2^G}{\alpha} - q_1$ and

$$F[Tw(x_1,x_2)] \quad = \quad \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2) H(i,j) > \frac{c_1^G - c_2^G}{\alpha} - q_1.$$

Cases 2 and 3: If $x_1 + x_2 = b$ and $x_1 > 0$, we have $D_1[Tw(x_1, x_2)] = c_1^G$ and $D_2[Tw(x_1, x_2)] = c_1^G - \Delta[Tw(x_1 - 1, x_2)]$, and hence,

$$\begin{aligned} F[Tw(x_1, x_2)] &= (r_1 - q_2)D_1[Tw(x_1, x_2)] + (p_1 - r_2)D_2[Tw(x_1, x_2)] \\ &= (r_1 - q_2)c_1^G + (p_1 - r_2)\big[c_1^G - \Delta[Tw(x_1 - 1, x_2)]\big] \\ &= (p_2 - q_1)c_1^G - (p_1 - r_2)\Delta[Tw(x_1 - 1, x_2)] \\ &\geq (p_2 - q_1)D_1[Tw(x_1 - 1, x_2)] - (p_1 - r_2)\Delta[Tw(x_1 - 1, x_2)] \\ &= (r_1 - q_2)D_1[Tw(x_1 - 1, x_2)] + (p_1 - r_2)D_2[Tw(x_1 - 1, x_2)] \\ &= F[Tw(x_1 - 1, x_2)] > \frac{c_1^G - c_2^G}{\alpha} - q_1, \end{aligned}$$

where the last inequality follows from Case 1.

Case 4: If $x_1 = 0$ and $x_2 = b$, we have $D_1[Tw(x_1, x_2)] = c_1^G$ and $D_2[Tw(x_1, x_2)] = c_2^G$, and then,

$$\begin{aligned} F[Tw(x_1, x_2)] &= (r_1 - q_2)D_1 Tw(x_1, x_2) + (p_1 - r_2)D_2 Tw(x_1, x_2) \\ &= (r_1 - q_2)c_1^G + (p_1 - r_2)c_2^G = (r_2 - p_1)(c_1^G - c_2^G) + (p_2 - q_1)c_1^G \\ &\geq (r_2 - p_1)(c_1 - c_2) + (p_2 - q_1)c_1 = \frac{c_1 - c_2}{\alpha} - q_1 > \frac{c_1^G - c_2^G}{\alpha} - q_1, \end{aligned}$$

where the first inequality follows from $r_2 - p_1 < 0$, $c_1 - c_2 > c_1^G - c_2^G$, $p_2 - q_1 \geq 0$ and $c_1^G \geq c_1$.

Thus, Conditions 4 and 5 hold for Tw, and hence $Tw \in \mathcal{V}_1$.

(c) From Theorem 11.5 of (Porteus 2002), we need to verify three conditions:

(i) Completeness. The proof is very similar to that of Lemma B.6 and thus is skipped.

(ii) Attainment. For any function $w \in \mathcal{V}_1$, we define a decision rule, which in state (x_1, x_2), discharges no patient when $x_1 + x_2 \leq b$, discharges a stage 2 patient when $x_1 = 0, x_2 = b + 1$ and discharges a stage 1 patient when $x_1 + x_2 = b + 1$ and $x_1 > 0$. This rule attains the minimum of $w = Tw$ from (B.15).

(iii) Preservation. This follows immediately from part (b).

Then, from Theorem 11.5 of (Porteus 2002), we can conclude that the optimal value function $v_\alpha \in \mathcal{V}_1$.

□

Definition B.8. Let $\mathcal{V}_2$ be a set of functions such that if $w \in \mathcal{V}_2$, then $w \in \mathcal{V}$ and

$$\text{Condition 6: } \Delta w(x_1, x_2) = w(x_1 + 1, x_2) - w(x_1, x_2 + 1) \leq c_1^G - c_2^G \qquad \text{for } x_1 + x_2 \leq b,$$

$$\text{Condition 7: } Fw(x_1, x_2) \leq \frac{c_1^G - c_2^G}{\alpha} - q_2 \qquad \text{for } x_1 + x_2 \leq b.$$

Lemma B.8. *If* (3.10) *and (3.18) hold, then for any function* $w \in \mathcal{V}_2$,

(a) $\Delta\Gamma w(x_1, x_2) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1$, *for all* $x_1 + x_2 \leq b - 1$.

(b) $Tw \in \mathcal{V}_2$.

(c) $v_\alpha \in \mathcal{V}_2$.

Proof of Lemma B.8: From (3.4), we can show that (3.18) is equivalent to

$$c_1 - c_2 \leq c_1^G - c_2^G, \; q_1 \geq p_2. \tag{B.17}$$

(a) By conditioning on how x_1 stage 1 patients and x_2 stage 2 patients evolve, we have for $x_1 + x_2 \leq b - 1$,

$$\Delta\Gamma w(x_1, x_2) = \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i, j|x_1, x_2) F w(i, j) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1,$$

where the inequality follows since we know that Condition 7 holds for $w \in \mathcal{V}_2$.

(b) To show $Tw \in \mathcal{V}_2$, it is sufficient to show Conditions 6 and 7 hold for Tw since we have already shown $Tw \in \mathcal{V}$ for $w \in \mathcal{V}$.

We first establish the expression of $\Delta\left[Tw(x_1, x_2)\right]$ for the four different cases described in Definition 3.2.

Case 1: If $x_1 + x_2 \leq b - 1$, we have, $\Delta T w(i, j) = \alpha[q_1 + \Delta\Gamma w(x_1, x_2)]$.

Case 2: If $x_1 + x_2 = b$ and $x_1 > 0, x_2 > 0$, then from part (a) we have

$$\Delta\Gamma w(x_1, x_2 - 1) = \Gamma w(x_1 + 1, x_2 - 1) - \Gamma w(x_1, x_2) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1,$$

which is equivalent to

$$c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right] \geq c_2^G + \alpha\left[q_1(x_1 + 1) + \Gamma w(x_1 + 1, x_2 - 1)\right].$$

Then,

$$\begin{aligned} Tw(x_1 + 1, x_2) &= \min\Big\{c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right], \\ &\qquad\qquad c_2^G + \alpha\left[q_1(x_1 + 1) + \Gamma w(x_1 + 1, x_2 - 1)\right]\Big\} \\ &= c_2^G + \alpha\left[q_1(x_1 + 1) + \Gamma w(x_1 + 1, x_2 - 1)\right] \\ &= c_2^G + Tw(x_1 + 1, x_2 - 1), \end{aligned}$$

and thus $D_1\left[Tw(x_1, x_2)\right] = Tw(x_1 + 1, x_2) - Tw(x_1, x_2) = c_2^G + \Delta\left[Tw(x_1, x_2 - 1)\right]$.

Similarly,

$$\begin{aligned} Tw(x_1, x_2+1) &= \min\Big\{c_1^G + \alpha\left[q_1(x_1-1) + \Gamma w(x_1-1, x_2+1)\right], \\ &\qquad\qquad c_2^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right]\Big\} \\ &= c_2^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right] = c_2^G + Tw(x_1, x_2), \end{aligned}$$

and thus $D_2\left[Tw(x_1, x_2)\right] = Tw(x_1, x_2+1) - Tw(x_1, x_2) = c_2^G$.

Hence, $\Delta Tw(x_1, x_2) = D_1\left[Tw(x_1, x_2)\right] - D_2\left[Tw(x_1, x_2)\right] = \Delta Tw(x_1, x_2-1)$.

Case 3: If $x_1 = b$ and $x_2 = 0$, then $Tw(x_1, x_2+1)$ and $Tw(x_1, x_2)$ have the same expressions as in Case 2 and thus $D_2\left[Tw(x_1, x_2)\right] = c_2^G$, and

$$Tw(x_1+1, x_2) = c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right] = c_1^G + T(x_1, x_2),$$

and thus $D_1\left[Tw(x_1, x_2)\right] = c_1^G$. Then, $\Delta Tw(x_1, x_2) = c_1^G - c_2^G$.

Case 4: If $x_1 = 0$ and $x_2 = b$, then from Definition 3.2, $Tw(x_1+1, x_2)$ and $Tw(x_1, x_2)$ have the same expressions as in Case 2 and thus $D_1\left[Tw(x_1, x_2)\right] = c_2^G + \Delta\left[Tw(x_1, x_2-1)\right]$ and

$$Tw(x_1, x_2+1) = c_2^G + \alpha\left[q_1 x_1 + \Gamma w(x_1, x_2)\right] = c_2^G + Tw(x_1, x_2).$$

Then, $D_2\left[Tw(x_1, x_2)\right] = c_2^G$. Hence, $\Delta Tw(x_1, x_2) = \Delta\left[Tw(x_1, x_2-1)\right]$.

Proof of Condition 6:

We consider four different cases as before.

Case 1: If $x_1 + x_2 \leq b-1$, we have, $\Delta Tw(i,j) = \alpha[q_1 + \Delta\Gamma w(x_1, x_2)] \leq c_1^G - c_2^G$, where the inequality follows from part (a).

Cases 2 and 4: If $x_1 + x_2 = b$ and $x_2 > 0$, then $\Delta Tw(x_1, x_2) = \Delta\left[Tw(x_1, x_2-1)\right] \leq c_1^G - c_2^G$, where the inequality follows from Case 1.

Case 3: If $x_1 = b$ and $x_2 = 0$, then $\Delta Tw(x_1, x_2) = c_1^G - c_2^G$.

Thus, Condition 6 holds for Tw.

Proof of Condition 7:

(i) If $p_1 \leq r_2$,

$$\begin{aligned}
F[Tw(i,j)] =& (p_2 - q_1)D_1[Tw(i,j)] - (p_1 - r_2)\Delta[Tw(i,j)] \\
\leq& (p_2 - q_1)c_1 - (p_1 - r_2)(c_1^G - c_2^G) \\
=& (c_1^G - c_2^G) - (p_1 + p_2 + q_2)(c_1^G - c_2^G) + (p_2 - q_1)c_1 \\
\leq& (c_1^G - c_2^G) - (p_1 + p_2 + q_2)(c_1 - c_2) + (p_2 - q_1)c_1 \\
=& (c_1^G - c_2^G) - (c_1 - c_2) + (p_1 - r_2)c_2 + (r_1 - q_2)c_1 \\
=& \frac{c_1^G - c_2^G}{\alpha} + (c_1^G - c_2^G)\left[1 - \frac{1}{\alpha}\right] - (c_1 - c_2) + \frac{c_1 - c_2}{\alpha} - q_1 \\
=& \frac{c_1^G - c_2^G}{\alpha} - q_1 + \left[1 - \frac{1}{\alpha}\right]\left[(c_1^G - c_2^G) - (c_1 - c_2)\right] \\
\leq& \frac{c_1^G - c_2^G}{\alpha} - q_1,
\end{aligned}$$

where the first inequality follows from $\Delta Tw(x_1, x_2) \leq c_1^G - c_2^G$, (i.e., that Condition 6 holds for Tw, as we established earlier), $D_1 Tw(x_1, x_2) \geq c_1$ since $Tw \in \mathcal{V}_b$ and $p_2 - q_1 \leq 0$ from (B.17). The second inequality follows from $(c_1^G - c_2^G) - (c_1 - c_2) \geq 0$ and $1 - \frac{1}{\alpha} < 0$.

(ii) If $p_1 > r_2$ and since $q_1 \geq p_2$ from (B.17), we have $r_1 < q_2$ using the fact that $p_1 + q_1 + r_1 = p_2 + q_2 + r_2 = 1$. Then, $p_1 q_2 > r_1 r_2$.

