Submitted to *Operations Research*
manuscript OPRE-2015-12-707

Allocation of Intensive Care Unit Beds in Periods of High Demand

Huiyin Ouyang
Faculty of Business and Economics, The University of Hong Kong, Hong Kong, oyhy@hku.hk

Nilay Tanık Argon, Serhan Ziya
Department of Statistics and Operations Research, The University of North Carolina at Chapel Hill, NC 27599, nilay@email.unc.edu, ziya@email.unc.edu

The objective of this paper is to use mathematical modeling and analysis to develop insights into and policies for making bed allocation decisions in an Intensive Care Unit (ICU) of a hospital during periods when the patient demand is high. We first develop a stylized mathematical model in which patients' health conditions change over time according to a Markov chain. In this model, each patient is in one of two possible health stages, one representing the *critical* and the other representing the *highly critical* health stage. The ICU has limited bed availability and therefore when a patient arrives and no beds are available, a decision needs to be made as to whether the patient should be admitted to the ICU and if so which patient in the ICU should be transferred to the general ward. With the objective of minimizing the long-run average mortality rate, we provide analytical characterizations of the optimal policy under certain conditions. Then, based on these analytical results, we propose heuristic methods, which can be used under assumptions that are more general than what is assumed for the mathematical model. Finally, we demonstrate that the proposed heuristic methods work well by a simulation study, which relaxes some of the restrictive assumptions of the mathematical model by considering a more complex transition structure for patient health and allowing for patients to be possibly queued for admission to the ICU and readmitted from the general ward after they are discharged.

Key words: Health care operations; Dynamic control; Markov decision processes
History: This paper was first submitted on Dec 14, 2015.

1. Introduction

Efficient management of Intensive Care Unit (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 expected 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" against 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 paper is to use mathematical modeling and analysis to develop insights and policies which can be useful when making these complex decisions in practice particularly under conditions where there is significantly high demand for limited ICU bed capacity.

The general framework we use to fulfill the objective we outlined above is as follows. We first develop a stylized mathematical formulation for the ICU. This relatively simple formulation (compared with the full complexity of the actual problem) allows us to provide characterizations for the optimal policy under certain conditions. These characterizations not only provide overall insights into "good" ICU admit/discharge decisions but also lead to the development of several heuristic policies that can potentially be used in practice. Finally, we test the performances of these policies with a simulation study relaxing some of the restrictive assumptions of the stylized mathematical formulation and find that the policies we propose perform quite well in comparison with some alternative benchmarks.

Our mathematical formulation assumes that each patient's health condition changes over time. Specifically, there are two discrete-time Markov chains with one representing the evolution of the patients in the ICU and one representing the evolution of the patients outside the ICU. Each Markov chain has four states corresponding to *death*, *highly critical*, *critical*, and *survival*, where death and survival states are absorbing 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 and therefore any patient in the system can only be in one of the two health stages, critical or highly critical. 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 discharge any of the highly critical or critical patients to the general ward early. The objective is to minimize the long-run average number of deaths.

We start our mathematical analysis by first considering an extreme setting, where the ICU has a single bed. The main insight that comes out of this analysis is 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 that which one of these two factors is more dominant depends on the overall level of demand for the ICU. We then consider the general setting, where the ICU has some arbitrary but finite number of beds. We formulate the decision problem as a Markov decision process (MDP) and prove that in general the optimal policy is a state-dependent policy, where the admission/discharge decisions depend on the mix of patients present in the ICU at the time the decisions are made.

While our mathematical analysis leads to useful insights into ICU patient admit/discharge decisions, it does not directly answer the question of how one can turn these insights into practical policies and how such policies would perform under realistic conditions. To address that, we introduce a simulation model, which enriches the mathematical model in a number of directions making it a more realistic environment for proposing and testing heuristics. Specifically, for this simulation model, we assume that patients can be in one of six health stages and they can transition from one stage to another according to a transition probability structure that is more complex than the one assumed in the mathematical model. Unlike the case in the mathematical model, patients who have already been discharged to the general ward are also considered for readmission to the ICU and patients who are initially admitted to the general ward can be admitted to the ICU later on. Finally, in accordance with our focus on bed allocation decisions during periods of high demand, the model considers a 36-week time horizon with a 12-week period in the middle during which the ICU observes more than usual demand levels with the arrival rate of patients first increasing and then decreasing and going back to regular levels. (The scenario is created based on the estimates of the US Centers for Disease Control and Prevention for flu seasons.) All of these additional features lead to an environment which is significantly different from the one assumed by our mathematical model. Nevertheless, the relative simplicity of our structural results make it possible for us to propose policies that can be used under more general conditions, such as those assumed by our simulation framework.

Specifically, we propose three different heuristic policies and compare their performances with those of four benchmarks. The three heuristics are named the Ratio Policy (RP), the Aggregated Ratio Policy (ARP), and the Aggregated Optimal Policy (AOP). RP is the policy that prioritizes patients according to their expected net benefit from ICU (increase in survival probability as a result of being treated in ICU) divided by their expected length-of-stay, ARP is a version of RP that assumes cruder patient health classifications (same as assumed in the mathematical model), and AOP is the policy that essentially uses the optimal policy for the mathematical model by assuming the same crude classifications as ARP. Note that RP and ARP are both state-independent policies while AOP is a state-dependent policy. Our simulation results indicate that even though all three heuristics perform well compared with the benchmarks, RP is the best policy overall. The fact that the best-performing policy is state-independent suggests that the optimality of state-dependent policies established for the mathematical model may not hold in general or that it might be difficult to identify “good” state-dependent policies. However, it is important to note that, as we explain in detail in the paper, even though our simulation study helps us further develop our intuition into what kind of policies are likely to perform well, one should refrain from reaching definite practical conclusions mainly because research on ICU patients is not at a level where we have a clear understanding of how one should model the health condition of a patient and its evolution and as a result there is significant uncertainty as to what the “right” simulation model is. Therefore, one should not ignore the possibility that state-dependent policies might be superior and future studies should continue to consider them. In any case, however, the good performance of RP in our simulation study is promising for the future as it suggests that simple policies like RP, which only requires estimates on patients’ survival probability and expected length of stay, could be good enough and there may not be a need for more sophisticated decision making tools.

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 paper, but only highlight a few papers as examples.

Strand and Flaatten (2008) provide 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 (Sequential Organ-Failure Assessment) (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 study the benefits of ICU care and the effects of rationing beds in times of limited availability. Sinuff et al. (2004) review past studies on bed rationing and find that admission to the ICU is associated with lower mortality. Shmueli and Sprung (2005) study the potential survival benefit for patients of different types and severity (measured by APACHE II score) and Kim et al. (2014) quantify 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 carry out a simulation study to test various patient admission policies and find 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 patient outcomes. For example, Chalfin et al. (2007) and Cardoso et al. (2011) study patients immediately admitted to ICU and those who had delayed admissions (i.e., waited longer than 6 hours for admission) and conclude 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) find patients are discharged more quickly when ICU occupancy is high, and such patients are 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 develop and analyze models with the goal of generating insights into capacity related questions for ICUs and Step Down Units (SDUs) and how patient admission and discharge decisions should be made. 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. Armony et al. (2018) develop a queueing model for an ICU together with an SDU and using this model provide insights into the optimal size for the SDU.

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 main reason why these differences are 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 paper 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 paper 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 in the formulation. While Ulukus et al. (2011) consider a more general form for the termination cost and multiple job classes, they 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 paper, 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. Model Description

In this section, we describe the stylized mathematical formulation we use to generate insights into "good" bed allocation decisions and develop practical heuristic methods. Specifically, in this model, 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 health 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. Note that we use the general ward to represent any non-ICU care unit, which includes actual hospital wards, step-down or transitionary care units, nursing homes, and any other facility that can accommodate the patients but cannot provide an ICU-level service to the patients. In our model, we assume that all these non-ICU beds are identical and the capacity of the general ward is infinite.

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 $i = 1, 2$; $\lambda \equiv \lambda_1 + \lambda_2$ denote the probability that there will be a patient arrival; let $\bar{\lambda} \equiv 1 - \lambda$ denote the probability of no arrival, where we assume $\lambda < 1$. 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 corresponds to the death of the patient while stage 3 represents the patient's survival. 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 1.

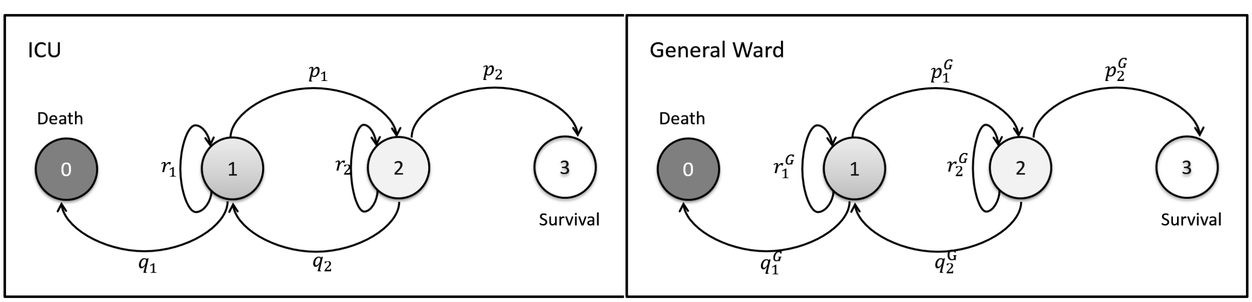

Figure 1 Transition diagram of patient evolution in the ICU and general ward

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 justified in some contexts because even in practice such simplifications are made to bring highly complex decision problems to manageable levels. 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. Nevertheless, in Section 6, we consider a more detailed and arguably more "realistic" evolution model for patients' health condition and demonstrate how our analysis based on this rather simplified structure would be useful.

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 $\mathbf{a} = (a_1, a_2)$, 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

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

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

$$\phi_1 = \frac{\beta_1 + \beta_1\beta_2}{1+\beta_1+\beta_1\beta_2}, \quad \phi_2 = \frac{\beta_1\beta_2}{1+\beta_1+\beta_1\beta_2}. \tag{2}$$

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

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

Note that while a hospital could possibly also have financial considerations when making patient admit/discharge decisions particularly under non-emergency conditions, in this paper, in parallel with our focus on periods during which there is excessively high demand, we restrict our focus to policies that aim to minimize the number of deaths.

Let $P_{(a_1,a_2)}(x_1,x_2,y_1,y_2)$ denote the probability that the system will transition to state (y_1,y_2) from state (x_1,x_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.,

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

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 $J^{\pi^*}(\mathbf{x})\le J^\pi(\mathbf{x})$ for any $\pi\in\Pi$ and $\mathbf{x}\in\mathcal{S}$. Note that this MDP is a 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

$$h(x_1,x_2)+g = \min_{(a_1,a_2)\in\mathcal{A}(x_1,x_2)}\Big\{c(x_1,x_2,a_1,a_2)+\sum_{(y_1,y_2)\in\mathcal{S}}P_{(a_1,a_2)}(x_1,x_2,y_1,y_2)h(y_1,y_2)\Big\}, \qquad (3)$$

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) for each $(x_1,x_2)\in\mathcal{S}$.

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. It will also provide support for one of the heuristic policies we propose in Section 6.

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 is 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. (The proof for the proposition as well as the proofs of all the other analytical results in the paper are provided in the Online Appendix.) 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

$$L_1 = \frac{p_1 + p_2 + q_2}{p_1 p_2 + q_1 p_2 + q_1 q_2}, \; L_2 = \frac{p_1 + q_1 + q_2}{p_1 p_2 + q_1 p_2 + q_1 q_2}. \tag{4}$$

Proposition 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

$$\lambda \leq \frac{(\phi_i^G - \phi_i) - (\phi_{3-i}^G - \phi_{3-i})}{(\phi_i^G - \phi_i) - (\phi_{3-i}^G - \phi_{3-i}) + \left[L_i(\phi_{3-i}^G - \phi_{3-i}) - L_{3-i}(\phi_i^G - \phi_i)\right]}. \tag{5}$$

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 $(\phi_i^G - \phi_i)/L_i$ 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 larger $(\phi_i^G - \phi_i)/L_i$ as "high-value-rate" patients. Then, according to Proposition 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 Proposition 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.

5. Analysis of the multi-bed ICU model

In this section, we consider the long-run average cost optimization problem with optimality equations given in (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 $\mathcal{A}^*(x_1, x_2)$:

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

Since the state space and action space are finite and costs are bounded, $\mathcal{A}^*$ is non-empty. In general, the set $A^*(x_1, x_2)$ 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 $a^*(x_1, x_2) = (a_1^*(x_1, x_2), a_2^*(x_1, x_2))$, where

$$a_1^*(x_1,x_2) = \min\{\bar{a}_1 : (\bar{a}_1,\bar{a}_2) \in A^*(x_1,x_2)\}, \text{ and } a_2^*(x_1,x_2) = \min\{\bar{a}_2 : (a_1^*(x_1,x_2),\bar{a}_2) \in A^*(x_1,x_2)\}.$$

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 1, 2, and 3 presented in this section below characterize the structure of the optimal policy. The proofs of these theorems are provided in the Appendix, where we first analyze the system with 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.

5.1. 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.

THEOREM 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 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. We found numerical examples that show that when this condition does not hold, the optimal policy is not necessarily non-idling meaning that the ICU would only accept patients from a particular stage and keep some of the ICU beds empty even when there is demand from patients of the other stage.

5.2. 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 the assumption that $\beta_i < \beta_i^G$ for $i = 1, 2$, 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 theorem, it turns out that the optimal decision has a threshold structure.

THEOREM 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 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, which depends on all the model parameter values and thus survival probabilities as well as lengths of stay, 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. (We should note however that it is not clear whether the potential benefits of using such a state-dependent policy can be realized in practice. We investigate and discuss this issue in detail in Section 6.1.)

The fact that in general the optimal policy can be state-dependent 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, then 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 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. *Suppose that* $\beta_i < \beta_i^G$ *for* $i = 1, 2$, *and for some fixed* $k \in \{1, 2\}$

$$\phi_k^G - \phi_k < \phi_{3-k}^G - \phi_{3-k} \text{ and } L_k \geq L_{3-k}. \tag{6}$$

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 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.

6. Simulation study

In Sections 4 and 5, we analyzed relatively simple formulations with the objective of generating insights and coming up with heuristic methods, which are flexible enough to be used under more general and realistic conditions. In this section, we have two main goals. First, to demonstrate how one can construct heuristic policies based on our analysis assuming that we know how health status of patients evolve in the ICU and in the general ward, and propose specific policies for the assumed evolution model. Second, to report the findings of our simulation study where we investigated how the policies we generated perform. Our simulation model relaxes some of the restrictive assumptions of the mathematical model of Section 3. In particular, we consider a more detailed and realistic health evolution model, a non-stationary patient arrival process, the possibility of patients to wait for admission to the ICU, and possible readmission of patients who have already been discharged from the ICU. We start with describing the health evolution model used in our simulation model.

6.1. Simulation model

Given what is known in the medical literature, it is not possible to construct a detailed, realistic model for describing how each patient's health status evolves in and outside the ICU. This obviously poses a significant challenge in reaching the two goals outlined above. While our mathematical model, which assumes two health stages and possible transitions between the two, broadly captures what happens in practice and is in fact in line with the only proposed classification protocol developed (see Christian et al. (2006)), it is also very likely that the model, with its mathematically convenient construction like having Markovian transition probabilities, fails to capture some of the features that one might see in reality. For example, two patients might be in the same "health stage" with respect to some objective criterion (one can think of a classification based on the SOFA score as used by Christian et al. (2006)) but assuming they would have the same stochastic evolution in the future could be an oversimplification if, for instance, one of the patients has just arrived and the other patient has been in the ICU for hours or one of the patients' health status has been gradually improving suggesting a positive trend while the other's health has been declining. Thus, there are a number of ways our mathematical model can be generalized in order to make it more "realistic."

The health evolution model we used in our simulation study, which is depicted in Figure 2, helped us capture some of the features described above. The model assumes four levels of "criticality" but also takes into account the direction of the last transition for the intermediate two criticality levels. Explicit consideration of the last transition makes it possible to at least partially capture the effect of trend in the evolution of patients' health status. More specifically, we assume that at any point in time patients in the ICU or the general ward belong to one of the six stages $\{1, 2H, 2L, 3H, 3L, 4\}$, where stages 1 and 4 denote the most critical and the least critical levels, stages 2H and 3H denote the intermediate criticality levels for patients whose health condition has been declining (i.e., the last transition was from a healthier stage) and stages 2L and 3L denote the intermediate criticality levels for patients whose health condition has been improving (i.e, the last transition was from a less healthy stage). Within one time period, which is assumed to be one hour (and thus one day consists of 24 time periods), the health condition of patients in stage $i \in \{1, 2H, 2L, 3H, 3L, 4\}$ can improve with probability p_i, decline with probability q_i, or stay the same with probabilities $r_i = 1 - p_i - q_i$. Stages 0 and 5 are two absorbing stages where 0 corresponds to death and 5 corresponds to survival. (See Figure 2 to see the probabilities corresponding to each transition.)

When choosing the values for transition probabilities, rather than setting them completely randomly, we set them in a way that the system at least conforms to what we know from the medical

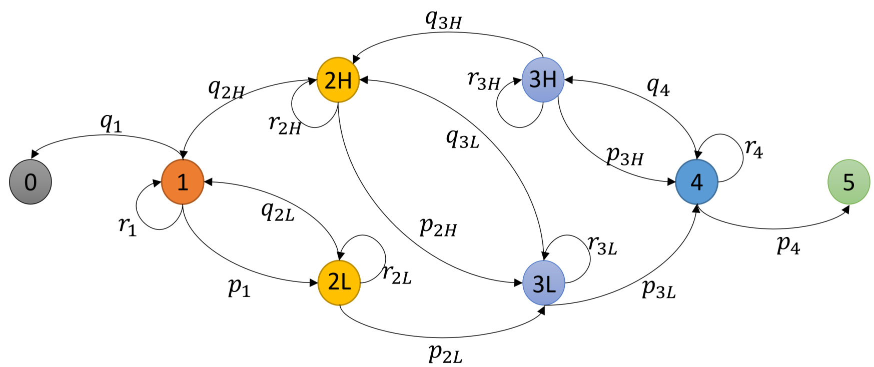

Figure 2 Transition diagram for patient evolution in the ICU (Patients in the general ward follow the same transition model with corresponding transition probabilities indicated by the superscript "G.")

literature. Several articles in the literature provide estimates on ICU length of stay and survival probabilities. 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. Therefore, we randomly generated scenarios so that the expected ICU death probability over all the scenarios is approximately 0.17 and the expected length of stay (with no early discharge) for the same is approximately $24 \times 12 = 288$ hours. In addition, we ensured that the generated scenarios satisfied the condition that patients who were previously in "healthier" stages are more likely to get "better."

When generating the random scenarios we first identified a baseline setting that conforms to the description above and then made random choices around this baseline. More specifically, we set $p_1 = 0.016$, $p_{2L} = 0.032U_{2L}$, $p_{2H} = 0.032U_{2H}$, $p_{3L} = 0.016U_{3L}$, $p_{3H} = 0.016U_{3H}$, $p_4 = 0.012$ and $q_1 = 0.0072$, $q_{2L} = 0.01V_{2L}$, $q_{2H} = 0.01V_{2H}$, $q_{3L} = 0.012V_{3L}$, $q_{3H} = 0.012V_{3H}$, $q_4 = 0.016$ where $U_{2L}, U_{3L}, V_{2H}, V_{3H}$ are independent random variables each uniformly distributed over $(0.5, 1)$, and $V_{2L}, V_{3L}, U_{2H}, U_{3H}$ are independent random variables each uniformly distributed over $(1, 1.5)$. (Note that the baseline level corresponds to the case where each random variable is set to 1.)

Kumar et al. (2009) do not provide any estimates on what the survival probabilities for the ICU patients would be if they were treated outside the ICU. In the absence of such estimates, recognizing that the condition of patients treated in non-ICU wards would be more likely to become worse and less likely to become better, for each $i \in \{1, 2H, 2L, 3H, 3L, 4\}$, we obtained q_i^G by multiplying q_i by a random coefficient uniformly distributed over $(1, 2)$, and p_i^G by multiplying p_i by a random coefficient uniformly distributed over $(0.5, 1)$.

In the simulation study, we focused on a time period during which the hospital experiences the flu season. To model patient arrivals realistically, we used Centers for Disease Control (CDC) flu season reports as well as FluSurge 2.0, the influenza patient demand prediction tool developed by CDC. As one can observe from Figure A1 in the Online Appendix, the flu season typically starts with a period where the arrival rate is mostly stationary, which is followed by an outbreak period, and ends with another stationary period. We considered a 36-week time period where during the first 12 weeks and the last 12 weeks patient demand is stationary (with an arrival probability of λ_{st} in each time period) while the outbreak and the non-stationary demand period is observed during the middle 12 weeks. According to the default scenario assumed by FluSurge 2.0, in this middle 12-week period, the Daily Percentage Change in Demand (DPCD) (i.e., percentage change in the expected number of new patient arrivals) is 3% during the first 6 weeks and -3% during the next 6 weeks. In our study, we considered two different settings, one with DPCD value of 3%, and the other with 5% (six weeks of increase followed by six weeks of decrease with the same absolute value for the rate). For the baseline stationary arrival rate, which the ICU observes during the first 12 weeks and the last 12 weeks, we considered three different levels. Specifically, we let the ICU load $\rho_{\text{st}} \triangleq \lambda_{\text{st}} E[L]/b$ (where b is the number of ICU beds) to be either 0.5, 0.8, or 1. The choice of baseline load on the ICU also determines the overall demand level during the outbreak period since the arrival rates of patients will increase starting from these baseline levels. As for the health stages for the new patients, rather than assuming that they all come in a given state, we assumed that there is patient heterogeneity. Specifically, letting θ_i denote the probability that the initial health stage for an incoming random patient is i, when generating scenarios, we let $\theta_i = (U_i^A + 1)/\sum_{j\in\{1,2L,2H,3L,3H,4\}}(U_j^A + 1)$, where $U_i^A \sim U(0,1)$ for each stage i.

We assumed that the ICU has 20 beds for our simulation study. (Note that the choice of a 20-bed ICU together with the three different load levels we consider in our study are consistent with the range of possible demand predictions of FluSurge 2.0 and a typical population/ICU bed ratio in the US.) We also assumed that, as in the mathematical model, there is no limit on the number of patients who can be accommodated in the general ward. (Note that while general ward beds are also limited in numbers in reality, they are more widely available than ICU beds and the key issue typically is the effective management of ICU beds.) However, in the simulation model, in accordance with commonly observed practice, we assumed that when a bed becomes available in the ICU and there are patients in the general ward, one of those patients is admitted to the newly vacated bed. With this feature, the simulation model allows the possibility of readmitting patients who were previously discharged from the ICU to the general ward back to the ICU and having patients who find the ICU full to queue up in the general ward for possible admission later on.

6.2. Proposed policies and benchmarks

In this section, we propose policies which are based on our mathematical analysis but are meant to be used in the more general construction assumed in the simulation model. By doing that, we will also be illustrating more generally in what way our mathematical results and insights can be used to develop heuristics that can be used under any future patient health evolution model that is supported by medical research and data. It is important to note that the policies we propose assume that patient demand is so high that ICU admits and discharges are done throughout the day as needed unlike some of the common practices in place under regular operating conditions, which restrict such decisions to be made and actions to be taken only during certain times of the day.

The first two policies described below are included mainly because they can serve as benchmark policies and do not necessarily represent policies that are used in practice.

First-Come-First-Served (FCFS): Patients are admitted to the ICU beds in the order they arrive. None of the patients are discharged early to the general ward when a new patient finds the ICU full. In such a case, the patient is admitted to the general ward and waits for an opening in the ICU. When a patient in the ICU leaves (as a result of death or survival), among the patients who are still in the general ward, the one who was first admitted to the general ward is admitted to the newly vacated bed in the ICU. This policy would clearly capture the policy of not being proactive about making the best possible use of the ICU and opting for a policy, which could largely be considered as "fair" rather than aiming to maximize "the greatest good for the greatest number."

Random Discharge Policy (RDP): Under this policy, if the ICU is fully occupied when a patient arrives one of the patients among the patients already in the ICU and the patient who has just arrived is randomly chosen and transferred to the general ward. When a patient in the ICU leaves (as a result of death or survival), one of the patients in the general ward is randomly chosen for admission to the newly vacated bed.

Greedy Policy (GP): This is an index policy, which gives priority according to the order determined by the differences $\phi_i^G - \phi_i$ where ϕ_i and ϕ_i^G denote the probability of death in the ICU and the general ward, respectively, for a patient in health stage i. Whenever an arriving patient finds the ICU full, the policy discharges the patient whose survival probability will have the smallest drop as a result of being treated in the general ward as opposed to the ICU. Similarly, when a patient leaves the ICU (as a result of death or survival), among the patients in the general ward, the patient with the most to benefit is chosen. Note that based on our mathematical analysis of the single-bed scenario, specifically Proposition 1, it might be reasonable to expect that GP would perform well when the patient demand is relatively low but when demand is high, as we assume in our simulation study, because the policy ignores the expected lengths of stay, we would not expect the policy to perform well.

Ratio Policy (RP): This is an index policy, which gives priority according to the order determined by $(\phi_i^G - \phi_i)/L_i$ where L_i is the expected length-of-stay for patients in health stage i. Whenever an arriving patient finds the ICU full, the policy discharges one of the patients with the smallest expected drop in the survival probability divided by the expected length of ICU stay. Similarly, when a patient leaves the ICU (as a result of death or survival), among the patients in the general ward, the patient with the largest value of $(\phi_i^G - \phi_i)/L_i$ is chosen. Our mathematical analysis provides strong support for this heuristic particularly when demand is high and thus one would expect good performance from this policy in the simulation study. Specifically, Proposition 1, which assumes the simplistic single-bed setting, finds that this policy is optimal when the arrival rate is sufficiently high. For the multi-bed scenario, we know from Theorem 2 that the optimal policy has a threshold structure, which would still be in line with RP, but we also have examples that show that RP is not optimal in general and that the optimal policy is state-dependent. Nevertheless, Theorem 3 finds that under a condition for which RP and GP would be in complete agreement, RP would be optimal.

An important assumption underlying GP and RP is that we can observe each patient's health stage precisely. Practically, however, this may not be possible. Patients could still be evolving according to some more sophisticated transition probability structure like the one we assumed in the simulation model but we might only be able to do some rough classification and make decisions accordingly without knowing precisely in which health stage each patient is in. In fact, this would most likely be the case at least in the foreseeable future as it is very difficult if not impossible to come up with a classification system that perfectly captures patient evolution at a very detailed level. To get a sense of how the policies we propose would perform in such a case, we also consider "aggregated" versions of GP and RP. They are aggregated in the sense that, as shown in Figure 3, if the patient is in one of the health stages 1, 2H, or 2L, the decision maker knows that the patient is in one of these stages but not the exact health stage. Thus, the decision maker assumes that the patient is in some aggregated stage $A1$. Similarly, if the patient is in one of the health stages 3H, 3L, or 4, the decision maker, not knowing the exact health stage of the patient, assumes that the patient is in the aggregated stage $A2$. This means that the decision maker puts patients in only one of two health stages as in the case of our mathematical formulation. This allows us to consider a setting where the "reality" is complicated (as described in the simulation study) but the decision maker follows the policies suggested by our mathematical analysis. Note that using the aggregated stages requires estimation of transition probabilities among the aggregated stages $A1$ and $A2$ as well as the death and survival stages 0 and 5. We explain how the decision maker makes this estimation in Online Appendix A5.3.