Case 1: If $x_1 + x_2 \leq b - 1$,

$$\begin{aligned}
F[Tw(x_1, x_2)] &= (r_1 - q_2)D_1 Tw(x_1, x_2) + (p_1 - r_2)D_2 Tw(x_1, x_2) \\
&= \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i, j|x_1, x_2)H(i, j),
\end{aligned}$$

where

$$\begin{aligned}
H(i,j) = (r_1 - q_2)\alpha\left[q_1 + r_1 D_1 w(i,j) + p_1 D_2 w(i,j)\right] \\
+ (p_1 - r_2)\alpha\left[q_2 D_1 w(i,j) + r_2 D_2 w(i,j)\right].
\end{aligned}$$

$$\begin{aligned}
\frac{H(i,j)}{\alpha} =& (r_1 - q_2)\left[q_1 + r_1 D_1 w(i,j) + p_1 D_2 w(i,j)\right] + (p_1 - r_2)\left[q_2 D_1 w(i,j) + r_2 D_2 w(i,j)\right] \\
=& -q_1(r_2+q_2) + (r_1+r_2)\left[q_1 + (r_1-q_2)D_1 w(i,j) + (p_1-r_2)D_2 w(i,j)\right] \\
&+ (p_1 q_2 - r_1 r_2)\left[D_1 w(i,j) - D_2 w(i,j)\right] \\
=& -q_1(r_2+q_2) + (r_1+r_2)[q_1 + Fw(i,j)] + (p_1 q_2 - r_1 r_2)\Delta w(i,j) \\
\leq& -q_1(r_2+q_2) + (r_1+r_2)\left(\frac{c_1^G - c_2^G}{\alpha}\right) + (p_1 q_2 - r_1 r_2)(c_1^G - c_2^G),
\end{aligned}$$

where the inequality follows from the fact that Conditions 6 and 7 hold for $w \in \mathcal{V}_2$. Then,

$$\begin{aligned}
& \alpha H(i,j) - \left[c_1^G - c_2^G - \alpha q_1\right] \\
\leq\ & \alpha q_1\Big[(1-\alpha(r_2+q_2)\Big] + \Big[\alpha(r_1+r_2) + \alpha^2(p_1 q_2 - r_1 r_2) - 1\Big](c_1^G - c_2^G) \\
=\ & \alpha q_1(1 - \alpha r_2 - \alpha q_2) - \left[1 - \alpha(r_1+r_2) + \alpha^2 r_1 r_2 - \alpha^2 p_1 q_2\right](c_1^G - c_2^G) \\
=\ & \alpha q_1(1 - \alpha r_2 - \alpha q_2) - \left[(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2\right](c_1^G - c_2^G) \\
=\ & \left[(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2\right]\left[\frac{\alpha q_1(1-\alpha r_2 - \alpha q_2)}{(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2} - (c_1^G - c_2^G)\right] \\
=\ & \left[(1-\alpha r_1)(1-\alpha r_2) - \alpha^2 p_1 q_2\right]\left[(c_1 - c_2) - (c_1^G - c_2^G)\right] \leq 0,
\end{aligned}$$

since $(1-\alpha r_1)(1-\alpha r_2)) - \alpha^2 p_1 q_2 = (1-\alpha+\alpha q_1+\alpha p_1)(1-\alpha+\alpha p_2+\alpha q_2) - \alpha^2 p_1 q_2 \geq 0$ and $c_1 - c_2 \leq c_1^G - c_2^G$. Hence, $H(i,j) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1$, and

$$F[Tw(x_1,x_2)] = \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2) H(i,j) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1.$$

Cases 2 and 4: If $x_1 + x_2 = b$ and $x_2 > 0$, we have, $D_1[Tw(x_1,x_2)] = c_2^G + \Delta[Tw(x_1,x_2-1)]$ and $D_2[Tw(x_1,x_2)] = c_2^G$, and hence,

$$\begin{aligned}
F[Tw(x_1,x_2)] &= (r_1 - q_2) D_1[Tw(x_1,x_2)] + (p_1 - r_2) D_2[Tw(x_1,x_2)] \\
&= (r_1 - q_2)\left[c_2^G + \Delta[Tw(x_1,x_2-1)]\right] + (p_1 - r_2) c_2^G \\
&= (r_1 - q_2)\Delta[Tw(x_1,x_2-1)] + (p_2 - q_1) c_2^G \\
&\leq (r_1 - q_2)\Delta[Tw(x_1,x_2-1)] + (p_2 - q_1) D_2[Tw(x_1,x_2-1)] \\
&= (r_1 - q_2) D_1[Tw(x_1,x_2-1)] + (p_1 - r_2) D_2[Tw(x_1,x_2-1)] \\
&= F[Tw(x_1,x_2-1)] \leq \frac{c_1^G - c_2^G}{\alpha} - q_1,
\end{aligned}$$

where the first inequality follows from $D_2[Tw(x_1,x_2-1)] \leq c_2^G$ since $Tw \in \mathcal{V}_b$ and $q_1 \geq p_2$, and the second inequality follows from Case 1.

Case 3: If $x_1 = b$ and $x_2 = 0$, we have $D_1[Tw(x_1, x_2)] = c_1^G$ and $D_2[Tw(x_1, x_2)] = c_2^G$. Then,

$$\begin{aligned} F[Tw(x_1, x_2)] &= (r_1 - q_2)D_1 Tw(x_1, x_2) + (p_1 - r_2)D_2 Tw(x_1, x_2) \\ &= (r_1 - q_2)c_1^G + (p_1 - r_2)c_2^G = (r_2 - p_1)(c_1^G - c_2^G) + (p_2 - q_1)c_1^G \\ &\leq (r_2 - p_1)(c_1 - c_2) + (p_2 - q_1)c_1 = \frac{c_1 - c_2}{\alpha} - q_1 \leq \frac{c_1^G - c_2^G}{\alpha} - q_1, \end{aligned}$$

where the first inequality follows from $r_2 - p_1 < 0$, $c_1^G - c_2^G \geq c_1 - c_2$, $p_2 - q_1 \leq 0$, and $c_1^G \geq c_1$.

Thus, Conditions 6 and 7 hold for Tw, and hence $Tw \in \mathcal{V}_2$.

(c) From Theorem 11.5 of (Porteus 2002), we need to verify three conditions:

(i) Completeness. The proof is very similar to that of Lemma B.6 and thus is skipped.

(ii) Attainment. For any function $w \in \mathcal{V}_2$, we define a decision rule, which, in state (x_1, x_2), discharges no patient if $x_1 + x_2 \leq b$, discharges a stage 1 patient if $x_1 = b + 1, x_2 = 0$, and discharges a stage 2 patient if $x_1 + x_2 = b + 1$ and $x_2 > 0$. This rule attains the minimum in Tw from Definition 3.2.

(iii) Preservation. This follows immediately from part (b).

Then, from Theorem 11.5 of (Porteus 2002), we can conclude that the optimal value function $v_\alpha \in \mathcal{V}_2$.

□

Proof of Proposition 3.4: (a) Suppose that (3.10) and (3.17) hold, then from Lemma B.7(c), $v_\alpha \in \mathcal{V}_1$. It then follows from Lemma B.7(a) that, for any $x_1 + x_2 \leq b - 1$,

$$\Delta\Gamma v_\alpha(x_1, x_2) > \frac{c_1^G - c_2^G}{\alpha} - q_1.$$

Hence, for all $x_1 + x_2 = b + 1$ and $x_1 > 0, x_2 > 0$, we have $\delta_{b-1}(x_1 - 1) = \Delta\Gamma v_\alpha(x_1 - 1, b - x_1) > \frac{c_1^G - c_2^G}{\alpha} - q_1$, and thus $a_\alpha^*(x_1, x_2) = (1, 0)$ from (B.14).

(b) Suppose that (3.10) and (3.18) hold, then following Lemma B.8(c), the optimal value function $v_\alpha \in \mathcal{V}_2$ and then we have from Lemma B.8(a), for all $x_1 + x_2 \leq b - 1$

$$\Delta\Gamma v_\alpha(x_1, x_2) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1.$$

Hence, for all $x_1 + x_2 = b + 1$ and $x_1 > 0, x_2 > 0$, we have $\delta_{b-1}(x_1 - 1) = \Delta\Gamma v_\alpha(x_1 - 1, b - x_1) \leq \frac{c_1^G - c_2^G}{\alpha} - q_1$, and thus $a_\alpha^*(x_1, x_2) = (0, 1)$ from (B.14).

□

B.2.4 Proof of Theorem 3.1

Lemma B.9. *If $\beta_i < \beta_i^G$ for both $i = 1, 2$, then there exists an $\alpha_0 \in (0, 1)$ such that (3.10) holds for all $\alpha \in [\alpha_0, 1]$.*

Proof of Lemma B.9: Let $f_1(\alpha) = \alpha\big[q_1 + r_1 c_1^G + p_1 c_2^G\big] - c_1^G$ and $f_2(\alpha) = \alpha(q_2 c_1^G + r_2 c_2^G) - c_2^G$ for $\alpha \in [0, 1]$, where c_1^G and c_2^G as expressed in (3.7) are continuous in α. Then $f_1(\alpha)$ and $f_2(\alpha)$ are both continuous in α. When $\alpha = 1$, we have

$$c_1^G = \frac{q_1^G(p_2^G + q_2^G)}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G},\ c_2^G = \frac{q_1^G q_2^G}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G}.$$

Then,

$$\begin{aligned} f_1(1) &= q_1 + r_1 c_1^G + p_1 c_2^G - c_1^G = q_1(1 - c_1^G) - p_1(c_1^G - c_2^G) \\ &= \frac{q_1 p_1^G p_2^G}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} - \frac{p_1 q_1^G p_2^G}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} = \frac{p_1 p_1^G p_2^G(\beta_1 - \beta_1^G)}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G}, \end{aligned}$$

and

$$\begin{aligned} f_2(1) &= q_2 c_1^G + r_2 c_2^G - c_2^G = q_2(c_1^G - c_2^G) - p_2 c_2^G \\ &= \frac{q_2 q_1^G p_2^G}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} - \frac{p_2 q_1^G q_2^G}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} = \frac{q_2 p_1^G p_2^G\left(\beta_2^G - \beta_2\right)}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} \end{aligned}$$

When $\alpha = 1$, $f_1(\alpha)$ and $f_2(\alpha)$ are strictly negative since $\beta_i < \beta_i^G$ for $i = 1, 2$, and they are continuous in α, then there must exist some $\alpha_1 \in (0, 1)$ and $\alpha_2 \in (0, 1)$, such that $f_1(\alpha) \le 0$ for all $\alpha \in [\alpha_1, 1]$ and $f_2(\alpha) \le 0$ for all $\alpha \in [\alpha_2, 1]$. Let $\alpha_0 = \max\{\alpha_1, \alpha_2\}$. Then, if $\beta_i < \beta_i^G$, (3.10) holds for all $\alpha \in [\alpha_0, 1]$.

□

Proof of Theorem 3.1: The MDP model we introduced in Section 3.3 has finite state space and every stationary policy induces a unichain. Thus, we know from Proposition 6.4.1 of (Sennott 1999) that $J(i) \equiv J$ for $i \in S$.

Let $h_\alpha(x_1, x_2) = v_\alpha(x_1, x_2) - v_\alpha(0, 0)$ for any $(x_1, x_2) \in S$. We have shown in Lemma B.4(c) that $D_i v_\alpha(x_1, x_2) \le c_i^G$. Then,

$$\begin{aligned} h_\alpha(x_1, x_2) = v_\alpha(x_1, x_2) - v_\alpha(0, 0) = \sum_{i=0}^{x_1 - 1} D_1 v_\alpha(i, 0) + \sum_{j=0}^{x_2 - 1} D_2 v_\alpha(x_1, j) \\ \le x_1 c_1^G + x_2 c_2^G \le (b + 1)(c_1^G + c_2^G). \end{aligned}$$

It follows from Theorem 6.4.2 of (Sennott 1999) that

$$h(x_1, x_2) = \lim_{\alpha \to 1^-} h_\alpha(x_1, x_2) = \lim_{\alpha \to 1^-} [v_\alpha(x_1, x_2) - v_\alpha(0, 0)],$$

where $h(\cdot)$ is the bias function as defined in (3.3). Using (3.9), the average cost optimality equation (3.3) can be rewritten as

$$h(x_1, x_2) + g = \min_{(a_1, a_2) \in \mathcal{A}(x_1, x_2)} \Big\{ a_1 \phi_1^G + a_2 \phi_2^G + q_1(x_1 - a_1) + \Gamma h(x_1 - a_1, x_2 - a_2) \Big\}. \tag{B.18}$$

For $x + y \leq b$,

$$\begin{aligned}
\Gamma h(x, y) &= \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i, j | x, y) \lim_{\alpha \to 1^-} h_\alpha(i, j) \\
&= \lim_{\alpha \to 1^-} \left[\sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i, j | x, y) h_\alpha(i, j) \right] \\
&= \lim_{\alpha \to 1^-} \left[\sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i, j | x, y) v_\alpha(i, j) - \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i, j | x, y) v_\alpha(0, 0) \right] \\
&= \lim_{\alpha \to 1^-} \big[\Gamma v_\alpha(x, y) - v_\alpha(0, 0) \big].
\end{aligned}$$

Then, we have for $x + y \leq b - 1$,

$$\begin{aligned}
D_1 \Gamma h(x, y) &= \Gamma h(x + 1, y) - \Gamma h(x, y) \\
&= \lim_{\alpha \to 1^-} \big[\Gamma v_\alpha(x + 1, y) - v_\alpha(0, 0) \big] - \lim_{\alpha \to 1^-} \big[\Gamma v_\alpha(x, y) - v_\alpha(0, 0) \big] \\
&= \lim_{\alpha \to 1^-} \big[\Gamma v_\alpha(x + 1, y) - \Gamma v_\alpha(x, y) \big] = \lim_{\alpha \to 1^-} D_1 \Gamma v_\alpha(x, y), \\
D_2 \Gamma h(x, y) &= \Gamma h(x, y + 1) - \Gamma h(x, y) \\
&= \lim_{\alpha \to 1^-} \big[\Gamma v_\alpha(x, y + 1) - \Gamma v_\alpha(x, y) \big] = \lim_{\alpha \to 1^-} D_2 \Gamma v_\alpha(x, y),
\end{aligned}$$

If (3.10) holds, then from (B.4) and (B.5), for any $x + y \leq b - 1$,

$$D_1 \Gamma v_\alpha(x, y) \leq \frac{c_1^G}{\alpha} - q_1, \quad D_2 \Gamma v_\alpha(x, y) \leq \frac{c_2^G}{\alpha}.$$

Then, since $\beta_i < \beta_i^G$ for both $i = 1, 2$, we know from Lemma B.9 that there exists α_0 such that for $\alpha \in [\alpha_0, 1]$, (3.10) holds and thus for such α we have,

$$D_1 \Gamma h(x,y) = \lim_{\alpha \to 1^-} D_1 \Gamma v_\alpha(x,y) \leq \lim_{\alpha \to 1^-} \frac{c_1^G}{\alpha} - q_1 = \phi_1^G - q_1, \tag{B.19}$$

$$D_2 \Gamma h(x,y) = \lim_{\alpha \to 1^-} D_2 \Gamma v_\alpha(x,y) \leq \lim_{\alpha \to 1^-} \frac{c_2^G}{\alpha} = \phi_2^G. \tag{B.20}$$

where the last equalities follow from (3.1), (3.7), and

$$\begin{aligned}
\lim_{\alpha \to 1^-} c_1^G &= \lim_{\alpha \to 1^-} \frac{\alpha q_1^G (1 - \alpha r_2^G)}{(1 - \alpha r_1^G)(1 - \alpha r_2^G) - \alpha^2 p_1^G q_2^G} = \frac{q_1^G (p_2^G + q_2^G)}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} = \phi_1^G, \\
\lim_{\alpha \to 1^-} c_2^G &= \lim_{\alpha \to 1^-} \frac{\alpha^2 q_1^G q_2^G}{(1 - \alpha r_1^G)(1 - \alpha r_2^G) - \alpha^2 p_1^G q_2^G} = \frac{q_1^G q_2^G}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} = \phi_2^G.
\end{aligned}$$

If (a_1, a_2) and $(a_1 + 1, a_2)$ are both feasible actions in state (x_1, x_2), then (B.19) implies that

$$\begin{aligned}
&(a_1 + 1)\phi_1^G + a_2 \phi_2^G + q_1(x_1 - a_1 - 1) + \Gamma h(x_1 - a_1 - 1, x_2 - a_2) \\
&\geq a_1 \phi_1^G + a_2 \phi_2^G + q_1(x_1 - a_1) + \Gamma h(x_1 - a_1, x_2 - a_2),
\end{aligned}$$

which means the cost does not increase if we discharge $a_1 + 1$ type 1 patients as opposed to a_1 type 1 patients.

Similarly, if (a_1, a_2) and $(a_1, a_2 + 1)$ are both feasible actions in state (x_1, x_2), then (B.20) implies

$$\begin{aligned}
&a_1 \phi_1^G + (a_2 + 1)\phi_2^G + q_1(x_1 - a_1) + \Gamma h(x_1 - a_1, x_2 - a_2 - 1) \\
&\geq a_1 \phi_1^G + a_2 \phi_2^G + q_1(x_1 - a_1) + \Gamma h(x_1 - a_1, x_2 - a_2),
\end{aligned}$$

which means that the cost does not increase if we discharge $a_2 + 1$ type 2 patients as opposed to a_2 type 2 patients.

Hence, the result follows.

□

B.2.5 Proof of Theorem 3.2

From Theorem 3.1, we know that there exists an optimal policy which is non-idling when $\beta_i < \beta_i^G$. Thus, we can restrict ourselves to the set of policies which are non-idling. Then, we can rewrite the optimality equations (B.18) as

(i) if $x_1 + x_2 \leq b - 1$,

$$h(x_1, x_2) + g \quad = \quad q_1 x_1 + \Gamma h(x_1, x_2),$$

(ii) if $x_1 = b + 1$ and $x_2 = 0$,

$$h(x_1, x_2) + g \quad = \quad \phi_1^G + q_1(x_1 - 1) + \Gamma h(x_1 - 1, x_2),$$

(iii) if $x_1 = 0$ and $x_2 = b + 1$,

$$h(x_1, x_2) + g \quad = \quad \phi_2^G + q_1 x_1 + \Gamma h(x_1, x_2 - 1),$$

(iv) if $x_1 + x_2 = b + 1$ and $x_1 > 0, x_2 > 0$,

$$h(x_1, x_2) \; + \; g \quad = \quad \min\Big\{\phi_1^G \; + \; q_1(x_1 \; - \; 1) \; + \; \Gamma h(x_1 \; - \; 1, x_2), \;\; \phi_2^G \; + \; q_1 x_1 \; + \; \Gamma h(x_1, x_2 \; - \; 1)\Big\}.$$

Let $\bar{\delta}_n(x_1) = \Delta\Gamma h(x_1, n - x_1)$ for $0 \leq x_1 \leq n \leq b - 1$. Then, for state (x_1, x_2) where $x_1 + x_2 = b + 1$ and $x_1 > 0, x_2 > 0$, we can conclude that $a^*(x_1, x_2) = (1, 0)$ if and only if

$$\bar{\delta}_{b-1}(x_1 - 1) = \Delta\Gamma h(x_1 - 1, x_2 - 1) = \Gamma h(x_1, x_2 - 1) - \Gamma h(x_1 - 1, x_2) > \phi_1^G - \phi_2^G - q_1. \tag{B.21}$$

As in the proof of Theorem 3.1, using Theorem 6.4.2 of (Sennott 1999) we can write, for $x, y \geq 0, x+y \leq b-1$,

$$\begin{aligned}\Delta\Gamma h(x, y) &= \Gamma h(x + 1, y) - \Gamma h(x, y + 1) \\ &= \lim_{\alpha \to 1^-} \big[\Gamma v_\alpha(x + 1, y) - \Gamma v_\alpha(x, y + 1)\big] = \lim_{\alpha \to 1^-} \Delta\Gamma v_\alpha(x, y),\end{aligned}$$

Since $\beta_i < \beta_i^G$ for $i = 1, 2$, we know from Lemma (B.9) that there exists $\alpha_0 < 1$ such that (3.10) holds for $\alpha \in [\alpha_0, 1]$. As in the proof or Proposition 3.3, we can also conclude from Lemma B.6 that for $x + y \leq b - 1$ and $x > 0$, $\Delta\Gamma v_\alpha(x, y) \geq \Delta\Gamma v_\alpha(x - 1, y + 1)$ for $\alpha \in [\alpha_0, 1]$. Then, it follows that for $n = x + y$,

$$\bar{\delta}_n(x) = \Delta\Gamma h(x, y) \geq \Delta\Gamma h(x - 1, y + 1) = \bar{\delta}_n(x - 1). \tag{B.22}$$

Thus, for fixed n, $\bar{\delta}_n(x)$ is non-decreasing in x for $0 \leq x \leq n$.

First, suppose that there exists $x_1 \in [1, b]$ such that $\bar{\delta}_{b-1}(x_1 - 1) = \Delta\Gamma h(x_1 - 1, x_2 - 1) > \phi_1^G - \phi_2^G - q_1$, and let

$$x^* = \min\left\{x_1 : 1 \le x_1 \le b \text{ and } \bar{\delta}_{b-1}(x_1 - 1) > \phi_1^G - \phi_2^G - q_1\right\}.$$

Then, from (B.22), we have $\bar{\delta}_{b-1}(x_1 - 1) > \frac{c_1^G - c_2^G}{\alpha} - q_1$ for all $x_1 \in [x^*, b]$, and $\bar{\delta}_{b-1}(x_1 - 1) \le \frac{c_1^G - c_2^G}{\alpha} - q_1$ for all $x_1 \in [1, x^*)$.

Thus, from (B.21) we have $a^*(x_1, x_2) = (1, 0)$ for $x_1 + x_2 = b + 1$ and $x_1 \ge x^*$, and $a^*(x_1, x_2) = (0, 1)$ for $x_1 + x_2 = b + 1$ and $x_1 < x^*$.

Now suppose that $\bar{\delta}_{b-1}(x_1 - 1) = \Delta\Gamma h(x_1 - 1, x_2 - 1) < \phi_1^G - \phi_2^G - q_1$ for all $x_1 + x_2 = b + 1$ and $x_1 > 0, x_2 > 0$, then let $x^* = b + 1$ and the result follows.

□

B.2.6 Proof of Theorem 3.3

For fixed $k \in \{1, 2\}$ and $\alpha \in (0, 1)$, let

$$\tilde{f}_k(\alpha) = \left(c_k^G - c_k\right) - \left(c_{3-k}^G - c_{3-k}\right).$$

From (3.7) and (3.16), c_k^G, c_k, c_{3-k}^G and c_{3-k} are all continuous functions of α, and when $\alpha = 1$ by comparing (3.7) and (3.16) with (3.1) and (3.2) we have,

$$c_k^G = \phi_k^G,\ c_k = \phi_k,\ c_{3-k}^G = \phi_{3-k}^G,\ c_{3-k} = \phi_{3-k}^G.$$

Then, $\tilde{f}_k(\alpha)$ is a continuous function of $\alpha \in [0, 1]$ and $\tilde{f}_k(1) = \left(\phi_k^G - \phi_k\right) - \left(\phi_{3-k}^G - \phi_{3-k}\right)$.

If $\phi_k^G - \phi_k < \phi_{3-k}^G - \phi_{3-k}$, $\tilde{f}_k(1)$ is negative. Then, there must exist $\alpha_0' \in (0, 1)$ such that for any $\alpha \in [\alpha_0', 1]$, $\tilde{f}_k(\alpha)$ is negative, which is equivalent to $c_k^G - c_k < c_{3-k}^G - c_{3-k}$ for such α. Furthermore, according to Lemma B.9, if $\beta_i < \beta_i^G$ for $i = 1, 2$, then there exists $\alpha_0 \in (0, 1)$ such that (3.10) holds for all $\alpha \in [\alpha_0, 1]$.

Let $\bar{\alpha} = \max\{\alpha_0, \alpha_0'\}$. Then, if $\beta_1 < \beta_1^G$ and $\beta_2 < \beta_2^G$, and (3.19) holds for $k = 1$, then (3.10) and (3.17) hold for all $\alpha \in [\bar{\alpha}, 1]$. Then, for all x_1, x_2 such that $x_1 > 0$, $x_2 > 0$ and $x_1 + x_2 = b + 1$,

$$\Delta\Gamma h(x_1 - 1, x_2 - 1) = \lim_{\alpha \to 1^-} \Delta\Gamma v_\alpha(x_1 - 1, x_2 - 1) > \lim_{\alpha \to 1^-} \frac{c_1^G - c_2^G}{\alpha} - q_1 = \phi_1^G - \phi_2^G - q_1,$$

where the inequality follows from $\Delta\Gamma v_\alpha(x_1 - 1, x_2 - 1) > \frac{c_1^G - c_2^G}{\alpha}$, which has been established in the proof of Proposition 3.4 (a). Hence, $a^*(x_1, x_2) = (1, 0)$ according to (B.21) for all x_1, x_2 such that $x_1 > 0$, $x_2 > 0$ and $x_1 + x_2 = b + 1$.

Similarly if $\beta_1 < \beta_1^G$ and $\beta_2 < \beta_2^G$, and (3.19) holds for $k = 2$, then (3.10) and (3.18) hold for all $\alpha \in [\bar{\alpha}, 1]$. Then, for all x_1, x_2 such that $x_1 > 0$, $x_2 > 0$ and $x_1 + x_2 = b + 1$,

$$\Delta\Gamma h(x_1 - 1, x_2 - 1) = \lim_{\alpha \to 1^-} \Delta\Gamma v_\alpha(x_1 - 1, x_2 - 1) \leq \lim_{\alpha \to 1^-} \frac{c_1^G - c_2^G}{\alpha} - q_1 = \phi_1^G - \phi_2^G - q_1,$$

where the inequality follows from $\Delta\Gamma v_\alpha(x_1 - 1, x_2 - 1) \leq \frac{c_1^G - c_2^G}{\alpha}$, which has been established in the proof of Proposition 3.4 (b). Hence, $a^*(x_1, x_2) = (0, 1)$ according to (B.21) for all x_1, x_2 such that $x_1 > 0$, $x_2 > 0$ and $x_1 + x_2 = b + 1$.