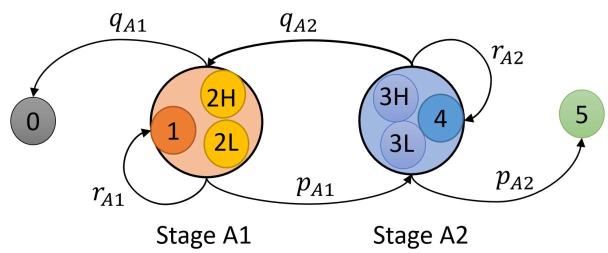

Figure 3 Aggregated two-stage transition diagram for patient evolution in the ICU.

Aggregated Greedy Policy (AGP): This policy is the same as GP except that the policy is applied over the aggregated classifications. When a patient from a particular aggregated stage is to be discharged or admitted from the general ward, one of the patients from that aggregated health stage is chosen randomly.

Aggregated Ratio Policy (ARP): This policy is the same as RP except that the policy is applied over the aggregated classifications. As in the case of AGP, when a patient from a particular aggregated stage is to be discharged or admitted from the general ward, one of the patients from that aggregated health stage is chosen randomly.

Aggregated Optimal Policy (AOP): When there are two health stages only and under the additional assumptions that when there are no patient readmissions from the general ward and patient arrival process is stationary, we can determine the optimal policy by solving the MDP formulation described in Section 5. AOP basically uses the actions this optimal policy suggests (when the arrival rate is set to the current arrival rate in the simulation model). As in the cases of AGP and ARP, AOP randomly picks among the patients who belong to the same aggregated health stage.

6.3. Results of the simulation study

In the simulation study, we considered two different DPCD values (3% and 5%), and three different levels for the baseline ICU load (0.5, 0.8, and 1) as described in Section 6.1. Thus, in total, we considered six different combinations. The performance measure for each policy π (described in Section 6.2) was chosen to be the mortality rate, M_π, which we define to be the percentage of deaths among the patients who arrived at the ICU for possible admission during the 36-week period. We generated 30 different transition probability scenarios for each one of the six DPCD-load pairs as described in Section 6.1 and ran 100 replications for each scenario. In each replication, we randomly determined the initial state of the system. Specifically, the number of patients initially in the ICU was set assuming that the number is uniformly distributed over the integers from

0 to b and the health stage i of each patient is determined using the probability distribution $\{\theta_i, i \in \{1, 2L, 2H, 3L, 3H, 4\}\}$.

Using simulation results, we made pairwise performance comparisons between RP and every other policy. Specifically, we calculated the mean value for $M_\pi - M_{RP}$ for every policy π (over the 100 replications) for each scenario and constructed a 95% confidence interval for the mean difference. In Figures 4, 5, and 6, we provide these confidence intervals along with the box plots, where we also indicate the 1st and 3rd quantiles, the minimum, and the maximum values.

From the figures, we can observe that RP has a superior performance overall when compared with the other policies. The good performance of RP is evident particularly when the comparison is made with respect to benchmark policies FCFS, RDP, GP, and AGP and the load on the ICU is very high. Given our mathematical analysis, the effect of system load on the performance of RP is not surprising. As we discussed before, when the system load is high, policies like GP, which exclusively takes into account the immediate benefit for the patients while ignoring system level factors such as the expected length of stay for the patients, are more likely to perform badly.

If we compare RP with the aggregated-type policies ARP and AOP, we observe that, even though these two policies perform better than the benchmarks, the performance of RP is again statistically better with the differences in the performances getting larger as the load on the system increases. Note that this comparison is important because as we discussed in Section 6.2, the model with which we are making decisions (e.g., our mathematical model) could be simpler than the "reality" (e.g., our simulation model as we assume in this paper). As the decision maker, we may not even know which specific health stage the patient is in but could only have some rough idea about the patient's health condition. With this comparison, we see that there is a benefit to knowing the health conditions of the patients in more detail especially when the system is heavily loaded.

To get a better sense as to why RP performs well and its performance gets better with increased ICU load recall Proposition 1 (particularly part (b)), which provides a necessary and sufficient condition for the optimality of RP for the special case where there is a single bed. According to the proposition, RP is optimal if the arrival probability exceeds a particular level, i.e., if the inequality (5) is violated. This suggests that in general RP could be more preferable when the ICU is highly loaded. In our simulation study, there are multiple beds, the arrival probability changes with time, and the inequality is written specifically for a model that assumes two health stages. Therefore, Condition (5) is not well-defined in the context of our simulation model. However, one can still get some general sense for whether or not the arrival probability is high enough to favor RP by adjusting the arrival probability λ by λ/b, using the aggregated version of the transition probabilities under each scenario, and then determining the percentage of time the inequality (5) is violated. Following this procedure, we obtained Table EC.1 in Online Appendix A5.4.

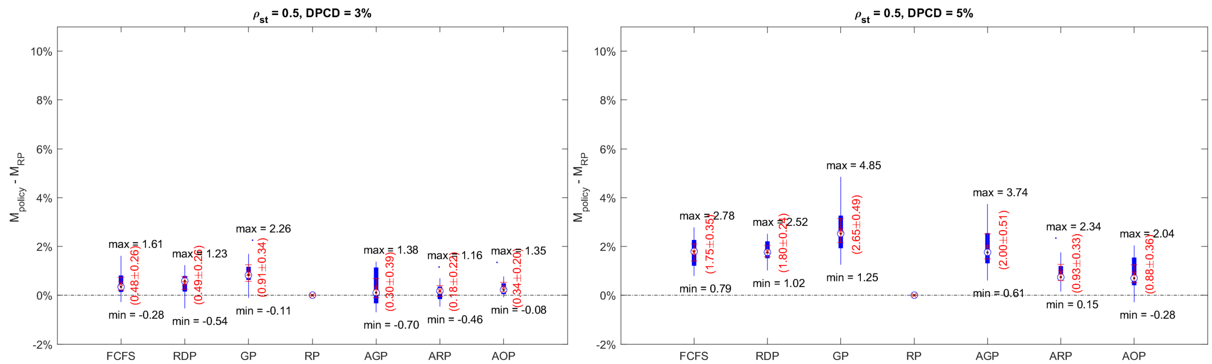

Figure 4 Pairwise comparisons of the differences in *average* mortality rates between other policies and RP for each scenario with $\rho_{\mathrm{st}} = 0.5$, where average is taken over 100 replications.

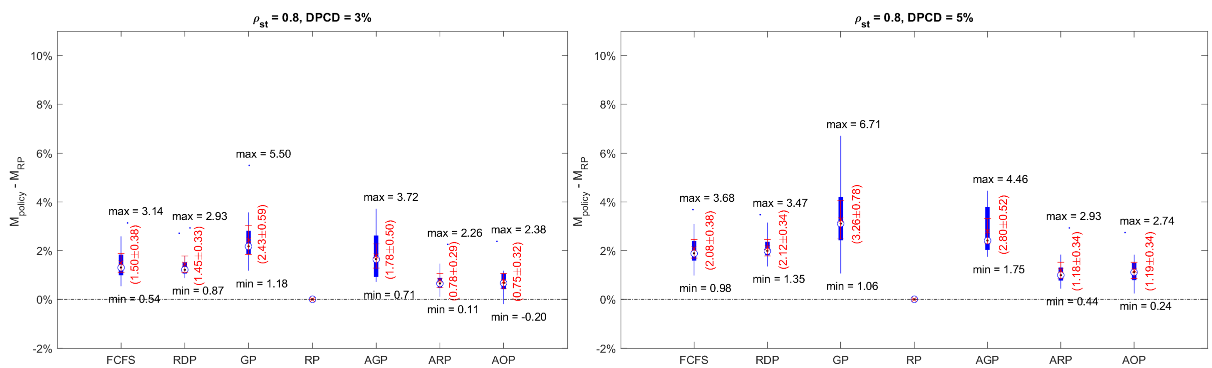

Figure 5 Pairwise comparisons of the differences in *average* mortality rates between other policies and RP for each scenario with $\rho_{\mathrm{st}} = 0.8$, where average is taken over 100 replications.

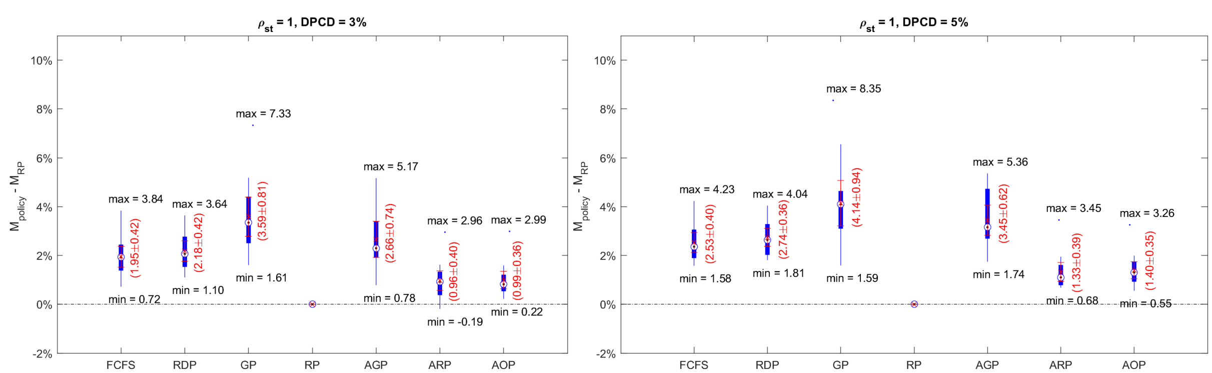

Figure 6 Pairwise comparisons of the differences in *average* mortality rates between other policies and RP for each scenario with $\rho_{\mathrm{st}} = 1$, where average is taken over 100 replications.

As we can observe from the table, in many of the scenarios considered in our simulation study, the fraction of time the adjusted version of Condition (5) is violated is either 1 or close to 1 providing some explanation as to why RP has such a good performance. It is also important to note that as the load on the system increases, the fractions under each scenario are also non-decreasing, which might explain why the performance of RP is more dominant when ICU load is higher.

Going back to Figures 4, 5, and 6, another observation we can make is that among the aggregated-type policies, ARP and AOP appear to perform better than AGP except for the case the load on the ICU is the smallest. If we compare ARP with AOP, we see that even though the mean performance of AOP is better than that of ARP for all ICU load levels, the differences are not statistically significant. This suggests that even when patients can only be classified at the aggregate level as described above and thus RP is not an option, using the policy that is optimal (for our stylized formulation) may not be justified and the aggregated version of RP might be acceptable.

The observations that RP performs better than AOP and AOP does not seem to have a statistically strong advantage over ARP highlight two important questions that are closely intertwined with each other: In searching for the discharge/admit policy to use in practice, can we restrict ourselves to policies that are state-independent? Given that the simulation study suggests that the best policy is state-independent does Theorem 2, which states that in general the optimal policy is of threshold-type, have any practical value? For several reasons, it is difficult to provide definite answers to these questions. First of all, we know that AOP is optimal for our mathematical model but this obviously does not mean that it would continue to perform well when we change the underlying model from one with two health stages (as in the mathematical model) to one with six health stages and a more complex transition structure (as in the simulation model). It is not even clear whether AOP is the best among all the aggregated-type policies one could use for the model assumed in the simulation study. In short, AOP may not have performed as well as one would hope but this does not rule out the possibility of the existence of a different state-dependent policy that performs better. But more importantly, even though our patient health evolution formulation assumed in the simulation study is highly likely to be an improvement over the one we assumed in the mathematical model we do not know how well this particular model captures reality. As we discussed in detail in Section 6.1, existing research on ICU patients is not at a level where we have a clear understanding of how ICU patients can be classified and how their health conditions evolve inside and outside the ICU. The model we used in the simulation study is only one possibility among the many plausible. Therefore, even though our simulation study provides some useful insights and directions for future work it would not be reasonable to make immediate generalizations from our observations. With more research in this area, we will have an increasingly better understanding of ICU patients and be able to develop models that are increasingly better representations of reality. It is possible that with changes in the health evolution model, performances of the policies relative to each other will also change and it will be prudent to construct potentially good new state-dependent policies and investigate their performances. The insights that come out of the optimal policy characterizations given in Theorem 2, which describe how the composition of the patients in the ICU should influence admit/discharge decisions, can be very helpful in the construction of such policies.

Even though our simulation study cannot provide a definite answer to the question of which policy would work better in practice, the fact that RP had the best performance is good news. The policy is simple, easily generalizable, intuitive, and does not need to keep track of system state information. It is also important to note that the policy only requires the estimation of expected net benefits and the expected lengths-of-stay for each health stage, not the individual transition probabilities. This not only makes it much easier to implement RP in practice but also means that the policy is highly robust to transition probability estimates and the assumptions made regarding the underlying patient health and transition formulation.

Finally, in this section, we investigate patients' lengths-of-stay in the ICU under each policy. Figures A2, A3, and A4 Figures A2, A3, and A4 given in Online Appendix A5.5 summarize the results of our analysis. We can observe from the figures that if we leave aside FCFS, there are no notable differences between the policies with respect to average lengths of ICU stay. Long lengths-of-stay under FCFS is not surprising because under that policy patients leave the ICU only when they are dead or they reach the survival stage. They are never discharged early to accommodate other patients. Under any of the other policies, patients can be discharged from the ICU even though they still need ICU care and this results in shorter lengths-of-stay. We can also observe from the figures that the lengths-of-stay under every policy except FCFS decrease as the ICU load increases. This is because except in the case of FCFS, the more patients there are in need of ICU, the higher the chances that any given patient's ICU stay is cut short, which ultimately leads to shorter average lengths-of-stay.