□

APPENDIX C: PROOFS OF RESULTS IN CHAPTER 4

C.1 Proof of results in Section 4.4

Proof of Lemma 4.4. Let $P^{\pi_i} = \left[p_{j,k}^{\pi_i}\right]_{2\times 2}$, where $p_{j,k}^{\pi_i}$ is the probability that this customer will transit from stage j to stage k under policy π_i. Then,

$$P^{\pi_0} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} P^{\pi_1} = \begin{bmatrix} q_{11} & q_{12} \\ p_{21} & p_{22} \end{bmatrix} \text{ and } P^{\pi_2} = \begin{bmatrix} p_{11} & p_{12} \\ q_{21} & q_{22} \end{bmatrix}.$$

Then, $\mathbf{r}^{\pi_i}$ can be obtained by solving the following linear equations, which come from the first step analysis:

$$\mathbf{r}^{\pi_i} = \left(\begin{bmatrix} p_{10} r_1 \\ p_{20} r_2 \end{bmatrix} - p_{i0} r_i \mathbf{e_i} \right) + \alpha P^{\pi_i} \mathbf{r}^{\pi_i},$$

which gives us (I is a 2×2 identity matrix)

$$\mathbf{r}^{\pi_i} = (I - \alpha P^{\pi_i})^{-1} \left(\begin{bmatrix} p_{10} r_1 \\ p_{20} r_2 \end{bmatrix} - p_{i0} r_i \mathbf{e_i} \right). \tag{C.1}$$

We have,

$$I - \alpha P^{\pi_i} = \begin{bmatrix} A^{\pi_i} & B^{\pi_i} \\ C^{\pi_i} & D^{\pi_i} \end{bmatrix} \Rightarrow (I - \alpha P^{\pi_i})^{-1} = \frac{1}{A^{\pi_i} D^{\pi_i} - B^{\pi_i} C^{\pi_i}} \begin{bmatrix} D^{\pi_i} & -B^{\pi_i} \\ -C^{\pi_i} & A^{\pi_i} \end{bmatrix},$$

where we have

$$A^{\pi_0} = A^{\pi_2} = 1 - \alpha p_{11},\ B^{\pi_0} = B^{\pi_2} = -\alpha p_{12},\ C^{\pi_0} = C^{\pi_1} = -\alpha p_{21},\ D^{\pi_0} = D^{\pi_1} = 1 - \alpha p_{22},$$

and $A^{\pi_1} = 1 - \alpha q_{11},\ B^{\pi_1} = -\alpha q_{12}, C^{\pi_2} = -\alpha q_{21},\ D^{\pi_2} = 1 - \alpha q_{22}$. For notation simplicity, we use $A,\ B,\ C,\ D$ to denote the respective quantities related to p_{ij}, and $A',\ B',\ C'\ D'$ to denote the respective quantities related to q_{ij}. Then, we have

$$\mathbf{r}^{\pi_0} = \frac{1}{AD - BC} \begin{bmatrix} D & -B \\ -C & A \end{bmatrix} \begin{bmatrix} p_{10}R_1 \\ p_{20}R_2 \end{bmatrix},$$
$$\mathbf{r}^{\pi_1} = \frac{1}{A'D - B'C} \begin{bmatrix} D & -B' \\ -C & A' \end{bmatrix} \begin{bmatrix} 0 \\ p_{20}R_2 \end{bmatrix},$$
$$\mathbf{r}^{\pi_2} = \frac{1}{AD' - BC'} \begin{bmatrix} D' & -B \\ -C' & A \end{bmatrix} \begin{bmatrix} p_{10}R_1 \\ 0 \end{bmatrix}.$$

Take the difference of $\mathbf{r}^{\pi_0}$ and $\mathbf{r}^{\pi_i}$ for $i \in \{1, 2\}$, we have

$$\begin{aligned}
\mathbf{r}^{\pi_0} - \mathbf{r}^{\pi_1} =& \frac{1}{AD - BC} \begin{bmatrix} D & -B \\ -C & A \end{bmatrix} \begin{bmatrix} p_{10}R_1 \\ p_{20}R_2 \end{bmatrix} - \frac{1}{A'D - B'C} \begin{bmatrix} D & -B' \\ -C & A' \end{bmatrix} \begin{bmatrix} 0 \\ p_{20}R_2 \end{bmatrix}, \\
=& \frac{p_{10}R_1}{AD - BC} \begin{bmatrix} D \\ -C \end{bmatrix} + \frac{p_{20}R_2(AB' - A'B)}{(AD - BC)(A'D - B'C)} \begin{bmatrix} D \\ -C \end{bmatrix} \\
=& \frac{p_{10}R_1(A'D - B'C) + p_{20}R_2(AB' - A'B)}{(AD - BC)(A'D - B'C)} \begin{bmatrix} D \\ -C \end{bmatrix}.
\end{aligned}$$

We have $D = 1 - \alpha p_{22} > 0$, $C = -\alpha p_{21} \leq 0$ and $AD - BC > 0, A'D - B'C > 0$. Hence, $\mathbf{r}^{\pi_0} \geq \mathbf{r}^{\pi_1}$ if and only if

$$p_{10}R_1(A'D - B'C) + p_{20}R_2(AB' - A'B) \geq 0. \tag{C.2}$$

Plugging the expressions for A, B, C, D and A', B', C', D', we get

$$p_{10}R_1((1 - \alpha q_{11})(1 - \alpha p_{22})\alpha^2 q_{12}p_{21}) \geq p_{20}R_2(\alpha q_{12}(1 - \alpha p_{11}) - \alpha p_{12}(1 - \alpha q_{11})).$$

Similarly, we compare $\mathbf{r}^{\pi_0}$ and $\mathbf{r}^{\pi_2}$,

$$\begin{aligned}\mathbf{r}^{\pi_0} - \mathbf{r}^{\pi_2} =& \frac{1}{AD-BC}\begin{bmatrix} D & -B \\ -C & A \end{bmatrix}\begin{bmatrix} p_{10}R_1 \\ p_{20}R_2 \end{bmatrix} - \frac{1}{AD'-BC'}\begin{bmatrix} D' & -B \\ -C' & A \end{bmatrix}\begin{bmatrix} p_{10}R_1 \\ 0 \end{bmatrix} \\ =& \frac{p_{20}R_2}{AD-BC}\begin{bmatrix} -B \\ A \end{bmatrix} + \frac{p_{10}R_1}{(AD-BC)(AD'-BC')}\begin{bmatrix} B(-DC'+D'C) \\ A(-CD'+C'D) \end{bmatrix} \\ =& \frac{p_{10}R_1(C'D-CD')+p_{20}R_2(AD'-BC')}{(AD-BC)(AD'-BC')}\begin{bmatrix} -B \\ A \end{bmatrix}.\end{aligned}$$

Then, we have $\mathbf{r}^{\pi_0} \geq \mathbf{r}^{\pi_2}$ if and only if

$$p_{10}R_1(C'D-CD') + p_{20}R_2(AD'-BC') \geq 0, \tag{C.3}$$

which is equivalent to

$$p_{20}R_2((1-\alpha q_{22})(1-\alpha p_{11}) - \alpha^2 q_{21}p_{12}) \geq p_{10}R_1(\alpha q_{21}(1-\alpha p_{22}) - \alpha p_{21}(1-\alpha q_{22})).$$

□

Proof of Lemma 4.5. Taking the difference of $\bar{\mathbf{R}}$ and $\bar{\mathbf{R}}^Q$, we have,

$$\begin{aligned}\bar{\mathbf{R}} - \bar{\mathbf{R}}^Q =& \begin{bmatrix} 1-\alpha q_{11} & -\alpha q_{12} \\ -\alpha q_{21} & 1-\alpha q_{22} \end{bmatrix}\mathbf{r}^{\pi_0} = \frac{1}{AD-BC}\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}\begin{bmatrix} D & -B \\ -C & A \end{bmatrix}\begin{bmatrix} p_{10}R_1 \\ p_{20}R_2 \end{bmatrix} \\ =& \frac{1}{AD-BC}\begin{bmatrix} p_{10}R_1(A'D-B'C) + p_{20}R_2(AB'-A'B) \\ p_{10}R_1(C'D-CD') + p_{20}R_2(AD'-BC') \end{bmatrix},\end{aligned}$$

where A, B, C, D and A', B', C', D' are as defined in the proof of Lemma 4.4. Then, the result follows from (C.2) and (C.3). □

Lemma C.1. *For $i = 1, 2$, $D_i V_n(\mathbf{x}) \geq 0$ for all $n \geq 0$, and $D_i V_\alpha(\mathbf{x}) \geq 0$.*

Proof of Lemma C.1: We prove by induction. When $n = 0$, the inequality holds since $D_i V_0(\mathbf{x}) = 0$ for all $\mathbf{x}$. Suppose for some $n \geq 0$, the inequality holds, i.e., $D_i V_n(\mathbf{x}) \geq 0$ for both $i = 1, 2$.

Suppose $\mathbf{a}^*$ is an optimal action at state $\mathbf{x}$ at time $n+1$, then

$$V_{n+1}(\mathbf{x}) = T_{\mathbf{a}^*}V_n(\mathbf{x}) = R(\mathbf{x}, \mathbf{a}^*) + \alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x}, \mathbf{y})V_n(\mathbf{y}).$$

Note that $A(\mathbf{x}) \subset A(\mathbf{x}+\mathbf{e_i})$ for $i = 1, 2$. Then, $\mathbf{a}^* \in A(\mathbf{x}+\mathbf{e_i})$. Hence, from (4.5) we have,

$$V_{n+1}(\mathbf{x}+\mathbf{e_i}) \geq T_{\mathbf{a}^*} V_n(\mathbf{x}+\mathbf{e_i}) = R(\mathbf{x}+\mathbf{e_i}, \mathbf{a}^*) + \alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x}+\mathbf{e_i}, \mathbf{y}) V_n(\mathbf{y}).$$

Thus,

$$D_i V_{n+1}(\mathbf{x}) \geq R(\mathbf{x}+\mathbf{e_i}, \mathbf{a}^*) - R(\mathbf{x}, \mathbf{a}^*) + \alpha \left[\sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x}+\mathbf{e_i}, \mathbf{y}) V_n(\mathbf{y}) - \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x}, \mathbf{y}) V_n(\mathbf{y}) \right]$$

By the definition of the transition probability, we have for any $\mathbf{a} \in A(\mathbf{x})$,

$$\sum_{\mathbf{y}\in S} P_{\mathbf{a}}(\mathbf{x}+\mathbf{e_i}, \mathbf{y}) V_n(\mathbf{y}) = \sum_{\mathbf{y}\in S} P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) \sum_{j=0}^{2} q_{ij} V_n(\mathbf{y}+\mathbf{e_j}).$$

The customer combination $\mathbf{x}+\mathbf{e_i}$ can be divided into two independent groups, one with combination $\mathbf{x}$ and the other with only a stage i customer, and the action $\mathbf{a}$ is applied to the group with combination $\mathbf{x}$. Then, the above equation follows from the fact that the transitions of the two groups are independent of each other.

Besides, $R(\mathbf{x}+\mathbf{e_i}, \mathbf{a}^*) - R(\mathbf{x}, \mathbf{a}^*) = 0$. Then,

$$\begin{aligned} D_i V_{n+1}(\mathbf{x}) &\geq \alpha \left[\sum_{\mathbf{y}\in S} P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) \sum_{j=0}^{2} q_{ij} \left(V_n(\mathbf{y}+\mathbf{e_j}) - V_n(\mathbf{y}) \right) \right] \\ &= \alpha \left[\sum_{\mathbf{y}\in S} P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) \sum_{j=1}^{2} q_{ij} D_j V_n(\mathbf{y}) \right] \geq 0, \end{aligned}$$

where the last inequality follows from the induction hypothesis.

From Lemma 4.3, we have $V_\alpha(\mathbf{x}) = \lim_{n\to\infty} V_n(\mathbf{x})$, then

$$D_i V_\alpha(\mathbf{x}) = V_\alpha(\mathbf{x}+\mathbf{e_i}) - V_\alpha(\mathbf{x}) = \lim_{n\to\infty} \left[V_n(\mathbf{x}+\mathbf{e_i}) - V_n(\mathbf{x}) \right] = \lim_{n\to\infty} D_i V_n(\mathbf{x}) \geq 0.$$

□

Lemma C.2. *For $i = 1, 2$, $D_i V_n(\mathbf{x}) \leq \bar{R}_i$ for all $n \geq 0$, and $D_i V_\alpha(\mathbf{x}) \leq \bar{R}_i$.*

Proof of Lemma C.2: We prove by induction. When $n = 0$, the inequality holds since $D_i V_0(\mathbf{x}) = 0 < \bar{R}_i$ for all $\mathbf{x}$. Suppose for some $n \geq 0$, the inequality holds, i.e., $D_i V_n(\mathbf{x}) < \bar{R}_i$ for both $i = 1, 2$ and all $\mathbf{x} \in S$.

Suppose $\mathbf{a}^* = (a_1^*, a_2^*)$ is an optimal action at state $\mathbf{x}+\mathbf{e_i}$ at time $n+1$, then

$$V_{n+1}(\mathbf{x}+\mathbf{e_i}) = T_{\mathbf{a}^*} V_n(\mathbf{x}+\mathbf{e_i}) = R(\mathbf{x}+\mathbf{e_i}, \mathbf{a}^*) + \alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x}+\mathbf{e_i}, \mathbf{y}) V_n(\mathbf{y}).$$

If $a_i^* \geq 1$, then $\mathbf{a} - \mathbf{e_i} \in A(\mathbf{x})$, and we have

$$V_{n+1}(\mathbf{x}) \geq T_{\mathbf{a}^* - \mathbf{e_i}} V_n(\mathbf{x}) = R(\mathbf{x}, \mathbf{a}^* - \mathbf{e_i}) + \alpha \sum_{\mathbf{y} \in S} P_{\mathbf{a}^* - \mathbf{e_i}}(\mathbf{x}, \mathbf{y}) V_n(\mathbf{y}).$$

Then,

$$\begin{aligned} D_i V_{n+1}(\mathbf{x}) &\leq R(\mathbf{x} + \mathbf{e_i}, \mathbf{a}^*) - R(\mathbf{x}, \mathbf{a}^* - \mathbf{e_i}) \\ &\qquad + \alpha \left[\sum_{\mathbf{y} \in S} P_{\mathbf{a}^*}(\mathbf{x} + \mathbf{e_i}, \mathbf{y}) V_n(\mathbf{y}) - \sum_{\mathbf{y} \in S} P_{\mathbf{a}^* - \mathbf{e_i}}(\mathbf{x}, \mathbf{y}) V_n(\mathbf{y}) \right] \\ &= p_{i0} R_i + \alpha \left[\sum_{y \in S} P_{\mathbf{a}^* - \mathbf{e_i}}(\mathbf{x}, \mathbf{y}) \sum_{j=1}^{2} p_{ij} D_j V_n(\mathbf{y}) \right] \\ &< p_{i0} R_i + \alpha \left[\sum_{y \in S} P_{\mathbf{a}^* - \mathbf{e_i}}(\mathbf{x}, \mathbf{y}) \sum_{j=1}^{2} p_{ij} \bar{R}_j \right] = \bar{R}_i. \end{aligned}$$

If $a_i^* = 0$, then $\mathbf{a} \in A(\mathbf{x})$, and $V_{n+1}(\mathbf{x}) \geq T_{\mathbf{a}^*} V_n(\mathbf{x})$. Then,

$$\begin{aligned} D_i V_{n+1}(\mathbf{x}) &\leq T_{\mathbf{a}^*} V_n(\mathbf{x} + \mathbf{e_i}) - T_{\mathbf{a}^*} V_n(\mathbf{x}) \\ &= \alpha \left[\sum_{\mathbf{y} \in S} P_{\mathbf{a}^*}(\mathbf{x}, \mathbf{y}) \sum_{j=1}^{2} q_{ij} D_j V_n(\mathbf{y}) \right] < \alpha \sum_{j=1}^{2} q_{ij} \bar{R}_j = \bar{R}_i^Q \leq \bar{R}_i, \end{aligned}$$

where the first inequality follows from $V_{n+1}(\mathbf{x}) \geq T_{\mathbf{a}^*} V_n(\mathbf{x})$, the second from the induction hypothesis and the last from Assumption 4.2. This finishes the induction proof.

From Lemma 4.3, we have $V_\alpha(\mathbf{x}) = \lim_{n \to \infty} V_n(\mathbf{x})$, then

$$D_i V_\alpha(\mathbf{x}) = V_\alpha(\mathbf{x} + \mathbf{e_i}) - V_\alpha(\mathbf{x}) = \lim_{n \to \infty} [V_n(\mathbf{x} + \mathbf{e_i}) - V_n(\mathbf{x})] = \lim_{n \to \infty} D_i V_n(\mathbf{x}) < \bar{R}_i.$$

□

Proof of Proposition 4.1 . Suppose $\mathbf{a} + \mathbf{e_i} \in A(\mathbf{x})$, which indicates $\mathbf{x} \geq \mathbf{e_i}$ and $\mathbf{a} \in A(\mathbf{x})$, then,

$$\begin{aligned} T_{\mathbf{a}+\mathbf{e_i}} V_\alpha(\mathbf{x}) - T_{\mathbf{a}} V_\alpha(\mathbf{x}) &= p_{i0} R_i + \alpha \left(\sum_{\mathbf{y} \in S} P_{\mathbf{a}+\mathbf{e_i}}(\mathbf{x}, \mathbf{y}) V_\alpha(\mathbf{y}) - \sum_{\mathbf{y} \in S} P_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) V_\alpha(\mathbf{y}) \right) \\ &= p_{i0} R_i + \alpha \sum_{\mathbf{y} \in S} P_{\mathbf{a}}(\mathbf{x} - \mathbf{e_i}, \mathbf{y}) \sum_{j=1}^{2} (p_{ij} - q_{ij}) D_j V_\alpha(\mathbf{y}). \end{aligned} \tag{C.4}$$

(a) If $p_{ij} \geq q_{ij}$ for $j = 1, 2$, then

$$T_{\mathbf{a}+\mathbf{e_i}} V_\alpha(\mathbf{x}) - T_{\mathbf{a}} V_\alpha(\mathbf{x}) \geq p_{i0} R_i > 0,$$

since $D_j V_\alpha(\mathbf{y}) \geq 0$ from Lemma C.1.

(b) If $p_{ij} < q_{ij}$ for $j = 1, 2$, then

$$T_{\mathbf{a}+\mathbf{e_i}} V_\alpha(\mathbf{x}) - T_{\mathbf{a}} V_\alpha(\mathbf{x}) > p_{i0} R_i + \alpha \sum_{\mathbf{y} \in S} P_{\mathbf{a}}(\mathbf{x} - \mathbf{e_i}, \mathbf{y}) \sum_{j=1}^{2} (p_{ij} - q_{ij}) \bar{R}_j = \bar{R}_i - \bar{R}_i^Q \geq 0,$$

which follows from Lemma C.1 and Assumption 4.2.

(c) If $p_{i1} \geq q_{i1}$, $p_{i2} < q_{i2}$ and $\bar{R}_i - \bar{R}_i^Q \geq \alpha(p_{i1} - q_{i1})\bar{R}_1$, then

$$T_{\mathbf{a}+\mathbf{e_i}} V_\alpha(\mathbf{x}) - T_{\mathbf{a}} V_\alpha(\mathbf{x}) > p_{i0} R_i + \alpha(p_{i2} - q_{i2})\bar{R}_2 = \bar{R}_i - \bar{R}_i^Q - \alpha(p_{i1} - q_{i1})\bar{R}_1 \geq 0.$$

(d) If $p_{i1} < q_{i1}$, $p_{i2} \geq q_{i2}$ and $\bar{R}_i - \bar{R}_i^Q \geq \alpha(p_{i2} - q_{i2})\bar{R}_2$, then

$$T_{\mathbf{a}+\mathbf{e_i}} V_\alpha(\mathbf{x}) - T_{\mathbf{a}} V_\alpha(\mathbf{x}) > p_{i0} R_i + \alpha(p_{i1} - q_{i1})\bar{R}_1 = \bar{R}_i - \bar{R}_i^Q - \alpha(p_{i2} - q_{i2})\bar{R}_2 \geq 0.$$

Hence, $T_{\mathbf{a}+\mathbf{e_i}} V_\alpha(\mathbf{x}) > T_{\mathbf{a}} V_\alpha(\mathbf{x})$ if Assumption 4.3(i) holds for $i \in \{1, 2\}$ and $\mathbf{a} + \mathbf{e_i} \in A(\mathbf{x})$, which means serving one more stage i customer if available is always better. Hence, serving type i customers is better than idling the servers. □

Proof of Proposition 4.2: If $\mathbf{a} + \mathbf{e_1} \in A(\mathbf{x})$ and $\mathbf{a} + \mathbf{e_2} \in A(\mathbf{x})$, then $\mathbf{x} \geq \mathbf{e_1} + \mathbf{e_2}$ and $\mathbf{a} \in A(\mathbf{x} - \mathbf{e_1} - \mathbf{e_2})$. Since all customers are changing types independently, we have

$$\begin{aligned}
&T_{\mathbf{a}+\mathbf{e_1}} V_\alpha(\mathbf{x}) - T_{\mathbf{a}+\mathbf{e_2}} V_\alpha(\mathbf{x}) \\
&= p_{10} R_1 - p_{20} R_2 + \alpha \sum_{\mathbf{y} \in S} P_{\mathbf{a}}(\mathbf{x} - \mathbf{e_1} - \mathbf{e_2}, \mathbf{y}) \left(\sum_{j=0}^{2} \sum_{k=0}^{2} (p_{1j} q_{2k} - q_{1j} p_{2k}) V_\alpha(\mathbf{y} + \mathbf{e_j} + \mathbf{e_k}) \right),
\end{aligned}$$

Since $p_{10} R_1 - p_{20} R_2 \geq 0$, it is sufficient to show that under these conditions, for any $\mathbf{y} \in S$,

$$\sum_{j=0}^{2} \sum_{k=0}^{2} (p_{1j} q_{2k} - q_{1j} p_{2k}) V_\alpha(\mathbf{y} + \mathbf{e_j} + \mathbf{e_k}) \geq 0.$$

(i) Assume $p_{1j}q_{2k} - q_{1j}p_{2k} \geq 0$, and $p_{1j} - q_{1j} + p_{10}q_{2j} - q_{10}p_{2j} \geq 0$ for all $j, k \in \{1, 2\}$.

$$\sum_{j=0}^{2}\sum_{k=0}^{2}(p_{1j}q_{2k} - q_{1j}p_{2k})V_\alpha(\mathbf{y} + \mathbf{e_j} + \mathbf{e_k}) = \sum_{j=1}^{2}\sum_{k=1}^{2}(p_{1j}q_{2k} - q_{1j}p_{2k})V_\alpha(\mathbf{y} + \mathbf{e_j} + \mathbf{e_k})$$
$$+ (p_{10}q_{20} - q_{10}p_{20})V_\alpha(\mathbf{y}) + \sum_{j=1}^{2}(p_{1j}q_{20} - q_{1j}p_{20})V_\alpha(\mathbf{y} + \mathbf{e_j})$$
$$+ \sum_{k=1}^{2}(p_{10}q_{2k} - q_{10}p_{2k})V_\alpha(\mathbf{y} + \mathbf{e_k}).$$

We have,

$$p_{10}q_{20} - q_{10}p_{20} = p_{10}(1 - \sum_{k=1}^{2} q_{2k}) - q_{10}(1 - \sum_{k=1}^{2} p_{2k}) = p_{10} - q_{10} - \sum_{k=1}^{2}(p_{10}q_{2k} - q_{10}p_{2k})$$
$$=(1 - \sum_{j=1}^{2} p_{1j}) - (1 - \sum_{j=1}^{2} q_{1j}) - \sum_{k=1}^{2}(p_{10}q_{2k} - q_{10}p_{2k})$$
$$= -\sum_{j=1}^{2}(p_{1j} - q_{1j}) - \sum_{k=1}^{2}(p_{10}q_{2k} - q_{10}p_{2k}),$$

and for $j = 1, 2$,

$$p_{1j}q_{20} - q_{1j}p_{20} = p_{1j}(1 - \sum_{k=1}^{2} q_{2k}) - q_{1j}(1 - \sum_{k=1}^{2} p_{2k}) = p_{1j} - q_{1j} - \sum_{k=1}^{2}(p_{1j}q_{2k} - q_{1j}p_{2k}).$$

Then,

$$
\begin{aligned}
&\sum_{j=0}^{2}\sum_{k=0}^{2}(p_{1j}q_{2k}-q_{1j}p_{2k})V_\alpha(\mathbf{y}+\mathbf{e_j}+\mathbf{e_k})\\
=&\sum_{j=1}^{2}\sum_{k=1}^{2}(p_{1j}q_{2k}-q_{1j}p_{2k})V_\alpha(\mathbf{y}+\mathbf{e_j}+\mathbf{e_k})\\
&-\left[\sum_{j=1}^{2}(p_{1j}-q_{1j})+\sum_{k=1}^{2}(p_{10}q_{2k}-q_{10}p_{2k})\right]V_\alpha(\mathbf{y})\\
&+\sum_{j=1}^{2}\left[(p_{1j}-q_{1j})-\sum_{k=1}^{2}(p_{1j}q_{2k}-q_{1j}p_{2k})\right]V_\alpha(\mathbf{y}+\mathbf{e_j})\\
&+\sum_{k=1}^{2}(p_{10}q_{2k}-q_{10}p_{2k})V_\alpha(\mathbf{y}+\mathbf{e_k})\\
=&\sum_{j=1}^{2}\sum_{k=1}^{2}(p_{1j}q_{2k}-q_{1j}p_{2k})D_kV_\alpha(\mathbf{y}+\mathbf{e_j})\\
&+\sum_{j=1}^{2}(p_{1j}-q_{1j})D_jV_\alpha(\mathbf{y})+\sum_{k=1}^{2}(p_{10}q_{2k}-q_{10}p_{2k})D_kV_\alpha(\mathbf{y}).
\end{aligned}
$$

From the conditions in this part, and the fact that $D_jV_\alpha(\mathbf{y})\geq 0$ for $j\in\{1,2\}$ and $\mathbf{y}\in S$, we have $T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x})\geq T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x})$.