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 paper has been to provide insights into and develop policies for making such allocation decisions.

What mainly sets our analysis apart from prior work is that in our model we allow the patients to move from one health stage to another and allocation decisions are made based on the patients' updated health conditions. This formulation captures an important feature of the actual problem and nicely fits with the triage protocol proposed by Christian et al. (2006). But more importantly, the model allowed us to go deeper and establish properties that appear to be difficult to identify using formulations considered in prior work. For example, 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 patients in the ICU.

Considering the complexity of the actual decision problem we are interested in, the mathematical model we analyze in this paper is stylized and therefore it is natural to question the generalizability of the main insights. Indeed, our simulation study, which, unlike the mathematical model, allows readmissions from the general ward and considers a more complex patient health evolution formulation, suggests that there may not be a justification for searching for a complex policy that prioritizes based on the patient mix in the ICU. On the other hand, the simulation study also shows that some of the policies that are proposed based on our mathematical analysis performs well even under the more general conditions of the simulation model. The fact is that not enough is known about the ICU patients for us to be able to construct a very realistic description of patient evolution. The more complex model considered in the simulation study is another simplification at best. Therefore, not only the conjectures on the generalizability of the policies should be taken with a grain of salt, but one should also not quickly conclude from our simulation study that in practice there is no need to consider policies that take patient mix into account. Nevertheless, our results provide some confidence that despite the complexity of the decision problem in practice, relatively simple policies might work well and the paper provides some useful guidance for what future research and data collection efforts should focus on in order to develop useful patient classification and triage protocols and ultimately decision support tools that can be implemented in practice.

Acknowledgments

The authors would like to thank the department editor, the associate editor, and the referees whose comments and suggestions greatly improved this paper. This research was supported by NSF grants CMMI-1234212 and CMMI-1635574.

References

Armony, M., C.W. Chan, B. Zhu. 2018. Critical care capacity management: Understanding the role of a step down unit. *Production and Operations Management* **27**(5) 859–883.

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.

Dobson, G., H.H. Lee, E. Pinker. 2010. A model of ICU bumping. *Operations Research* **58**(6) 1564–1576.

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.

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.

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.

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.

Sennott, L.I. 1999. *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley & Sons.

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.

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.

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.

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.

Online Appendix

This document provides the proofs of the analytical results and supplemental information for the simulation study.

Appendix A1: Proof of the analytical result in Section 4

Proof of Proposition 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,

$$
\begin{aligned}
&P^1_{00}=\bar{\lambda}+\lambda_1 q_1+\lambda_2 p_2,\ P^1_{01}=\lambda_1 r_1+\lambda_2 q_2,\ P^1_{02}=\lambda_1 p_1+\lambda_2 r_2,\\
&P^1_{10}=(\bar{\lambda}+\lambda_1)q_1+\lambda_2 p_2,\ P^1_{11}=(\bar{\lambda}+\lambda_1)r_1+\lambda_2 q_2,\ P^1_{12}=(\bar{\lambda}+\lambda_1)p_1+\lambda_2 r_2,\\
&P^1_{20}=p_2,\ P^1_{21}=q_2,\ P^1_{22}=r_2.
\end{aligned}
$$

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 $a = q_1q_2 + q_1p_2 + p_1p_2$ and

$$D^{\bar{\pi}_1} = \bar{\lambda}\Big[a\bar{\lambda} + \lambda_1(p_1 + p_2 + q_2 + a) + \lambda_2(p_1 + q_1 + p_2 + q_2)\Big] + (\lambda_1 + \lambda_2)\Big[\lambda_1(p_1 + p_2 + q_2) + \lambda_2\Big].$$

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 $J^{\bar{\pi}_1} = u_0m_0 + u_1m_1 + u_2m_2$. We can then show that

$$\begin{aligned} J^{\bar{\pi}_1}D^{\bar{\pi}_1} = (\bar{\lambda}a + \lambda_1 a + \lambda_2 p_2)\lambda_1 q_1 + \Big[\lambda_1(p_2 + q_2 - a) + \lambda_2 q_2\Big]\Big[\bar{\lambda}q_1 + \lambda_1(q_1 + \phi_1^G) + \lambda_2\phi_1^G\Big] \\ + \Big[\bar{\lambda}(\lambda_1 p_1 + \lambda_2(p_1 + q_1 - a)) + (\lambda_1 + \lambda_2)(\lambda_1 p_1 + \lambda_2 r_2)\Big]\Big[\lambda_1\phi_1^G + \lambda_2\phi_2^G\Big]. \quad \text{(A1)} \end{aligned}$$

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,

$$\begin{aligned} &P_{00}^2 = \bar{\lambda} + \lambda_1 q_1 + \lambda_2 p_2,\ P_{01}^2 = \lambda_1 r_1 + \lambda_2 q_2,\ P_{02}^2 = \lambda_1 p_1 + \lambda_2 r_2, \\ &P_{10}^2 = q_1,\ P_{11}^2 = r_1,\ P_{12}^2 = p_1, \\ &P_{20}^2 = (\bar{\lambda} + \lambda_2)p_2 + \lambda_1 q_1,\ P_{21}^2 = (\bar{\lambda} + \lambda_2)q_2 + \lambda_1 r_1,\ P_{22}^2 = (\bar{\lambda} + \lambda_2)r_2 + \lambda_1 p_1. \end{aligned}$$

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

$$\begin{aligned} (\lambda_1(1 - q_1) + \lambda_2(1 - p_2))u_0' &= q_1u_1' + ((\bar{\lambda} + \lambda_2)p_2 + \lambda_1 q_1)u_2', \\ (q_1 + p_1)u_1' &= (\lambda_1 r_1 + \lambda_2 q_2)u_0' + ((\bar{\lambda} + \lambda_2)q_2 + \lambda_1 r_1)u_2', \\ u_0' + u_1' + u_2' &= 1. \end{aligned}$$

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

$$D^{\bar{\pi}_2} = \bar{\lambda}\Big[a\bar{\lambda} + (p_1 + q_1 + p_2 + q_2)\lambda_1 + (p_1 + q_1 + q_2 + a)\lambda_2\Big] + (\lambda_1 + \lambda_2)\Big[\lambda_1 + (p_1 + q_1 + q_2)\lambda_2\Big].$$

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

$$J^{\bar{\pi}_2} D^{\bar{\pi}_2} = (\bar{\lambda}a + \lambda_1 q_1 + \lambda_2 a)\lambda_1 q_1 + \Big[\lambda_1 p_1 + \lambda_2(p_1 + q_1 - a)\Big]\Big[\lambda_1(q_1 + \phi_2^G) + \lambda_2\phi_2^G\Big] + \Big[\bar{\lambda}(\lambda_1(p_2 + q_2 - a) + \lambda_2 q_2) + (\lambda_1 + \lambda_2)(\lambda_1 r_1 + \lambda_2 q_2)\Big]\Big[q_1 + \lambda_1\phi_1^G + \lambda_2\phi_2^G\Big]. \quad \text{(A2)}$$

Next, take the difference of $J^{\bar{\pi}_1}$ (from A1) and $J^{\bar{\pi}_2}$ (from A2), after some algebra, we have,

$$J^{\bar{\pi}_1} - J^{\bar{\pi}_2} = \frac{M\Bigg((1-\lambda)p_1p_2\big[\lambda\big[p_1(\beta_1^G - \beta_1) + p_2\beta_1^G(\beta_2 - \beta_2^G)\big] + (\beta_1^G - \beta_1) + \beta_1\beta_1^G(\beta_2 - \beta_2^G)\big]\Bigg)}{D^{\bar{\pi}_1}D^{\bar{\pi}_2}(1 + \beta_1^G + \beta_1^G\beta_2^G)}$$
$$= \frac{Mp_1p_2\beta_1\beta_1^G}{D^{\bar{\pi}_1}D^{\bar{\pi}_2}(1 + \beta_1^G + \beta_1^G\beta_2^G)}\left(\lambda\left[\frac{\beta_1^G - \beta_1}{p_2\beta_1\beta_1^G} + \frac{\beta_2 - \beta_2^G}{q_1} + (1-\lambda)\left(\frac{\beta_1^G - \beta_1}{\beta_1\beta_1^G} + (\beta_2 - \beta_2^G)\right)\right]\right),$$

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 $a = p_1p_2 + q_1p_2 + q_1q_2 = (p_1 + q_1)(p_2 + q_2) - p_1q_2 \le (p_1 + q_1)(p_2 + q_2)$, we have $1 - a \ge 0$ and

$$p_1 + q_1 + p_2 + q_2 - 2a \ge p_1 + q_1 + p_2 + q_2 - 2(p_1 + q_1)(p_2 + q_2)$$
$$= (p_1 + q_1)(1 - p_2 - q_2) + (p_2 + q_2)(1 - p_1 - q_1) \ge 0,$$

and thus $M \ge 0$. Then, we have $J^{\bar{\pi}_1} - J^{\bar{\pi}_2} \le 0$ if and only if

$$\lambda\left[\frac{\beta_1^G - \beta_1}{p_2\beta_1\beta_1^G} - \frac{\beta_2^G - \beta_2}{q_1}\right] + (1-\lambda)\left[\frac{\beta_1^G - \beta_1}{\beta_1\beta_1^G} - (\beta_2^G - \beta_2)\right] \le 0. \quad \text{(A3)}$$

Next, we find that

$$\frac{\beta_1^G - \beta_1}{\beta_1^G\beta_1} - (\beta_2^G - \beta_2) = \frac{(\phi_1^G - \phi_2^G) - (\phi_1 - \phi_2)}{(\phi_1 - \phi_2)(\phi_1^G - \phi_2^G)}, \quad \frac{\beta_1^G - \beta_1}{p_2\beta_1^G\beta_1} - \frac{\beta_2^G - \beta_2}{q_1} = \frac{L_2(\phi_1^G - \phi_1) - L_1(\phi_2^G - \phi_2)}{(\phi_1 - \phi_2)(\phi_1^G - \phi_2^G)},$$

since

$$\frac{\beta_1^G - \beta_1}{\beta_1^G\beta_1} - (\beta_2^G - \beta_2) = \frac{1}{\beta_1} - \frac{1}{\beta_1^G} - (\beta_2^G - \beta_2) = \left(\frac{1}{\beta_1} + \beta_2 + 1\right) - \left(\frac{1}{\beta_1^G} + \beta_2^G + 1\right)$$
$$= \frac{1 + \beta_1 + \beta_1\beta_2}{\beta_1} - \frac{1 + \beta_1^G + \beta_1^G\beta_2^G}{\beta_1^G} = \frac{1}{\phi_1 - \phi_2} - \frac{1}{\phi_1^G - \phi_2^G} = \frac{(\phi_1^G - \phi_2^G) - (\phi_1 - \phi_2)}{(\phi_1 - \phi_2)(\phi_1^G - \phi_2^G)}$$

and

$$\frac{\beta_1^G - \beta_1}{p_2\beta_1^G\beta_1} - \frac{\beta_2^G - \beta_2}{q_1} = \frac{1}{q_1p_2}\left(\frac{q_1}{\beta_1} - \frac{q_1}{\beta_1^G} - p_2(\beta_2^G - \beta_2)\right) = \frac{1}{q_1p_2}\left(\frac{q_1}{\beta_1} + p_2\beta_2 - \frac{q_1}{\beta_1^G} - p_2\beta_2^G\right)$$
$$= \frac{1}{q_1p_2}\left(p_1 + q_2 - \frac{q_1}{\beta_1^G} - p_2\beta_2^G\right) = \frac{1}{q_1p_2}\left(p_1 + q_2 - \frac{q_1 + p_2\beta_1^G\beta_2^G}{\beta_1^G}\right)$$

$$
\begin{aligned}
&= \frac{1+\beta_1^G+\beta_1^G\beta_2^G}{q_1p_2\beta_1^G}\left(\frac{(p_1+q_2)\beta_1^G}{1+\beta_1^G+\beta_1^G\beta_2^G} - \frac{q_1+p_2\beta_1^G\beta_2^G}{1+\beta_1^G+\beta_1^G\beta_2^G}\right)\\
&= \frac{1}{q_1p_2(\phi_1^G-\phi_2^G)}\left((p_1+q_2)(\phi_1^G-\phi_2^G) - q_1(1-\phi_1^G) - p_2\phi_2^G\right)\\
&= \frac{1}{q_1p_2(\phi_1^G-\phi_2^G)}\left((p_1+q_1+q_2)\phi_1^G - (p_1+p_2+q_2)\phi_2^G - q_1\right)\\
&= \frac{p_1p_2+q_1p_2+q_1q_2}{q_1p_2(\phi_1^G-\phi_2^G)}\left(\frac{p_1+q_1+q_2}{p_1p_2+q_1p_2+q_1q_2}\phi_1^G - \frac{p_1+p_2+q_2}{p_1p_2+q_1p_2+q_1q_2}\phi_2^G - \frac{q_1}{p_1p_2+q_1p_2+q_1q_2}\right)\\
&\big(\text{Since } q_1=(p_1+q_1+q_2)\phi_1-(p_1+p_2+q_2)\phi_2\big)\\
&= \frac{1}{(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)}\left(\frac{p_1+q_1+q_2}{p_1p_2+q_1p_2+q_1q_2}\phi_1^G - \frac{p_1+p_2+q_2}{p_1p_2+q_1p_2+q_1q_2}\phi_2^G - \frac{(p_1+q_1+q_2)\phi_1-(p_1+p_2+q_2)\phi_2}{p_1p_2+q_1p_2+q_1q_2}\right)\\
&= \frac{1}{(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)}\left(L_2\phi_1^G - L_1\phi_2^G - L_2\phi_1 + L_1\phi_2\right)\\
&= \frac{L_2(\phi_1^G-\phi_1) - L_1(\phi_2^G-\phi_2)}{(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)}.
\end{aligned}
$$

Then, since $(\phi_1-\phi_2)(\phi_1^G-\phi_2^G)\geq 0$, (A3) is equivalent to

$$
\lambda\left[L_2(\phi_1^G-\phi_1) - L_1(\phi_2^G-\phi_2)\right] + (1-\lambda)\left[(\phi_1^G-\phi_1)-(\phi_2^G-\phi_2)\right]\leq 0. \tag{A4}
$$

Then,

(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 (A4) 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 (A4) 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., (A4) holds in the opposite direction if

$$
\lambda < \frac{(\phi_1^G-\phi_1)-(\phi_2^G-\phi_2)}{(\phi_1^G-\phi_1)-(\phi_2^G-\phi_2)+[L_1(\phi_2^G-\phi_2)-L_2(\phi_1^G-\phi_1)]}.
$$

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., (A4) holds if

$$
\lambda \leq \frac{(\phi_2^G-\phi_2)-(\phi_1^G-\phi_1)}{(\phi_2^G-\phi_2)-(\phi_1^G-\phi_1)+[L_2(\phi_1^G-\phi_1)-L_1(\phi_2^G-\phi_2)]}.
$$

□

Appendix A2: Infinite-horizon expected total discounted cost problem

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.