(ii) Assume $p_{1j} \geq q_{1j}$ and $p_{2j} \leq q_{2j}$ for all $j \in \{1,2\}$. Let $p_{1j} = q_{1j} + a_j$ and $q_{2j} = p_{2j} + b_j$ for $j = 1, 2$, where a_j, b_j are non-negative. Then, $q_{10} = p_{10} + \sum_{j=1}^{2} a_j$ and $p_{20} = q_{20} + \sum_{k=1}^{2} b_k$.

$$
\begin{aligned}
&\sum_{j=0}^{2}\sum_{k=0}^{2}(p_{1j}q_{2k} - q_{1j}p_{2k})V_\alpha(\mathbf{y}+\mathbf{e_j}+\mathbf{e_k}) \\
=&\sum_{j=1}^{2}\sum_{k=1}^{2}((a_j+q_{1j})(p_{2k}+b_k) - q_{1j}p_{2k})V_\alpha(\mathbf{y}+\mathbf{e_j}+\mathbf{e_k}) \qquad \text{(C.5)} \\
&+\left(p_{10}q_{20} - (\sum_{j=1}^{2}a_j + p_{10})(\sum_{k=1}^{2}b_k + q_{20})\right)V_\alpha(\mathbf{y}) \\
&+\sum_{j=1}^{2}\left((a_j+q_{1j})q_{20} - q_{1j}(q_{20}+\sum_{k=1}^{2}b_k)\right)V_\alpha(\mathbf{y}+\mathbf{e_j}) \qquad \text{(C.6)} \\
&+\sum_{k=1}^{2}\left(p_{10}(b_k+p_{2k}) - (\sum_{j=1}^{2}a_j + p_{10})p_{2k}\right)V_\alpha(\mathbf{y}+\mathbf{e_k}) \\
=&\sum_{j=1}^{2}\sum_{k=1}^{2}(a_jb_k + a_jp_{2k} + b_kq_{1j})V_\alpha(\mathbf{y}+\mathbf{e_j}+\mathbf{e_k}) \qquad \text{(C.7)} \\
&-\left(\sum_{j=1}^{2}a_jq_{20} + \sum_{k=1}^{2}b_kp_{10} + \sum_{j=1}^{2}\sum_{k=1}^{2}a_jb_k\right)V_\alpha(\mathbf{y}) \\
&+\sum_{j=1}^{2}\left(a_jq_{20} - \sum_{k=1}^{2}b_kq_{1j}\right)V_\alpha(\mathbf{y}+\mathbf{e_j}) + \sum_{k=1}^{2}\left(b_kp_{10} - \sum_{j=1}^{2}a_jp_{2k}\right)V_\alpha(\mathbf{y}+\mathbf{e_k}) \\
=&\sum_{j=1}^{2}\sum_{k=1}^{2}\Big[a_jb_k\left(D_jV_\alpha(\mathbf{y}+\mathbf{e_k}) + D_kV_\alpha(\mathbf{y})\right) + a_jp_{2k}D_jV_\alpha(\mathbf{y}+\mathbf{e_k}) + b_kq_{1j}D_kV_\alpha(\mathbf{y}+\mathbf{e_j})\Big] \\
&+\sum_{j=1}^{2}a_jq_{20}D_jV_\alpha(\mathbf{y}) + \sum_{k=1}^{2}b_kp_{10}D_kV_\alpha(\mathbf{y}) \geq 0 \qquad \text{(C.8)}
\end{aligned}
$$

since $D_jV_\alpha(\mathbf{x}) \geq 0$ and $D_kV_\alpha(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in S$ and $j, k \in \{1,2\}$.

Then, $T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x}) \geq T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x})$.

□

Proof of results in Section 4.6

Before proving Propositions 4.6 and 4.7, we first show the following Lemma.

Lemma C.3. *Assume $\mu_1 R_1 \geq \mu_2 R_2$ and $\beta_1 \geq \beta_2$ for Model II.*

(i) If $\mu_1 \leq \mu_2$, then for all $\mathbf{x} \in S$,

$$D_1V_\alpha(\mathbf{x}) - D_2V_\alpha(\mathbf{x}) \leq \bar{R}_1 - \bar{R}_2, \text{ and } D_1h(\mathbf{x}) - D_2h(\mathbf{x}) \leq R_1 - R_2.$$

(ii) If $\mu_1 \geq \mu_2$, then for all $\mathbf{x} \in S$

$$D_1 V_\alpha(\mathbf{x}) - D_2 V_\alpha(\mathbf{x}) \leq \frac{\mu_1 R_1 - \mu_2 R_2}{1-\alpha(1-\mu_1)}, \text{ and } D_1 h(\mathbf{x}) - D_2 h(\mathbf{x}) \leq \frac{\mu_1 R_1 - \mu_2 R_2}{\mu_1}.$$

Proof of Lemma C.3(i). For Model II, we have

$$\bar{R}_i = \frac{\mu_i R_i}{1-\alpha+\alpha\mu_i}, \text{ and } \hat{R}_i = \lim_{\alpha \to 1} \bar{R}_i = R_i.$$

Since $\mu_1 R_1 \geq \mu_2 R_2$ and $\mu_1 \leq \mu_2$, and then $R_1 \geq R_2$, and

$$\bar{R}_1 - \bar{R}_2 = \frac{(1-\alpha)(\mu_1 R_1 - \mu_2 R_2) + \alpha\mu_1\mu_2(R_1 - R_2)}{(1-\alpha+\alpha\mu_1)(1-\alpha+\alpha\mu_2)} \geq 0.$$

We establish Part (i) of Lemma C.3 by showing that $D_1 V_n(\mathbf{x}) - D_2 V_n(\mathbf{x}) \leq \bar{R}_1 - \bar{R}_2$ for all $n \geq 0$ where $V_n(\mathbf{x})$ are defined as in Lemma 4.3 with $V_0(\mathbf{x}) = 0$ for all $\mathbf{x} \in S$.

For $n = 0$, we have $D_1 V_0(\mathbf{x}) - D_2 V_0(\mathbf{x}) = 0 \leq \bar{R}_1 - \bar{R}_2$. Suppose $D_1 V_n(\mathbf{x}) - D_2 V_n(\mathbf{x}) \leq \bar{R}_1 - \bar{R}_2$ for some $n \geq 0$. Then,

$$\begin{aligned} D_1 V_{n+1}(\mathbf{x}) - D_2 V_{n+1}(\mathbf{x}) &= V_{n+1}(\mathbf{x}+\mathbf{e_1}) - V_{n+1}(\mathbf{x}+\mathbf{e_2}) \\ &= \max_{\mathbf{a}}\{T_{\mathbf{a}} V_n(\mathbf{x}+\mathbf{e_1})\} - \max_{\mathbf{a}}\{T_{\mathbf{a}} V_n(\mathbf{x}+\mathbf{e_2})\}. \end{aligned}$$

Suppose $\mathbf{a}^* = (a_1^*, a_2^*) = \arg\max_{\mathbf{a}}\{T_{\mathbf{a}} V_n(\mathbf{x}+\mathbf{e_1})\}$ is an optimal action at state $\mathbf{x}+\mathbf{e_1}$, i.e., $\max_{\mathbf{a}}\{T_{\mathbf{a}} V_{n+1}(\mathbf{x}+\mathbf{e_1})\} = T_{\mathbf{a}^*} V_n(\mathbf{x}+\mathbf{e_1})$.

If $a_1^* \leq x_1$, then $\mathbf{a}^*$ is a feasible action at state $\mathbf{x}+\mathbf{e_2}$. Then, $\max_{\mathbf{a}}\{T_{\mathbf{a}} V_{n+1}(\mathbf{x}+\mathbf{e_2})\} \geq T_{\mathbf{a}^*} V_n(\mathbf{x}+\mathbf{e_2})$, and hence,

$$D_1 V_{n+1}(\mathbf{x}) - D_2 V_{n+1}(\mathbf{x}) \leq T_{\mathbf{a}^*} V_n(\mathbf{x}+\mathbf{e_1}) - T_{\mathbf{a}^*} V_n(\mathbf{x}+\mathbf{e_2}),$$

where

$$
\begin{aligned}
&T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_1}) - T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_2}) \\
=&\alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x},\mathbf{y}) \left[\beta_1 V_n(\mathbf{y}) + (1-\beta_1)V_n(\mathbf{y}+\mathbf{e_1}) - \beta_2 V_n(\mathbf{y}) - (1-\beta_2)V_n(\mathbf{y}+\mathbf{e_2})\right] \\
=&\alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x},\mathbf{y}) \left[(1-\beta_1)D_1V_n(\mathbf{y}) - (1-\beta_2)D_2V_n(\mathbf{y})\right] \\
=&\alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x},\mathbf{y}) \left[(1-\beta_1)[D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y})] - (\beta_1-\beta_2)D_2V_n(\mathbf{y})\right] \\
\leq&\alpha(1-\beta_1)[\bar{R}_1 - \bar{R}_2] \leq \bar{R}_1 - \bar{R}_2,
\end{aligned}
$$

where the first inequality follows from $D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y}) \leq \bar{R}_1 - \bar{R}_2$, $D_2V_n(\mathbf{y}) \geq 0$ and $\beta_2 \leq \beta_1 \leq 1$, and the second inequality follows from $0 \leq \alpha, \beta_1 \leq 1$.

If $a_1^* = x_1 + 1$, then $\mathbf{a}^* - \mathbf{e_1} + \mathbf{e_2}$ is a feasible action at state $\mathbf{x} + \mathbf{e_2}$. We have

$$
D_1V_{n+1}(\mathbf{x}) - D_2V_{n+1}(\mathbf{x}) \leq T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_1}) - T_{\mathbf{a}^*-\mathbf{e_1}+\mathbf{e_2}}V_n(\mathbf{x}+\mathbf{e_2}),
$$

where

$$
\begin{aligned}
&T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_1}) - T_{\mathbf{a}^*-\mathbf{e_1}+\mathbf{e_2}}V_n(\mathbf{x}+\mathbf{e_2}) = \mu_1 R_1 - \mu_2 R_2 \\
&\quad + \alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*-\mathbf{e_1}}(\mathbf{x},\mathbf{y}) \left[\mu_1 V_n(\mathbf{y}) + (1-\mu_1)V_n(\mathbf{y}+\mathbf{e_1}) - \mu_2 V_n(\mathbf{y}) - (1-\mu_2)V_n(\mathbf{y}+\mathbf{e_2})\right] \\
=&\mu_1 R_1 - \mu_2 R_2 + \alpha \sum_{\mathbf{y}\in S} P_{\mathbf{a}^*-\mathbf{e_1}}(\mathbf{x},\mathbf{y}) \left[(1-\mu_1)D_1V_n(\mathbf{y}) - (1-\mu_2)D_2V_n(\mathbf{y})\right].
\end{aligned}
$$

We have

$$
\begin{aligned}
(1-\mu_1)D_1V_n(\mathbf{y}) - (1-\mu_2)D_2V_n(\mathbf{y}) &= (1-\mu_1)(D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y})) + (\mu_2-\mu_1)D_2V_n(\mathbf{y}) \\
&\leq (1-\mu_1)(\bar{R}_1 - \bar{R}_2) + (\mu_2-\mu_1)\bar{R}_2 = (1-\mu_1)\bar{R}_1 - (1-\mu_2)\bar{R}_2,
\end{aligned}
$$

where the inequality follows from $D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y}) \leq \bar{R}_1 - \bar{R}_2$, $D_2V_n(\mathbf{y}) \leq \bar{R}_2$ and $\mu_1 \leq \mu_2 \leq 1$. Then,

$$
\begin{aligned}
&T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_1}) - T_{\mathbf{a}^*-\mathbf{e_1}+\mathbf{e_2}}V_n(\mathbf{x}+\mathbf{e_2}) \\
&\qquad \leq \mu_1 R_1 - \mu_2 R_2 + \alpha\left[(1-\mu_1)\bar{R}_1 - (1-\mu_2)\bar{R}_2\right] = \bar{R}_1 - \bar{R}_2.
\end{aligned}
$$

Hence, $D_1V_{n+1}(\mathbf{x}) - D_2V_{n+1}(\mathbf{x}) \leq \bar{R}_1 - \bar{R}_2$ for all $n \geq 0$.