A2.1. Formulation of the infinite-horizon expected total discounted cost problem

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. 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

$$c_1^G = \frac{\alpha q_1^G(1-\alpha r_2^G)}{(1-\alpha r_1^G)(1-\alpha r_2^G)-\alpha^2 p_1^G q_2^G},\ c_2^G = \frac{\alpha^2 q_1^G q_2^G}{(1-\alpha r_1^G)(1-\alpha r_2^G)-\alpha^2 p_1^G q_2^G}. \tag{A5}$$

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:

$$v_\alpha(x_1,x_2) = \min_{(a_1,a_2)\in\mathcal{A}(x_1,x_2)} \{V_\alpha(x_1,x_2,a_1,a_2)\}, \tag{A6}$$

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 A1. For a function $w(x_1, x_2)$ with $x_1, x_2 \geq 0$ and $x_1 + x_2 \leq b$

$$\Gamma w(x_1, x_2) = \sum_{j_1=0}^{x_1+x_2+1} \sum_{j_2=0}^{x_1+x_2+1-j_1} P(j_1, j_2 | x_1, x_2) w(j_1, j_2). \tag{A7}$$

A2.2. Main results for the infinite-horizon expected total discounted cost problem

For the infinite-horizon expected total α-discounted cost problem (A6), we denote the set of optimal actions in state (x_1, x_2) by $\mathcal{A}^*_\alpha(x_1, x_2)$, i.e.,

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

Since the state space and action space are finite and costs are bounded, $\mathcal{A}^*_\alpha$ 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

$$a^*_{1\alpha}(x_1, x_2) = \min\{\bar{a}_1 : (\bar{a}_1, \bar{a}_2) \in A^*_\alpha(x_1, x_2)\}, \text{ and } a^*_{2\alpha}(x_1, x_2) = \min\{\bar{a}_2 : (a^*_{1\alpha}(x_1, x_2), \bar{a}_2) \in A^*_\alpha(x_1, x_2)\}.$$

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.

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 A1. *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*

$$\alpha(q_1 + r_1 c_1^G + p_1 c_2^G) \leq c_1^G,\ \alpha(q_2 c_1^G + r_2 c_2^G) \leq c_2^G. \tag{A8}$$

From Proposition A1, we know that if (A8) holds we can restrict ourselves to the set of policies that is non-idling. Then, the optimality equations (A6) can be reduced to

$$v_\alpha(x_1, x_2) = T v_\alpha(x_1, x_2) \text{ for } (x_1, x_2) \in S, \tag{A9}$$

where the optimality operator T is defined as follows:

Definition A2. 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 $x_1+x_2 \leq b$,

$$Tw(x_1,x_2) = \alpha\left[q_1x_1 + \Gamma w(x_1,x_2)\right], \tag{A10}$$

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

$$Tw(x_1,x_2) = c_1^G + \alpha\left[q_1(x_1-1) + \Gamma w(x_1-1,x_2)\right], \tag{A11}$$

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

$$Tw(x_1,x_2) = c_2^G + \alpha\left[q_1x_1 + \Gamma w(x_1,x_2-1)\right], \tag{A12}$$

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

$$Tw(x_1,x_2) = \min\left\{c_1^G + \alpha\left[q_1(x_1-1) + \Gamma w(x_1-1,x_2)\right],\; c_2^G + \alpha\left[q_1x_1 + \Gamma w(x_1,x_2-1)\right]\right\}. \tag{A13}$$

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 (A8), 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 A2. *Suppose that (A8) 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 A2 states that the decision of whether a stage 1 or stage 2 patient should be discharged early 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. 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. Nest, we identify some conditions under which that would be the case.

Conditions for the optimality of the greedy policy. 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

$$c_1 = \frac{\alpha q_1(1-\alpha r_2)}{(1-\alpha r_1)(1-\alpha r_2)-\alpha^2 p_1 q_2},\ c_2 = \frac{\alpha^2 q_1 q_2}{(1-\alpha r_1)(1-\alpha r_2)-\alpha^2 p_1 q_2}, \tag{A14}$$

using the same argument used to obtain c_i^G's in (A5).

Proposition A3. *Suppose that (A8) 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*

$$c_1^G - c_1 < c_2^G - c_2 \text{ and } L_1 \geq L_2, \tag{A15}$$

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

$$c_1^G - c_1 \geq c_2^G - c_2 \text{ and } L_1 \leq L_2, \tag{A16}$$

where L_i is given by (4) for $i=1,2$.

Proposition A3 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.

Appendix A3: Proofs of the analytical results in Appendix A2.

In this section, we prove Propositions A1, A2, and A3 for infinite-horizon expected total discounted cost problem. 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 A3. The first difference operators D_1 and D_2 are defined as

$$D_1 w(x_1,x_2) = w(x_1+1,x_2) - w(x_1,x_2),\ D_2 w(x_1,x_2) = w(x_1,x_2+1) - w(x_1,x_2) \text{ where } x_1+x_2 \leq b.$$

Definition A4. The operator Δ is defined as

$$\Delta w(x_1,x_2) = w(x_1+1,x_2) - w(x_1,x_2+1) = D_1 w(x_1,x_2) - D_2 w(x_1,x_2) \text{ where } x_1+x_2 \leq b.$$

Definition A5. The second difference operators $D_{11} = D_1D_1, D_{22} = D_2D_2$ and $D_{12} = D_1D_2 = D_2D_1$ are defined as

$$
\begin{aligned}
D_{11}w(x_1,x_2) &= D_1w(x_1+1,x_2) - D_1w(x_1,x_2) = w(x_1+2,x_2) - 2w(x_1+1,x_2) + w(x_1,x_2),\\
D_{22}w(x_1,x_2) &= D_2w(x_1,x_2+1) - D_2w(x_1,x_2) = w(x_1,x_2+2) - 2w(x_1,x_2+1) + w(x_1,x_2),\\
D_{12}w(x_1,x_2) &= D_1w(x_1,x_2+1) - D_1w(x_1,x_2) = D_2w(x_1+1,x_2) - D_2w(x_1,x_2)\\
&= w(x_1+1,x_2+1) - w(x_1,x_2+1) - w(x_1+1,x_2) + w(x_1,x_2) \text{ where } x_1+x_2 \le b-1.
\end{aligned}
$$

A3.1. Proof of Proposition A1

Lemma A1. *For $x_1 + x_2 \le b$, we have $D_1v_\alpha(x_1,x_2) \le c_1^G,\ D_2v_\alpha(x_1,x_2) \le c_2^G$.*

Proof of Lemma A1: 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,

$$v_\alpha^{\pi_1}(x_1+1,x_2) = c_1^G + v_\alpha(x_1,x_2) \ge v_\alpha(x_1+1,x_2) \Rightarrow v_\alpha(x_1+1,x_2) - v_\alpha(x_1,x_2) \le c_1^G.$$

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,

$$v_\alpha^{\pi_2}(x_1,x_2+1) = c_2^G + v_\alpha(x_1,x_2) \ge v_\alpha(x_1,x_2+1) \Rightarrow v_\alpha(x_1,x_2+1) - v_\alpha(x_1,x_2) \le c_2^G.$$

□

Lemma A2. *Suppose that (A8) holds and for any $w : \mathcal{S} \to \mathbb{R}$, we have (i) $D_1w(i,j) \le c_1^G$ for $i + j \le b$, and (ii) $D_2w(i,j) \le c_2^G$ for $i+j \le b$. Then, for $x+y \le b-1$, $D_1\Gamma w(x,y) \le \frac{c_1^G}{\alpha} - q_1$, and $D_2\Gamma w(x,y) \le \frac{c_2^G}{\alpha}$.*

Proof of Lemma A2: Let $x+y \le b-1$. Thus, $\Gamma w(x+1,y)$ and $\Gamma w(x,y+1)$ are well defined according to Definition A1.

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

$$\Gamma w(x+1,y) = \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y)\,[q_1w(i,j) + r_1w(i+1,j) + p_1w(i,j+1)].$$

Then,

$$D_1\Gamma w(x,y) = \Gamma w(x+1,y) - \Gamma w(x,y),$$
$$= \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y)\,[r_1 D_1 w(i,j) + p_1 D_2 w(i,j)] \le r_1 c_1^G + p_1 c_2^G \le \frac{c_1^G}{\alpha} - q_1,$$

where again the first inequality follows from the lemma assumptions (i) and (ii), and the second inequality follows from Condition (A8), which is another lemma assumption.

Similarly, we can write $\Gamma w(x,y+1)$

$$\Gamma w(x,y+1) = \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y)\,[p_2 w(i,j) + q_2 w(i+1,j) + r_2 w(i,j+1)],$$

and then

$$D_2\Gamma w(x,y) = \Gamma w(x,y+1) - \Gamma w(x,y)$$
$$= \sum_{i=0}^{x+y+1} \sum_{j=0}^{x+y+1-i} P(i,j|x,y)\,[q_2 D_1 w(i,j) + r_2 D_2 w(i,j)] \le q_2 c_1^G + r_2 c_2^G \le \frac{c_2^G}{\alpha},$$

where the first inequality follows from the lemma assumptions (i) and (ii), and the second inequality follows from Condition (A8), which is another lemma assumption.

□

Proof of Proposition A1: From Lemmas A1 and A2, we have for any $x+y \le b-1$,

$$D_1\Gamma v_\alpha(x,y) \le \frac{c_1^G}{\alpha} - q_1, \tag{A17}$$

$$D_2\Gamma v_\alpha(x,y) \le \frac{c_2^G}{\alpha}. \tag{A18}$$

Then, for any $x_1 \ge a_1 \ge 1$, $x_2 \ge a_2 \ge 0$ and $x_1 - (a_1 - 1) + x_2 - a_2 \le b$,

$$V_\alpha(x_1,x_2,a_1-1,a_2) - V_\alpha(x_1,x_2,a_1,a_2) = -c_1^G + \alpha\,[q_1 + \Gamma v_\alpha(x_1-a_1+1,x_2) - \Gamma v_\alpha(x_1-a_1,x_2-a_2)]$$
$$= \alpha\left[-\frac{c_1^G}{\alpha} + q_1 + D_1\Gamma v(x_1-a_1,x_2-a_2)\right] \le 0, \tag{A19}$$

where the inequality follows from (A17). We can then conclude that $V_\alpha(x_1,x_2,a_1-1,a_2) \le V_\alpha(x_1,x_2,a_1,a_2)$, which implies that decreasing the number of stage 1 discharges does not increase the expected cost.

Similarly, for any $x_1 \ge a_1 \ge 0$, $x_2 \ge a_2 \ge 1$ and $x_1 - a_1 + x_2 - (a_2 - 1) \le b$,

$$V_\alpha(x_1,x_2,a_1,a_2-1) - V_\alpha(x_1,x_2,a_1,a_2) = -c_2^G + \alpha\,[\Gamma v_\alpha(x_1-a_1,x_2-a_2+1) - \Gamma v_\alpha(x_1-a_1,x_2-a_2)]$$
$$= \alpha\left[-\frac{c_2^G}{\alpha} + D_2\Gamma v(x_1-a_1,x_2-a_2)\right] \le 0, \tag{A20}$$

where the inequality follows from (A18). We can then conclude that $V_\alpha(x_1,x_2,a_1,a_2-1) \le V_\alpha(x_1,x_2,a_1,a_2)$, which implies that decreasing the number of stage 2 discharges does not increase the expected cost, either. Then, the result follows.

□

A3.2. Proof of Proposition A2

Lemma A3. *If (A8) holds, then $c_i \le c_i^G$ for $i=1,2$, where c_i^G is given in* (A5) *and c_i is given in* (A14).

Proof of Lemma A3: Let $f_i(0)=c_i^G$ for $i=1,2$, and for $n\ge 0$ we define

$$f_1(n+1)=\alpha(q_1+r_1f_1(n)+p_1f_2(n)),\ f_2(n+1)=\alpha(q_2f_1(n)+r_2f_2(n)). \tag{A21}$$

Next, by induction, we show that $f_i(n+1)\le f_i(n)$ for all $n\ge 0$: From (A8) and (A21), 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,

$$f_1(k+2)=\alpha(q_1+r_1f_1(k+1)+p_1f_2(k+1))\le \alpha(q_1+r_1f_1(k)+p_1f_2(k))=f_1(k+1),$$

and

$$f_2(k+2)=\alpha(q_2f_1(k+1)+r_2f_2(k+1))\le \alpha(q_2f_1(k)+r_2f_2(k))=f_2(k+1).$$

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 (A21), we have

$$f_1=\alpha(q_1+r_1f_1+p_1f_2),\ f_2=\alpha(q_2f_1+r_2f_2).$$

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 A6. Let $\mathcal{V}_b$ be a set of functions such that if $w\in\mathcal{V}_b$, then

$$c_1\le D_1w(i,j)\le c_1^G,\ c_2\le D_2w(i,j)\le c_2^G \text{ for all } i+j\le b. \tag{A22}$$

Lemma A4. *If (A8) 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 A4: (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 A2, 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 $x_1+x_2 \le b-1$,

$$D_1 \Gamma w(x_1,x_2) = \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)\left[r_1 D_1 w(i,j) + p_1 D_2 w(i,j)\right] \ge r_1 c_1 + p_1 c_2 = \frac{c_1}{\alpha} - q_1,$$
$$D_2 \Gamma w(x_1,x_2) = \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)\left[q_2 D_1 w(i,j) + r_2 D_2 w(i,j)\right] \ge q_2 c_1 + r_2 c_2 = \frac{c_2}{\alpha},$$

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 A2.

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

$$c_1 \le D_1 T w(x_1,x_2) \le c_1^G,\ c_2 \le D_2 T w(x_1,x_2) \le c_2^G \text{ for } x_1+x_2 \le b. \tag{A23}$$

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 (A23),

$$c_2^G + T w(x_1+1,x_2-1) - T w(x_1,x_2) = c_2^G + D_1 T w(x_1,x_2-1) - D_2 T w(x_1,x_2-1) \ge c_1,$$

and $c_1^G \ge c_1$, then we can conclude that $D_1 T w(x_1,x_2) \ge c_1$.

Similarly,

$$D_2 T w(x_1,x_2) = \min\left\{c_1^G + T w(x_1-1,x_2+1) - T w(x_1,x_2),\ c_2^G\right\} \le c_2^G.$$

Furthermore, since $D_1 T w(x_1-1,x_2) \le c_1^G$ and $D_2 T w(x_1-1,x_2) \ge c_2$ from (A23),

$$c_1^G + T w(x_1-1,x_2+1) - T w(x_1,x_2) = c_1^G - D_1 T w(x_1-1,x_2) + D_2 T w(x_1-1,x_2) \ge c_2,$$

and $c_2^G \ge c_2$, we can conclude that $D_2 T w(x_1,x_2) \ge c_2$.

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 A7. 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: $D_{11}w(x_1,x_2)D_{22}w(y_1,y_2) + D_{11}w(y_1,y_2)D_{22}w(x_1,x_2) - 2D_{12}w(x_1,x_2)D_{12}w(y_1,y_2) \ge 0,$
for $x_1 + x_2 \le b-1$ 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 A5. *Suppose that (A8) holds, and $w \in \mathcal{V}$. Then,*
(i) for $x_1 + x_2 \le b-2$ and $y_1 + y_2 \le b-2$,

$$D_{11}\Gamma w(x_1,x_2) \ge 0,\ D_{22}\Gamma w(x_1,x_2) \ge 0,\ \text{and } D_{12}\Gamma w(x_1,x_2) \ge 0, \tag{A24}$$

$$D_{11}\Gamma w(x_1,x_2)D_{22}\Gamma w(y_1,y_2) + D_{11}\Gamma w(y_1,y_2)D_{22}\Gamma w(x_1,x_2) - 2D_{12}\Gamma w(x_1,x_2)D_{12}\Gamma w(y_1,y_2) \ge 0, \tag{A25}$$

$$D_{11}\Gamma w(x_1,x_2) + D_{22}\Gamma w(x_1,x_2) - 2D_{12}\Gamma w(x_1,x_2) \ge 0. \tag{A26}$$

(ii) $Tw \in \mathcal{V}$.

Proof of Lemma A5 (i): Establishing (A24), (A25), (A26).

Proof of (A24): Using the fact that patients' health conditions change independently of each other, we have for $x_1 + x_2 \le b-2$,

$$\begin{aligned}
\Gamma w(x_1+2,x_2) = & \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)\Big[q_1^2 w(i,j) + 2q_1r_1w(i+1,j) \\
& + 2q_1p_1w(i,j+1) + r_1^2w(i+2,j) + 2r_1p_1w(i+1,j+1) + p_1^2w(i,j+2)\Big], \\
\Gamma w(x_1+1,x_2) = & \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)\left[q_1w(i,j) + r_1w(i+1,j) + p_1w(i,j+1)\right], \\
\Gamma w(x_1,x_2) = & \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)w(i,j).
\end{aligned}$$

Then, we have

$$D_{11}\Gamma w(x_1,x_2) = \Gamma w(x_1+2,x_2) - 2\Gamma w(x_1+1,x_2) + \Gamma w(x_1,x_2)$$
$$= \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)g_{11}(i,j),$$

where

$$g_{11}(i,j) = \Big[(r_1+p_1)^2 w(i,j) - 2r_1(r_1+p_1)w(i+1,j) - 2p_1(r_1+p_1)w(i,j+1) + r_1^2 w(i+2,j)$$
$$+ 2r_1p_1w(i+1,j+1) + p_1^2 w(i,j+2)\Big] = r_1^2 D_{11}w(i,j) + 2p_1r_1D_{12}w(i,j) + p_1^2 D_{22}w(i,j).$$

Similarly, we have

$$D_{22}\Gamma w(x_1,x_2) = \Gamma w(x_1,x_2+2) - 2\Gamma w(x_1,x_2+1) + \Gamma w(x_1,x_2)$$
$$= \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)g_{22}(i,j),$$

$$D_{12}\Gamma w(x_1,x_2) = \Gamma w(x_1+1,x_2+1) - \Gamma w(x_1+1,x_2) - \Gamma w(x_1,x_2+1) + \Gamma w(x_1,x_2)$$
$$= \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)g_{12}(i,j),$$

where

$$g_{22}(i,j) = q_2^2 D_{11}w(i,j) + 2q_2r_2D_{12}w(i,j) + r_2^2 D_{22}w(i,j),$$
$$g_{12}(i,j) = r_1q_2D_{11}w(i,j) + (r_1r_2+p_1q_2)D_{12}w(i,j) + p_1r_2D_{22}w(i,j).$$

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 (A25): 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 $y_1+y_2\leq b-2$,

$$D_{11}\Gamma w(x_1,x_2)D_{22}\Gamma w(y_1,y_2) + D_{11}\Gamma w(y_1,y_2)D_{22}\Gamma w(x_1,x_2) - 2D_{12}\Gamma w(x_1,x_2)D_{12}\Gamma w(y_1,y_2)$$
$$= \sum_{i_1=0}^{x_1+x_2+1}\sum_{i_2=0}^{x_1+x_2+1-i_1}\sum_{j_1=0}^{y_1+y_2+1}\sum_{j_2=0}^{y_1+y_2+1-j_1} P(i_1,i_2|x_1,x_2)P(j_1,j_2|y_1,y_2)g(i_1,i_2,j_1,j_2),$$

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\qquad \times \Big[q_2^2 D_{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_2^2 D_{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\qquad \times \Big[r_1^2 D_{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]\\
&\quad + \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]\\
&\quad - \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]\\
&\quad + \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]\\
&\quad - \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]\\
&\quad + \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]\\
&\quad - \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) + 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\left[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)\right].
\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,

$$D_{11}\Gamma w(x_1,x_2)D_{22}\Gamma w(y_1,y_2) + D_{11}\Gamma w(y_1,y_2)D_{22}\Gamma w(x_1,x_2) - 2D_{12}\Gamma w(x_1,x_2)D_{12}\Gamma w(y_1,y_2) \geq 0.$$

Proof of (A26): 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 $i_1+i_2 \leq x_1+x_2+1 \leq b-1$,

$$
\begin{aligned}
f(i_1,i_2) &= g_{11}(i_1,i_2) + g_{22}(i_1,i_2) - 2g_{12}(i_1,i_2)\\
&= (r_1-q_2)^2D_{11}w(i_1,i_2) + 2(r_1-q_2)(p_1-r_2)D_{12}w(i_1,i_2) + (p_1-r_2)^2D_{22}w(i_1,i_2).
\end{aligned}
$$

Since Condition 2 holds for w, we have for $i_1+i_2 \le b-1$,

$$D_{11}w(i_1,i_2)D_{22}w(i_1,i_2)+D_{11}w(i_1,i_2)D_{22}w(i_1,i_2)-2D_{12}w(i_1,i_2)D_{12}w(i_1,i_2)\ge 0$$
$$\Rightarrow D_{12}w(i_1,i_2)\le \sqrt{D_{11}w(i_1,i_2)D_{22}w(i_1,i_2)}.$$

Then, if $(r_1-q_2)(p_1-r_2)<0$, we have

$$f(i_1,i_2)\ge (r_1-q_2)^2D_{11}w(i_1,i_2)+(p_1-r_2)^2D_{22}w(i_1,i_2)+2(r_1-q_2)(p_1-r_2)\sqrt{D_{11}w(i_1,i_2)D_{22}w(i_1,i_2)}$$
$$=\left[(r_1-q_2)\sqrt{D_{11}w(i_1,i_2)}+(p_1-r_2)\sqrt{D_{22}w(i_1,i_2)}\right]^2\ge 0.$$

On the other hand, if $(r_1-q_2)(p_1-r_2)\ge 0$, then since $D_{11}w(i_1,i_2), D_{22}w(i_1,i_2), D_{12}w(i_1,i_2)$ for $i_1+i_2\le b-1$ are nonnegative (from Condition 1), we have

$$f(i_1,i_2)=(r_1-q_2)^2D_{11}w(i_1,i_2)+2(r_1-q_2)(p_1-r_2)D_{12}w(i_1,i_2)+(p_1-r_2)^2D_{22}w(i_1,i_2)\ge 0.$$

Thus, $f(i_1,i_2)\ge 0$ for all $i_1+i_2\le b-1$, and we can conclude that

$$D_{11}\Gamma w(x_1,x_2)+D_{22}\Gamma w(x_1,x_2)-2D_{12}\Gamma w(x_1,x_2)\ge 0.$$

Proof of Lemma A5 (ii): We have already established in Lemma A4 that $Tw\in\mathcal{V}_b$, then, we only need to show that Conditions 1 to 3 given in Definition A7 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 A2.

Case 1: When $x_1+x_2+2\le b$, $Tw(i,j)=\alpha\left[q_1 i+\Gamma w(i,j)\right]$ 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,

$$D_{11}Tw(x_1,x_2)=Tw(x_1+2,x_2)-2Tw(x_1+1,x_2)+Tw(x_1,x_2),$$

and

$$\alpha\left[\Gamma w(x_1+2,x_2)-2\Gamma w(x_1+1,x_2)+\Gamma w(x_1,x_2)\right]=\alpha D_{11}\Gamma w(x_1,x_2).$$

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,

$$Tw(x_1+2,x_2)=\min\left\{c_1^G+\alpha\left[q_1(x_1+1)+\Gamma w(x_1+1,x_2)\right],\ c_2^G+\alpha\left[q_1(x_1+2)+\Gamma w(x_1+2,x_2-1)\right]\right\},$$
$$Tw(x_1+1,x_2+1)=\min\left\{c_1^G+\alpha\left[q_1x_1+\Gamma w(x_1,x_2+1)\right],\ c_2^G+\alpha\left[q_1(x_1+1)+\Gamma w(x_1+1,x_2)\right]\right\},$$
$$Tw(x_1,x_2+2)=\min\left\{c_1^G+\alpha\left[q_1(x_1-1)+\Gamma w(x_1-1,x_2+2)\right],\ c_2^G+\alpha\left[q_1x_1+\Gamma w(x_1,x_2+1)\right]\right\}.$$

Then we have,

$$
\begin{aligned}
D_{11}Tw(x_1,x_2) = \min\Big\{ & c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right], \\
& c_2^G - \alpha D_2\Gamma w(x_1,x_2) + \alpha\left[\Delta\Gamma w(x_1+1,x_2-1) - \Delta\Gamma w(x_1,x_2)\right] \Big\}, \\
D_{12}Tw(x_1,x_2) = \min\Big\{ & c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right], \quad c_2^G - \alpha D_2\Gamma w(x_1,x_2) \Big\}, \\
D_{22}Tw(x_1,x_2) = \min\Big\{ & c_1^G - \alpha\left[q_1 + D_1\Gamma w(x_1,x_2)\right] + \alpha\left[\Delta\Gamma w(x_1,x_2) - \Delta\Gamma w(x_1-1,x_2+1)\right], \\
& c_2^G - \alpha D_2\Gamma w(x_1,x_2) \Big\}.
\end{aligned}
$$

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 A2, $Tw(x_1+2,x_2) = c_1^G + \alpha[q_1(x_1+1) + \Gamma w(x_1+1,x_2)]$. 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 A2, $Tw(x_1,x_2+2) = c_2^G + \alpha[q_1x_1 + \Gamma w(x_1,x_2+1)]$. 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 (A24) 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 A7 we have $w \in \mathcal{V}_b$, and using Lemma A4(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 (A26) 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, $D_{11}T(x_1,x_2)\geq 0,\ D_{22}T(x_1,x_2)\geq 0,\ D_{12}T(x_1,x_2)\geq 0$ 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 $x_1+x_2\leq b-2$,

$$\begin{aligned}&D_{11}Tw(x_1,x_2)+D_{22}Tw(x_1,x_2)-2D_{12}Tw(x_1,x_2)\\&\qquad=\alpha\big[D_{11}\Gamma w(x_1,x_2)+D_{22}\Gamma w(x_1,x_2)-2D_{12}\Gamma w(x_1,x_2)\big]\geq 0,\end{aligned}$$

where the inequality follows from (A26) 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 (A25) 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

$$D_{11}Tw(x_1,x_2)D_{22}Tw(y_1,y_2)+D_{11}Tw(y_1,y_2)D_{22}Tw(x_1,x_2)-2D_{12}Tw(x_1,x_2)D_{12}Tw(y_1,y_2)\geq 0.$$

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 A6. *Suppose that (A8) 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 A6: (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

$$c_1\leq D_1v_n(x_1,x_2)\leq c_1^G,\ \ c_2\leq D_2v_n(x_1,x_2)\leq c_2^G \quad \text{for } x_1+x_2\leq b,$$

$$D_{11}v_n(x_1,x_2)\geq 0,\ \ D_{22}v_n(x_1,x_2)\geq 0, \text{ and } D_{12}v_n(x_1,x_2)\geq 0 \quad \text{for } x_1+x_2\leq b-1$$

$$D_{11}v_n(x_1,x_2)D_{22}v_n(y_1,y_2)+D_{11}v_n(y_1,y_2)D_{22}v_n(x_1,x_2)-2D_{12}v_n(x_1,x_2)D_{12}v_n(y_1,y_2)\geq 0,$$
$$\text{for } x_1+x_2\leq b-1 \text{ and } y_1+y_2\leq b-1,$$

$$D_{11}v_n(x_1,x_2)+D_{22}v_n(x_1,x_2)-2D_{12}v_n(x_1,x_2)\geq 0 \quad \text{for } x_1+x_2\leq b-1.$$

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 (A9), δ is optimal.