From Lemma 4.3, we have $V_\alpha(\mathbf{x}) = \lim_{n\to\infty} V_n(\mathbf{x})$, and then,

$$\begin{aligned} D_1V_\alpha(\mathbf{x}) - D_2V_\alpha(\mathbf{x}) &= V_\alpha(\mathbf{x}+\mathbf{e_1}) - V_\alpha(\mathbf{x}+\mathbf{e_2}) \\ &= \lim_{n\to\infty}\left[V_n(\mathbf{x}+\mathbf{e_1}) - V_n(\mathbf{x}+\mathbf{e_2})\right] = \lim_{n\to\infty}\left[D_1V_n(\mathbf{x}) - D_2V_n(\mathbf{x})\right] \le \bar{R}_1 - \bar{R}_2. \end{aligned}$$

From Lemma 4.2 we have $h(\mathbf{x}) = \lim_{\alpha\to 1} V_\alpha(\mathbf{x}) - V_\alpha(\mathbf{e_0})$, and then,

$$\begin{aligned} D_1h(\mathbf{x}) - D_2h(\mathbf{x}) = h(\mathbf{x}+\mathbf{e_1}) - h(\mathbf{x}+\mathbf{e_2}) &= \lim_{\alpha\to 1}\left[V_\alpha(\mathbf{x}+\mathbf{e_1}) - V_\alpha(\mathbf{x}+\mathbf{e_2})\right] \\ &= \lim_{\alpha\to 1}\left[D_1V_\alpha(\mathbf{x}) - D_2V_\alpha(\mathbf{x})\right] \le \lim_{\alpha\to 1}\left[\bar{R}_1 - \bar{R}_2\right] = R_1 - R_2. \end{aligned}$$

Proof of Lemma C.3(ii): Similar as Part (i), we first show by induction that for this case $D_1V_n(\mathbf{x}) - D_2V_n(\mathbf{x}) \le \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}$ for all $\mathbf{x}\in S$. For $n=0$, we have $D_1V_0(\mathbf{x}) - D_2V_0(\mathbf{x}) = 0 \le \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}$. Suppose $D_1V_n(\mathbf{x}) - D_2V_n(\mathbf{x}) \le \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}$ for some $n \ge 0$. Let $\mathbf{a}^* = (a_1^*, a_2^*) = \arg\max_{\mathbf{a}}\{T_{\mathbf{a}}V_n(\mathbf{x}+\mathbf{e_1})\}$. If $a_1^* \le x_1$, we have

$$\begin{aligned} D_1V_{n+1}(\mathbf{x}) - D_2V_{n+1}(\mathbf{x}) \le& T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_1}) - T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_2}), \\ =& \alpha\sum_{\mathbf{y}\in S} P_{\mathbf{a}^*}(\mathbf{x},\mathbf{y})\left[(1-\beta_1(D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y}))(\beta_1-\beta_2)D_2V_n(\mathbf{y})\right] \\ \le& \alpha(1-\beta_1)\left[\frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}\right] \le \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}, \end{aligned}$$

where the inequality follows from $D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y}) \le \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}$ from induction hypothesis.

If $a_1^* = x_1 + 1$, then

$$\begin{aligned} D_1V_{n+1}(\mathbf{x}) - D_2V_{n+1}(\mathbf{x}) &\le T_{\mathbf{a}^*}V_n(\mathbf{x}+\mathbf{e_1}) - T_{\mathbf{a}^*-\mathbf{e_1}+\mathbf{e_2}}V_n(\mathbf{x}+\mathbf{e_2}), \\ &= \mu_1R_1 - \mu_2R_2 + \alpha\sum_{\mathbf{y}\in S} P_{\mathbf{a}^*-\mathbf{e_1}}(\mathbf{x},\mathbf{y})\left[(1-\mu_1)D_1V_n(\mathbf{y}) - (1-\mu_2)D_2V_n(\mathbf{y})\right]. \end{aligned}$$

We have

$$\begin{aligned} &(1-\mu_1)D_1V_n(\mathbf{y}) - (1-\mu_2)D_2V_n(\mathbf{y}) \\ &\qquad = (1-\mu_1)(D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y})) + (\mu_2-\mu_1)D_2V_n(\mathbf{y}) \le (1-\mu_1)\left[\frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}\right], \end{aligned}$$

where the inequality follows from $D_1V_n(\mathbf{y}) - D_2V_n(\mathbf{y}) \leq \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}$ from induction hypothesis, $D_2V_n(\mathbf{y}) \geq 0$ from Lemma C.1, and $\mu_2 \leq \mu_1 \leq 1$. Then,

$$D_1V_{n+1}(\mathbf{x}) - D_2V_{n+1}(\mathbf{x}) \leq \mu_1R_1 - \mu_2R_2 + \alpha(1-\mu_1)\left[\frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}\right] = \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}.$$

Hence, $D_1V_n(\mathbf{x}) - D_2V_n(\mathbf{x}) \leq \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}$ for all $n \geq 0$.

From Lemma 4.3, we have $V_\alpha(\mathbf{x}) = \lim_{n\to\infty} V_n(\mathbf{x})$, and then,

$$D_1V_\alpha(\mathbf{x}) - D_2V_\alpha(\mathbf{x}) = \lim_{n\to\infty}[D_1V_n(\mathbf{x}) - D_2V_n(\mathbf{x})] \leq \frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}.$$

From Lemma 4.2 we have $h(\mathbf{x}) = \lim_{\alpha\to1}[V_\alpha(\mathbf{x}) - V_\alpha(\mathbf{e_0})]$, and then

$$D_1h(\mathbf{x}) - D_2h(\mathbf{x}) = \lim_{\alpha\to1}[D_1V_\alpha(\mathbf{x}) - D_2V_\alpha(\mathbf{x})] \leq \lim_{\alpha\to1}\frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)} = \frac{\mu_1R_1-\mu_2R_2}{\mu_1}.$$

□

Proof of Proposition 4.6. We assume $\mu_2 \geq \mu_1 \geq \beta_1 \geq \beta_2$ and $\beta_1R_1 \geq (\beta_1+\mu_2-\mu_1)R_2$. Then, $\mu_1R_1 - \mu_2R_2 = \frac{\mu_1(\beta_1+\mu_2-\mu_1)}{\beta_1} - \mu_2R_2 = \frac{\mu_1(\beta_1+\mu_2-\mu_1)-\mu_2\beta_1}{\beta_1}R_2 = \frac{(\mu_1-\beta_1)(\mu_2-\mu_1)}{\beta_1}R_2 \geq 0$. From Lemma C.3(i), we have $D_1V_\alpha(\mathbf{y}) - D_2V_\alpha(\mathbf{y}) \leq \bar{R}_1 - \bar{R}_2$ and $D_1h(\mathbf{x}) - D_2h(\mathbf{x}) \leq R_1 - R_2$.

If $\mathbf{a}+\mathbf{e_1} \in A(\mathbf{x})$, $\mathbf{a}+\mathbf{e_2} \in A(\mathbf{x})$, then, $\mathbf{x} \geq \mathbf{e_1}+\mathbf{e_2}$ and $\mathbf{a} \in A(\mathbf{x}-\mathbf{e_1}-\mathbf{e_2})$. Then,

$$T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x}) - T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x}) = \mu_1R_1 - \mu_2R_2 + \alpha\sum_{\mathbf{y}\in S}P_{\mathbf{a}}(\mathbf{x}-\mathbf{e_1}-\mathbf{e_2},\mathbf{y})G(\mathbf{y}),$$

where

$$\begin{aligned}
G(\mathbf{y}) =& \left[(1-\mu_1)(1-\beta_2) - (1-\mu_2)(1-\beta_1)\right]V_\alpha(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) \\
&+ \left[(1-\mu_1)\beta_2 - (1-\beta_1)\mu_2\right]V_\alpha(\mathbf{y}+e_1) \\
&+ \left[\mu_1(1-\beta_2) - \beta_1(1-\mu_2)\right]V_\alpha(\mathbf{y}+e_2) + \left[\mu_1\beta_2 - \beta_1\mu_2\right]V_\alpha(\mathbf{y}) \\
=& \left[(1-\mu_1)(1-\beta_2) - (1-\mu_2)(1-\beta_1)\right]D_2V_\alpha(\mathbf{y}+\mathbf{e_1}) \\
&+ (\beta_1-\mu_1)\left[D_1V_\alpha(\mathbf{y}) - D_2V_\alpha(\mathbf{y})\right] \\
&+ \left[\beta_1\mu_2 - \mu_1\beta_2\right]D_2V_\alpha(\mathbf{y}) \geq (\beta_1-\mu_1)(\bar{R}_1-\bar{R}_2),
\end{aligned}$$

since $(1-\mu_1)(1-\beta_2)-(1-\mu_2)(1-\beta_1)\geq 0$, $\beta_1-\mu_1\leq 0$, $\beta_1\mu_2-\mu_1\beta_2\geq 0$, $D_2V_\alpha(\mathbf{y})\geq 0$ and $D_1V_\alpha(\mathbf{y})-D_2V_\alpha(\mathbf{y})\leq \bar{R}_1-\bar{R}_2$. Hence,

$$\begin{aligned}&T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x})-T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x})\geq \mu_1R_1-\mu_2R_2+\alpha(\beta_1-\mu_1)(\bar{R}_1-\bar{R}_2)\\ =&\mu_1R_1-\mu_2R_2+\alpha(\beta_1-\mu_1)\left[\frac{(1-\alpha)(\mu_1R_1-\mu_2R_2)+\alpha\mu_1\mu_2(R_1-R_2)}{(1-\alpha+\alpha\mu_1)(1-\alpha+\alpha\mu_2)}\right]\end{aligned}$$

We write the above expression in the form of one fraction, where the denominator $(1-\alpha+\alpha\mu_1)(1-\alpha+\alpha\mu_2)$ is positive, and then we focus on the numerator, which equals to

$$\begin{aligned}&\left[(1-\alpha+\alpha\mu_1)(1-\alpha+\alpha\mu_2)+\alpha(\beta_1-\mu_1)(1-\alpha)\right](\mu_1R_1-\mu_2R_2)\\ &+\alpha^2(\beta_1-\mu_1)\mu_1\mu_2(R_1-R_2)\\ =&\left[(1-\alpha)^2+\alpha^2\mu_1\mu_2+\alpha(1-\alpha)(\mu_2+\beta_1)\right](\mu_1R_1-\mu_2R_2)\\ &+\alpha^2(\beta_1-\mu_1)\mu_1\mu_2(R_1-R_2)\\ =&\left[(1-\alpha)^2+\alpha(1-\alpha)(\mu_2+\beta_1)\right](\mu_1R_1-\mu_2R_2)+\alpha^2\mu_1\mu_2\left[\beta_1R_1-(\beta_1+\mu_2-\mu_1)R_2\right]\geq 0.\end{aligned}$$

Similarly, since $D_1h(\mathbf{y})-D_2h(\mathbf{y})\leq R_1-R_2$,

$$\begin{aligned}H_{\mathbf{a}+\mathbf{e_1}}h(\mathbf{x})-H_{\mathbf{a}+\mathbf{e_2}}h(\mathbf{x})=&\mu_1R_1-\mu_2R_2\\ &+\sum_{\mathbf{y}\in S}P_{\mathbf{a}}(\mathbf{x}-\mathbf{e_1}-\mathbf{e_2},\mathbf{y})\Big\{(\beta_1-\mu_1)\left[D_1h(\mathbf{y})-D_2h(\mathbf{y})\right]\\ &+\left[(1-\mu_1)(1-\beta_2)-(1-\mu_2)(1-\beta_1)\right]D_2h(\mathbf{y}+\mathbf{e_1})\\ &+\left[\beta_1\mu_2-\mu_1\beta_2\right]D_2h(\mathbf{y})\Big\}\\ \geq&\mu_1R_1-\mu_2R_2+(\beta_1-\mu_1)(R_1-R_2)\geq 0.\end{aligned}$$

□

Proof of Proposition 4.7. We assume $\mu_1R_1\geq\mu_2R_2$, $\frac{\beta_1}{\beta_2}\geq\frac{\mu_1}{\mu_2}\geq 1$ and $1\leq\frac{1-\beta_1}{1-\mu_1}\leq\frac{1-\beta_2}{1-\mu_2}$. Then, from Lemma C.3(ii), we have for all $\mathbf{x}\in S$,

$$D_1V_\alpha(\mathbf{x})-D_2V_\alpha(\mathbf{x})\leq\frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)},\text{ and }D_1h(\mathbf{x})-D_2h(\mathbf{x})\leq\frac{\mu_1R_1-\mu_2R_2}{\mu_1}.$$

Follow the same arguments in the proof of Proposition 4.6, we have for $\mathbf{a}+\mathbf{e_1}\in A(\mathbf{x})$, $\mathbf{a}+\mathbf{e_2}\in A(\mathbf{x})$,

$$T_{\mathbf{a}+\mathbf{e_1}}V_\alpha(\mathbf{x})-T_{\mathbf{a}+\mathbf{e_2}}V_\alpha(\mathbf{x})=\mu_1R_1-\mu_2R_2+\alpha\sum_{\mathbf{y}\in S}P_{\mathbf{a}}(\mathbf{x}-\mathbf{e_1}-\mathbf{e_2},\mathbf{y})G(\mathbf{y}),$$

where

$$\begin{aligned}G(\mathbf{y}) =& \left[(1-\mu_1)(1-\beta_2)-(1-\mu_2)(1-\beta_1)\right] D_2V_\alpha(\mathbf{y}+\mathbf{e_1})\\ &+(\beta_1-\mu_1)\left[D_1V_\alpha(\mathbf{y})-D_2V_\alpha(\mathbf{y})\right]+\left[\beta_1\mu_2-\mu_1\beta_2\right]D_2V_\alpha(\mathbf{y})\\ \geq&(\beta_1-\mu_1)\left[\frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}\right],\end{aligned}$$

and hence,

$$\begin{aligned}&T_{\mathbf{a+e_1}}V_\alpha(\mathbf{x})-T_{\mathbf{a+e_2}}V_\alpha(\mathbf{x})\geq\mu_1R_1-\mu_2R_2+\alpha(\beta_1-\mu_1)\left[\frac{\mu_1R_1-\mu_2R_2}{1-\alpha(1-\mu_1)}\right]\\ =&(\mu_1R_1-\mu_2R_2)\left[\frac{1-\alpha(1-\mu_1)+\alpha(\beta_1-\mu_1)}{1-\alpha(1-\mu_1)}\right]=(\mu_1R_1-\mu_2R_2)\left[\frac{1-\alpha(1-\beta_1)}{1-\alpha(1-\mu_1)}\right]\geq0,\end{aligned}$$

and

$$\begin{aligned}H_{\mathbf{a+e_1}}h(\mathbf{x})-H_{\mathbf{a+e_2}}h(\mathbf{x})\geq\mu_1R_1-\mu_2R_2+(\beta_1-\mu_1)\left[\frac{\mu_1R_1-\mu_2R_2}{\mu_1}\right]&\\ =\frac{\beta_1(\mu_1R_1-\mu_2R_2)}{\mu_1}\geq0.&\end{aligned}$$

□

Proof of Corollary 4.5. Let $\mu_i=\tilde{\mu}_i+\beta_i-\tilde{\mu}_i\beta_i$ and $R_i=\frac{\tilde{\mu}_i}{\mu_i}\tilde{R}_i$ for $i=1,2$. Then, we have

$$\begin{aligned}&\mu_1\beta_2-\mu_2\beta_1=\tilde{\mu}_1(1-\beta_1)\beta_2-\tilde{\mu}_2(1-\beta_2)\beta_1\leq0\Leftrightarrow\frac{\beta_1}{\beta_2}\geq\frac{\mu_1}{\mu_2},\\ &(\mu_2-\beta_1)(1-\beta_2)=\tilde{\mu}_2(1-\beta_1)(1-\beta_2)\geq\tilde{\mu}_1(1-\beta_1)(1-\beta_2)=(\mu_1-\beta_2)(1-\beta_1)\geq0,\end{aligned}$$

$$\begin{aligned}\mu_1R_1-\mu_2R_2=&\tilde{\mu}_1\tilde{R}_1-\tilde{\mu}_2\tilde{R}_2\\ \geq&\max\left\{0,\ \tilde{\mu}_1(1-\beta_1)\left[\frac{\tilde{\mu}_1\tilde{R}_1}{\tilde{\mu}_1+(1-\tilde{\mu}_1)\beta_1}-\frac{\tilde{\mu}_2\tilde{R}_2}{\tilde{\mu}_2+(1-\tilde{\mu}_2)\beta_2}\right]\right\}\\ =&\max\left\{0,\ (\mu_1-\beta_1)\left[R_1-R_2\right]\right\}.\end{aligned}$$

□

BIBLIOGRAPHY

Aféche, P., H. Mendelson. 2004. Pricing and priority auctions in queueing systems with a generalized delay cost structure. *Management Science* **50**(7) 869–882.

Akan, Mustafa, Oguzhan Alagoz, Baris Ata, Fatih Safa Erenay, Adnan Said. 2012. A broader view of designing the liver allocation system. *Operations research* **60**(4) 757–770.

Argon, Nilay Tanik, Serhan Ziya, Rhonda Righter. 2008. Scheduling impatient jobs in a clearing system with insights on patient triage in mass casualty incidents. *Probability in the Engineering and Informational Sciences* **22**(03) 301–332.

Atar, Rami, Chanit Giat, Nahum Shimkin. 2010. The cμ/θ rule for many-server queues with abandonment. *Operations Research* **58**(5) 1427–1439.

Cao, Ping, Jingui Xie. 2015. Optimal control of a multiclass queueing system when customers can change types. *Queueing Systems* 1–29.

Cardoso, L.T., C.M. Grion, T. Matsuo, E.H. Anami, I.A. Kauss, L. Seko, A. M. Bonametti. 2011. Impact of delayed admission to intensive care units on mortality of critically ill patients: a cohort study. *Critical Care* **15**(1) R28.

Chalfin, D.B., S. Trzeciak, A. Likourezos, B.M. Baumann, R.P. Dellinger, DELAY-ED study group. 2007. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit*. *Critical Care Medicine* **35**(6) 1477–1483.

Chan, C.W., V.F. Farias, N. Bambos, G.J. Escobar. 2012. Optimizing intensive care unit discharge decisions with patient readmissions. *Operations Research* **60**(6) 1323–1341.

Chan, C.W., G. Yom-Tov, G.J. Escobar. 2014. When to use speedup: An examination of service systems with returns. *Operations Research* **62**(2) 462–482.

Christian, M.D., L. Hawryluck, R.S. Wax, T. Cook, N.M. Lazar, M.S. Herridge, M.P. Muller, D.R. Gowans, W. Fortier, F.M. Burkle. 2006. Development of a triage protocol for critical care during an influenza pandemic. *Canadian Medical Association Journal* **175**(11) 1377–1381.

Cobham, A. 1954. Priority assignment in waiting line problems. *Journal of the Operations Research Society of America* **2**(1) 70–76.

Cox, D. R., W. L. Smith. 1961. *Queues*. Chapman & Hall.

Dewan, S., H. Mendelson. 1990. User delay costs and internal pricing for a service facility. *Management Science* **36**(12) 1502–1517.

Di Crescenzo, A. 1999. A probabilistic analogue of the mean value theorem and its applications to reliability theory. *Journal of Applied Probability* 706–719.

Dobson, G., H.H. Lee, E. Pinker. 2010. A model of ICU bumping. *Operations Research* **58**(6) 1564–1576.

Down, Douglas G, Ger Koole, Mark E Lewis. 2011. Dynamic control of a single-server system with abandonments. *Queueing Systems* **67**(1) 63–90.

Down, Douglas G, Mark E Lewis. 2010. The n-network model with upgrades. *Probability in the Engineering and Informational Sciences* **24**(02) 171–200.

El-Taha, M., S. Stidham. 1999. *Sample-Path Analysis of Queueing Systems*, vol. 11. Springer Science & Business Media.

Ghahramani, S., R. W. Wolff. 1989. A new proof of finite moment conditions for GI/G/1 busy periods. *Queueing Systems* **4**(2) 171–178.

Ghassemi, M., T. Naumann, F. Doshi-Velez, N. Brimmer, R. Joshi, A. Rumshisky, P. Szolovits. 2014. Unfolding physiological state: mortality modelling in intensive care units. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 75–84.

Gortzis, L.G., F. Sakellaropoulos, I. Ilias, K. Stamoulis, I. Dimopoulou. 2008. Predicting ICU survival: a meta-level approach. *BMC Health Services Research* **8**(1) 157.

Gross, D., J. F. Shortle, J. M. Thompson, C.M. Harris. 2008. *Fundamentals of Queueing Theory*. 4th ed. Chapman & Hall.

Haji, R., G. F. Newell. 1971. Optimal strategies for priority queues with nonlinear costs of delay. *SIAM Journal on Applied Mathematics* **20**(2) 224–240.

Hassin, R., J. Puerto, F. R. Fernandez. 2009. The use of relative priorities in optimizing the performance of a queueing system. *European Journal of Operational Research* **193** 476–483.

He, Qi-Ming, Jingui Xie, Xiaobo Zhao. 2012. Priority queue with customer upgrades. *Naval Research Logistics (NRL)* **59**(5) 362–375.

Jacobson, E.U. 2010. Scheduling in service systems with impatient customers and insights on mass-casualty triage. Ph.D. thesis, University of North Carolina at Chapel Hill.

Jacobson, Evin Uzun, Nilay Tanik Argon, Serhan Ziya. 2012. Priority assignment in emergency response. *Operations Research* **60**(4) 813–832.

Jaiswal, N. K. 1968. *Priority queues*. Academic Press, New York.

Kc, D.S., C. Terwiesch. 2012. An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing & Service Operations Management* **14**(1) 50–65.

Kim, S.H., C.W. Chan, M. Olivares, G. Escobar. 2014. ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science* **61**(1) 19–38.

Kulkarni, V.G. 2009. *Modeling and Analysis of Stochastic Systems*. 2nd ed. Chapman & Hall.

Kumar, A., R. Zarychanski, R. Pinto, D.J. Cook, J. Marshall, J. Lacroix, T. Stelfox, S. Bagshaw, K. Choong, F. Lamontagne. 2009. Critically ill patients with 2009 influenza a (H1N1) infection in Canada. *Jama* **302**(17) 1872–1879.

Mandelbaum, A., A. L. Stolyar. 2004. Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized $c\mu$-rule. *Operations Research* **52**(6) 836–855.

Miller, D. R. 1960. Priority queues. *The Annals of Mathematical Statistics* **31**(1) 86–103.

Moreno, R.P., P.G. Metnitz, E. Almeida, B. Jordan, P. Bauer, R.A. Campos, G. Iapichino, D. Edbrooke, M. Capuzzo, J.R. Le Gall. 2005. SAPS 3 From evaluation of the patient to evaluation of the intensive care unit. part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Medicine* **31**(10) 1345–1355.

Örmeci, E.L., A Burnetas. 2005. Dynamic admission control for loss systems with batch arrivals. *Advances in Applied Probability* **37**(4) 915–937.

Örmeci, E.L., A. Burnetas, J. van der Wal. 2001. Admission policies for a two class loss system. *Stochastic Models* **17**(4) 513–540.

Porteus, E.L. 2002. *Foundations of Stochastic Inventory Theory*. Stanford University Press.

Puterman, M.L. 2005. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.

Rocker, G., D. Cook, P. Sjokvist, B. Weaver, S. Finfer, E. McDonald, J. Marshall, A. Kirby, M. Levy, P. Dodek, et al. 2004. Clinician predictions of intensive care unit mortality. *Critical Care Medicine* **32**(5) 1149–1154.

Roman, Steven. 1980. The formula of Faa di Bruno. *American Mathematical Monthly* 805–809.

Ross, Sheldon M. 1983. *Introduction to Stochastic Dynamic Programming*. Academic press.

Rothkopf, M. H., S. A. Smith. 1984. There are no undiscovered priority index sequencing rules for minimizing total delay costs. *Operations Research* **32**(2) 451–456.

Sennott, L.I. 1999. *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley & Sons.

Shaked, M., J.G. Shanthikumar. 2007. *Stochastic Orders*. Springer.

Shmueli, A., C.L. Sprung. 2005. Assessing the in-hospital survival benefits of intensive care. *International Journal of Technology Assessment in Health Care* **21**(01) 66–72.

Shmueli, A., C.L. Sprung, E.H. Kaplan. 2003. Optimizing admissions to an intensive care unit. *Health Care Management Science* **6**(3) 131–136.

Sinuff, T., K. Kahnamoui, D.J. Cook, J.M. Luce, M.M. Levy, et al. 2004. Rationing critical care beds: A systematic review. *Critical Care Medicine* **32**(7) 1588–1597.

Strand, K., H. Flaatten. 2008. Severity scoring in the ICU: a review. *Acta Anaesthesiologica Scandinavica* **52**(4) 467–478.

Takács, Lajos. 1955. Investigation of waiting time problems by reduction to markov processes. *Acta Mathematica Hungarica* **6**(1) 101–129.

Ulukus, M.Y., R. Güllü, E.L. Örmeci. 2011. Admission and termination control of a two class loss system. *Stochastic Models* **27**(1) 2–25.

Van Mieghem, J. A. 1995. Dynamic scheduling with convex delay costs: The generalized cμ rule. *The Annals of Applied Probability* **5**(3) 809–833.

Vasicek, O. A. 1977. An inequality for the variance of waiting time under a general queuing discipline. *Operations Research* **25**(5) 879–884.

Vincent, J.L., R. Moreno, J. Takala, S. Willatts, A. De Mendonça, H. Bruining, C.K. Reinhart, P. Suter, L.G. Thijs. 1996. The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. *Intensive Care Medicine* **22**(7) 707–710.

Wagner, J., N.B. Gabler, S.J. Ratcliffe, S.E. Brown, B.L. Strom, S.D. Halpern. 2013. Outcomes among patients discharged from busy intensive care units. *Annals of Internal Medicine* **159**(7) 447–455.

Wishart, David MC. 1960. Queuing systems in which the discipline is last-come, first-served. *Operations Research* **8**(5) 591–599.

Wolff, R. W. 1989. *Stochastic modeling and the theory of queues*. Prentice Hall International Series in Industrial and Systems Engineering, Prentice Hall Inc.

Zimmerman, J.E., A.A. Kramer, D.S. McNair, F.M. Malila. 2006. Acute physiology and chronic health evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients. *Critical Care Medicine* **34**(5) 1297–1310.