(iii) **Preservation.** This follows immediately from Lemma A5.

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 A5 (i) that $\Gamma v_\alpha(x_1,x_2)$ satisfies (A26), 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 $x_1+x_2 \le b-1$ and $x_1>0$,

$$\begin{aligned}
&\Delta\Gamma v_\alpha(x_1,x_2)-\Delta\Gamma v_\alpha(x_1-1,x_2+1)\\
&\qquad=\Gamma v_\alpha(x_1+1,x_2)-\Gamma v_\alpha(x_1,x_2+1)-\Gamma v_\alpha(x_1,x_2+1)+\Gamma v_\alpha(x_1-1,x_2+2)\\
&\qquad\qquad=D_{11}\Gamma v_\alpha(x_1-1,x_2)+D_{22}\Gamma v_\alpha(x_1-1,x_2)-2D_{12}\Gamma v_\alpha(x_1-1,x_2)\ge 0.
\end{aligned}$$

□

Proof of Proposition A2: Define function $\delta_n(i)=\Delta\Gamma v_\alpha(i,n-i)$ for $1\le i\le n\le b-1$. Then, from Lemma A6 (b), we have $\delta_n(i)\ge\delta_n(i-1)$. Thus, for fixed n, $\delta_n(i)$ is non-decreasing in i for $1\le i\le n$.

From the optimality equations (A9) and Definition A2, 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

$$c_1^G+\alpha\left[q_1(x_1-1)+\Gamma v_\alpha(x_1-1,x_2)\right]<c_2^G+\alpha\left[q_1x_1+\Gamma v_\alpha(x_1,x_2-1)\right],$$

which can equivalently be written as

$$\delta_{b-1}(x_1-1)>\frac{c_1^G-c_2^G}{\alpha}-q_1. \tag{A27}$$

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

$$x_\alpha^*=\min\left\{x_1: 1\le x_1\le b \text{ and } \delta_{b-1}(x_1-1)>\frac{c_1^G-c_2^G}{\alpha}-q_1\right\}.$$

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 $\delta_{b-1}(x_1-1)\le\frac{c_1^G-c_2^G}{\alpha}-q_1$ 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 $x_1\ge x^*$, 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.

□

A3.3. Proof of Proposition A3

Definition A8. Define an operator F on w by

$$Fw(i,j)=(r_1-q_2)D_1w(i,j)+(p_1-r_2)D_2w(i,j).$$

Definition A9. Let $\mathcal{V}_1$ be a set of functions such that if $w\in\mathcal{V}_1$, then $w\in\mathcal{V}$ and

Condition 4: $\Delta w(x_1,x_2)=w(x_1+1,x_2)-w(x_1,x_2+1)\ge c_1^G-c_2^G$, for $x_1+x_2\le b$,

Condition 5: $Fw(x_1,x_2)>\dfrac{c_1^G-c_2^G}{\alpha}-q_2$, for $x_1+x_2\le b$.

Lemma A7. *If (A8) and (A15) 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* $x_1+x_2 \le b-1$.
(b) $Tw \in \mathcal{V}_1$.
(c) $v_\alpha \in \mathcal{V}_1$.

Proof of Lemma A7: From (4), we can show that (A15) is equivalent to

$$c_1 - c_2 > c_1^G - c_2^G,\ q_1 \le p_2. \tag{A28}$$

(a) By conditioning on how x_1 stage 1 patients and x_2 stage 2 patients evolve, we have for $x_1+x_2 \le b-1$,

$$\Delta\Gamma w(x_1,x_2) = \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)Fw(i,j) > \frac{c_1^G - c_2^G}{\alpha} - q_1,$$

where the inequality follows from Condition 5 for $w \in \mathcal{V}_1$ and $i+j \le x_1+x_1+1 \le b$.
(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 A2.
Case 1: If $x_1+x_2 \le b-1$, 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

$$c_1^G + \alpha[q_1x_1 + \Gamma w(x_1,x_2)] < c_2^G + \alpha[q_1(x_1+1) + \Gamma w(x_1+1,x_2-1)].$$

Then,

$$\begin{aligned}Tw(x_1+1,x_2) = \min\Big\{&c_1^G + \alpha[q_1x_1 + \Gamma w(x_1,x_2)], c_2^G + \alpha[q_1(x_1+1) + \Gamma w(x_1+1,x_2-1)]\Big\}\\ &= c_1^G + \alpha[q_1x_1 + \Gamma w(x_1,x_2)] = c_1^G + Tw(x_1,x_2),\end{aligned}$$

and thus $D_1[Tw(x_1,x_2)] = Tw(x_1+1,x_2) - Tw(x_1,x_2) = c_1^G$.
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,

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

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,

$$\Delta\left[Tw(x_1,x_2)\right] = D_1\left[Tw(x_1,x_2)\right] - D_2\left[Tw(x_1,x_2)\right] = \Delta\left[Tw(x_1-1,x_2)\right]. \tag{A29}$$

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 $D_2\left[Tw(x_1,x_2)\right] = c_1^G - \Delta\left[Tw(x_1-1,x_2)\right]$, and

$$Tw(x_1+1, x_2) = c_1^G + \alpha\left[q_1 x_1 + \Gamma w(x_1,x_2)\right] = c_1^G - \Delta\left[Tw(x_1,x_2)\right],$$

and thus $D_1\left[Tw(x_1,x_2)\right] = c_1^G$. Then,

$$\Delta\left[Tw(x_1,x_2)\right] = \Delta\left[Tw(x_1-1,x_2)\right].$$

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 $D_1\left[Tw(x_1,x_2)\right] = c_1^G$, and

$$Tw(x_1, x_2+1) = c_2^G + \alpha\left[q_1(x_1+1) + \Gamma w(x_1,x_2)\right],$$

and thus $D_2\left[Tw(x_1,x_2)\right] = c_2^G$. Then,

$$\Delta\left[Tw(x_1,x_2)\right] = c_1^G - c_2^G.$$

Proof of Condition 4:

Case 1: If $x_1 + x_2 \le b-1$, 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 $\Delta\left[Tw(x_1,x_2)\right] = \Delta\left[Tw(x_1-1,x_2)\right] > c_1^G - c_2^G$, 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, $\Delta Tw(x_1,x_2) \ge c_1^G - c_2^G$ for all $x_1 + x_2 \le b$.

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 (A28). 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 (A30), we have $p_1 > r_2$ using the fact that $p_1+q_1+r_1 = p_2+q_2+r_2 = 1$. Then, $p_1q_2 > r_1r_2$. Next we consider four different cases as before.

Case 1: If $x_1+x_2 \leq b-1$,

$$F[Tw(x_1,x_2)] = (r_1-q_2)D_1Tw(x_1,x_2) + (p_1-r_2)D_2Tw(x_1,x_2) = \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)H(i,j),$$

where

$$H(i,j) = (r_1-q_2)\alpha\left[q_1 + r_1D_1w(i,j) + p_1D_2w(i,j)\right] + (p_1-r_2)\alpha\left[q_2D_1w(i,j) + r_2D_2w(i,j)\right].$$

We have,

$$\begin{aligned}
\frac{H(i,j)}{\alpha} &= (r_1-q_2)\left[q_1 + r_1D_1w(i,j) + p_1D_2w(i,j)\right] + (p_1-r_2)\left[q_2D_1w(i,j) + r_2D_2w(i,j)\right] \\
&= -q_1(r_2+q_2) + (r_1+r_2)\left[q_1 + (r_1-q_2)D_1w(i,j) + (p_1-r_2)D_2w(i,j)\right] \\
&\quad + (p_1q_2 - r_1r_2)\left[D_1w(i,j) - D_2w(i,j)\right] \\
&= -q_1(r_2+q_2) + (r_1+r_2)[q_1 + Fw(i,j)] + (p_1q_2-r_1r_2)\Delta w(i,j) \\
&\geq -q_1(r_2+q_2) + (r_1+r_2)\left(\frac{c_1^G-c_2^G}{\alpha}\right) + (p_1q_2-r_1r_2)(c_1^G-c_2^G),
\end{aligned}$$

where the inequality follows from Conditions 4 and 5 for $w \in \mathcal{V}_1$. Then,

$$\begin{aligned}
\alpha H(i,j) - \left[c_1^G - c_2^G - \alpha q_1\right] &\geq \alpha q_1\left[(1-\alpha(r_2+q_2)\right] + \left[\alpha^2(r_1+r_2) + \alpha^2(p_1q_2-r_1r_2) - 1\right](c_1^G-c_2^G) \\
&= \alpha q_1(1-\alpha r_2 - \alpha q_2) - \left[1 - \alpha(r_1+r_2) + \alpha^2 r_1r_2 - \alpha^2 p_1q_2\right](c_1^G-c_2^G)
\end{aligned}$$

$$
\begin{aligned}
&= \alpha q_1(1-\alpha r_2-\alpha q_2) - \left[(1-\alpha r_1)(1-\alpha r_2)-\alpha^2 p_1 q_2\right](c_1^G-c_2^G)\\
&= \left[(1-\alpha r_1)(1-\alpha r_2)-\alpha^2 p_1 q_2\right]\left[\frac{\alpha q_1(1-\alpha r_2-\alpha q_2)}{(1-\alpha r_1)(1-\alpha r_2)-\alpha^2 p_1 q_2} - (c_1^G-c_2^G)\right]\\
&= \left[(1-\alpha r_1)(1-\alpha r_2)-\alpha^2 p_1 q_2\right]\left[(c_1-c_2)-(c_1^G-c_2^G)\right] > 0,
\end{aligned}
$$

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

$$
F[Tw(x_1,x_2)] = \sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2)H(i,j) > \frac{c_1^G-c_2^G}{\alpha} - q_1.
$$

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,

$$
\begin{aligned}
F[Tw(x_1,x_2)] &= (r_1-q_2)D_1[Tw(x_1,x_2)] + (p_1-r_2)D_2[Tw(x_1,x_2)]\\
&= (r_1-q_2)c_1^G + (p_1-r_2)\left[c_1^G - \Delta[Tw(x_1-1,x_2)]\right]\\
&= (p_2-q_1)c_1^G - (p_1-r_2)\Delta[Tw(x_1-1,x_2)]\\
&\ge (p_2-q_1)D_1[Tw(x_1-1,x_2)] - (p_1-r_2)\Delta[Tw(x_1-1,x_2)]\\
&= (r_1-q_2)D_1[Tw(x_1-1,x_2)] + (p_1-r_2)D_2[Tw(x_1-1,x_2)]\\
&= F[Tw(x_1-1,x_2)] > \frac{c_1^G-c_2^G}{\alpha} - q_1,
\end{aligned}
$$

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,

$$
\begin{aligned}
F[Tw(x_1,x_2)] &= (r_1-q_2)D_1Tw(x_1,x_2) + (p_1-r_2)D_2Tw(x_1,x_2)\\
&= (r_1-q_2)c_1^G + (p_1-r_2)c_2^G = (r_2-p_1)(c_1^G-c_2^G) + (p_2-q_1)c_1^G\\
&\ge (r_2-p_1)(c_1-c_2) + (p_2-q_1)c_1 = \frac{c_1-c_2}{\alpha} - q_1 > \frac{c_1^G-c_2^G}{\alpha} - q_1,
\end{aligned}
$$

where the first inequality follows from $r_2 - p_1 < 0$, $c_1 - c_2 > c_1^G - c_2^G$, $p_2 - q_1 \ge 0$ and $c_1^G \ge c_1$.

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 A6 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 $x_1+x_2 \le b$, 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 (A28).

(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 A10. Let $\mathcal{V}_2$ be a set of functions such that if $w \in \mathcal{V}_2$, then $w \in \mathcal{V}$ and

$$\text{Condition 6: } \Delta w(x_1,x_2) = w(x_1+1,x_2) - w(x_1,x_2+1) \le c_1^G - c_2^G \qquad \text{for } x_1+x_2 \le b,$$
$$\text{Condition 7: } Fw(x_1,x_2) \le \frac{c_1^G - c_2^G}{\alpha} - q_2 \qquad \text{for } x_1+x_2 \le b.$$

Lemma A8. *If* (A8) *and (A16) hold, then for any function* $w \in \mathcal{V}_2$,
(a) $\Delta\Gamma w(x_1,x_2) \le \frac{c_1^G - c_2^G}{\alpha} - q_1$, *for all* $x_1+x_2 \le b-1$.
(b) $Tw \in \mathcal{V}_2$.
(c) $v_\alpha \in \mathcal{V}_2$.

Proof of Lemma A8: From (4), we can show that (A16) is equivalent to

$$c_1 - c_2 \le c_1^G - c_2^G,\ q_1 \ge p_2. \tag{A30}$$

(a) By conditioning on how x_1 stage 1 patients and x_2 stage 2 patients evolve, we have for $x_1+x_2 \le b-1$,

$$\Delta\Gamma w(x_1,x_2) = \sum_{i=0}^{x_1+x_2+1} \sum_{j=0}^{x_1+x_2+1-i} P(i,j|x_1,x_2) Fw(i,j) \le \frac{c_1^G - c_2^G}{\alpha} - q_1,$$

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 $\Delta[Tw(x_1,x_2)]$ for the four different cases described in Definition A2.
Case 1: If $x_1+x_2 \le b-1$, 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

$$\Delta\Gamma w(x_1,x_2-1) = \Gamma w(x_1+1,x_2-1) - \Gamma w(x_1,x_2) \le \frac{c_1^G - c_2^G}{\alpha} - q_1,$$

which is equivalent to

$$c_1^G + \alpha\left[q_1x_1 + \Gamma w(x_1,x_2)\right] \ge c_2^G + \alpha\left[q_1(x_1+1) + \Gamma w(x_1+1,x_2-1)\right].$$

Then,

$$\begin{aligned} Tw(x_1+1,x_2) &= \min\left\{c_1^G + \alpha\left[q_1x_1 + \Gamma w(x_1,x_2)\right],\ c_2^G + \alpha\left[q_1(x_1+1) + \Gamma w(x_1+1,x_2-1)\right]\right\} \\ &= c_2^G + \alpha\left[q_1(x_1+1) + \Gamma w(x_1+1,x_2-1)\right] = c_2^G + Tw(x_1+1,x_2-1), \end{aligned}$$

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,

$$Tw(x_1,x_2+1)=\min\left\{c_1^G+\alpha\left[q_1(x_1-1)+\Gamma w(x_1-1,x_2+1)\right],\ c_2^G+\alpha\left[q_1x_1+\Gamma w(x_1,x_2)\right]\right\}$$
$$=c_2^G+\alpha\left[q_1x_1+\Gamma w(x_1,x_2)\right]=c_2^G+Tw(x_1,x_2),$$

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

$$Tw(x_1+1,x_2)=c_1^G+\alpha\left[q_1x_1+\Gamma w(x_1,x_2)\right]=c_1^G+T(x_1,x_2),$$

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 A2, $Tw(x_1+1,x_2)$ 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

$$Tw(x_1,x_2+1)=c_2^G+\alpha\left[q_1x_1+\Gamma w(x_1,x_2)\right]=c_2^G+Tw(x_1,x_2).$$

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 $x_1+x_2\le b-1$, we have, $\Delta Tw(i,j)=\alpha[q_1+\Delta\Gamma w(x_1,x_2)]\le c_1^G-c_2^G$, where the inequality follows from part (a).

Cases 2 and 4: If $x_1+x_2=b$ and $x_2>0$, then $\Delta Tw(x_1,x_2)=\Delta\left[Tw(x_1,x_2-1)\right]\le c_1^G-c_2^G$, 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 $p_1\le r_2$,

$$\begin{aligned}
F[Tw(i,j)] &= (p_2-q_1)D_1[Tw(i,j)]-(p_1-r_2)\Delta[Tw(i,j)]\\
&\le (p_2-q_1)c_1-(p_1-r_2)(c_1^G-c_2^G)\\
&= (c_1^G-c_2^G)-(p_1+p_2+q_2)(c_1^G-c_2^G)+(p_2-q_1)c_1\\
&\le (c_1^G-c_2^G)-(p_1+p_2+q_2)(c_1-c_2)+(p_2-q_1)c_1\\
&= (c_1^G-c_2^G)-(c_1-c_2)+(p_1-r_2)c_2+(r_1-q_2)c_1\\
&= \frac{c_1^G-c_2^G}{\alpha}+(c_1^G-c_2^G)\left[1-\frac{1}{\alpha}\right]-(c_1-c_2)+\frac{c_1-c_2}{\alpha}-q_1\\
&= \frac{c_1^G-c_2^G}{\alpha}-q_1+\left[1-\frac{1}{\alpha}\right]\left[(c_1^G-c_2^G)-(c_1-c_2)\right]\le\frac{c_1^G-c_2^G}{\alpha}-q_1,
\end{aligned}$$

where the first inequality follows from $\Delta Tw(x_1,x_2)\le c_1^G-c_2^G$, (i.e., follows from that Condition 6 holds for Tw, as we established earlier), $D_1Tw(x_1,x_2)\ge c_1$ since $Tw\in\mathcal{V}_b$ and $p_2-q_1\le 0$ from (A30). The second inequality follows from $(c_1^G-c_2^G)-(c_1-c_2)\ge 0$ and $1-\frac{1}{\alpha}<0$.

(ii) If $p_1>r_2$ and since $q_1\ge p_2$ from (A30), we have $r_1<q_2$ using the fact that $p_1+q_1+r_1=p_2+q_2+r_2=1$. Then, $p_1q_2>r_1r_2$.

Case 1: If $x_1+x_2\le b-1$,

$$F[Tw(x_1,x_2)]=(r_1-q_2)D_1Tw(x_1,x_2)+(p_1-r_2)D_2Tw(x_1,x_2)$$
$$=\sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i}P(i,j|x_1,x_2)H(i,j),$$

where

$$H(i,j)=(r_1-q_2)\alpha\left[q_1+r_1D_1w(i,j)+p_1D_2w(i,j)\right]+(p_1-r_2)\alpha\left[q_2D_1w(i,j)+r_2D_2w(i,j)\right].$$

We have,

$$\begin{aligned}
\frac{H(i,j)}{\alpha}&=(r_1-q_2)\left[q_1+r_1D_1w(i,j)+p_1D_2w(i,j)\right]+(p_1-r_2)\left[q_2D_1w(i,j)+r_2D_2w(i,j)\right]\\
&=-q_1(r_2+q_2)+(r_1+r_2)\left[q_1+(r_1-q_2)D_1w(i,j)+(p_1-r_2)D_2w(i,j)\right]\\
&\quad+(p_1q_2-r_1r_2)\left[D_1w(i,j)-D_2w(i,j)\right]\\
&=-q_1(r_2+q_2)+(r_1+r_2)[q_1+Fw(i,j)]+(p_1q_2-r_1r_2)\Delta w(i,j)\\
&\le -q_1(r_2+q_2)+(r_1+r_2)\left(\frac{c_1^G-c_2^G}{\alpha}\right)+(p_1q_2-r_1r_2)(c_1^G-c_2^G),
\end{aligned}$$

where the inequality follows from the fact that Conditions 6 and 7 hold for $w\in\mathcal{V}_2$. Then,

$$\begin{aligned}
\alpha H(i,j)-\left[c_1^G-c_2^G-\alpha q_1\right]&\le \alpha q_1\Big[(1-\alpha(r_2+q_2)\Big]+\Big[\alpha(r_1+r_2)+\alpha^2(p_1q_2-r_1r_2)-1\Big](c_1^G-c_2^G)\\
&=\alpha q_1(1-\alpha r_2-\alpha q_2)-\left[1-\alpha(r_1+r_2)+\alpha^2r_1r_2-\alpha^2p_1q_2\right](c_1^G-c_2^G)\\
&=\alpha q_1(1-\alpha r_2-\alpha q_2)-\left[(1-\alpha r_1)(1-\alpha r_2)-\alpha^2p_1q_2\right](c_1^G-c_2^G)\\
&=\left[(1-\alpha r_1)(1-\alpha r_2)-\alpha^2p_1q_2\right]\left[\frac{\alpha q_1(1-\alpha r_2-\alpha q_2)}{(1-\alpha r_1)(1-\alpha r_2)-\alpha^2p_1q_2}-(c_1^G-c_2^G)\right]\\
&=\left[(1-\alpha r_1)(1-\alpha r_2)-\alpha^2p_1q_2\right]\left[(c_1-c_2)-(c_1^G-c_2^G)\right]\le 0,
\end{aligned}$$

since $(1-\alpha r_1)(1-\alpha r_2))-\alpha^2p_1q_2=(1-\alpha+\alpha q_1+\alpha p_1)(1-\alpha+\alpha p_2+\alpha q_2)-\alpha^2p_1q_2\ge 0$ and $c_1-c_2\le c_1^G-c_2^G$. Hence, $H(i,j)\le\frac{c_1^G-c_2^G}{\alpha}-q_1$, and

$$F[Tw(x_1,x_2)]=\sum_{i=0}^{x_1+x_2+1}\sum_{j=0}^{x_1+x_2+1-i}P(i,j|x_1,x_2)H(i,j)\le\frac{c_1^G-c_2^G}{\alpha}-q_1.$$

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,

$$\begin{aligned}
F[Tw(x_1,x_2)] &= (r_1-q_2)D_1[Tw(x_1,x_2)]+(p_1-r_2)D_2[Tw(x_1,x_2)]\\
&= (r_1-q_2)\left[c_2^G+\Delta[Tw(x_1,x_2-1)]\right]+(p_1-r_2)c_2^G\\
&= (r_1-q_2)\Delta[Tw(x_1,x_2-1)]+(p_2-q_1)c_2^G\\
&\le (r_1-q_2)\Delta[Tw(x_1,x_2-1)]+(p_2-q_1)D_2[Tw(x_1,x_2-1)]\\
&= (r_1-q_2)D_1[Tw(x_1,x_2-1)]+(p_1-r_2)D_2[Tw(x_1,x_2-1)]\\
&= F[Tw(x_1,x_2-1)]\le \frac{c_1^G-c_2^G}{\alpha}-q_1,
\end{aligned}$$

where the first inequality follows from $D_2[Tw(x_1,x_2-1)]\le c_2^G$ since $Tw\in\mathcal{V}_b$ and $q_1\ge p_2$, 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,

$$\begin{aligned}
F[Tw(x_1,x_2)] &= (r_1-q_2)D_1Tw(x_1,x_2)+(p_1-r_2)D_2Tw(x_1,x_2)\\
&= (r_1-q_2)c_1^G+(p_1-r_2)c_2^G=(r_2-p_1)(c_1^G-c_2^G)+(p_2-q_1)c_1^G\\
&\le (r_2-p_1)(c_1-c_2)+(p_2-q_1)c_1=\frac{c_1-c_2}{\alpha}-q_1\le\frac{c_1^G-c_2^G}{\alpha}-q_1,
\end{aligned}$$

where the first inequality follows from $r_2-p_1<0$, $c_1^G-c_2^G\ge c_1-c_2$, $p_2-q_1\le 0$, and $c_1^G\ge c_1$.

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 A6 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 $x_1+x_2\le b$, 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 A2.

(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 A3: (a) Suppose that (A8) and (A15) hold, then from Lemma A7(c), $v_\alpha\in\mathcal{V}_1$. It then follows from Lemma A7(a) that, for any $x_1+x_2\le b-1$,

$$\Delta\Gamma v_\alpha(x_1,x_2)>\frac{c_1^G-c_2^G}{\alpha}-q_1.$$

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 (A27).

(b) Suppose that (A8) and (A16) hold, then following Lemma A8(c), the optimal value function $v_\alpha\in\mathcal{V}_2$ and then we have from Lemma A8(a), for all $x_1+x_2\le b-1$

$$\Delta\Gamma v_\alpha(x_1,x_2)\le\frac{c_1^G-c_2^G}{\alpha}-q_1.$$

Hence, for all $x_1+x_2=b+1$, $x_1>0$, and $x_2>0$, we have $\delta_{b-1}(x_1-1)=\Delta\Gamma v_\alpha(x_1-1,b-x_1)\le\frac{c_1^G-c_2^G}{\alpha}-q_1$, and thus $a^*_\alpha(x_1,x_2)=(0,1)$ from (A27).

□

Appendix A4: Proof of the analytical results in Section 5

A4.1. Proof of Theorem 1

LEMMA A9. *If $\beta_i<\beta_i^G$ for both $i=1,2$, then there exists an $\alpha_0\in(0,1)$ such that (A8) holds for all $\alpha\in[\alpha_0,1]$.*

Proof of Lemma A9: Let $f_1(\alpha)=\alpha\big[q_1+r_1c_1^G+p_1c_2^G\big]-c_1^G$ and $f_2(\alpha)=\alpha(q_2c_1^G+r_2c_2^G)-c_2^G$ for $\alpha\in[0,1]$, where c_1^G and c_2^G as expressed in (A5) are continuous in α. Then $f_1(\alpha)$ and $f_2(\alpha)$ are both continuous in α. When $\alpha=1$, we have

$$c_1^G=\frac{q_1^G(p_2^G+q_2^G)}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G},\ c_2^G=\frac{q_1^Gq_2^G}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G}.$$

Then,

$$\begin{aligned}f_1(1)&=q_1+r_1c_1^G+p_1c_2^G-c_1^G=q_1(1-c_1^G)-p_1(c_1^G-c_2^G)\\&=\frac{q_1p_1^Gp_2^G}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G}-\frac{p_1q_1^Gp_2^G}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G}=\frac{p_1p_1^Gp_2^G(\beta_1-\beta_1^G)}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G},\end{aligned}$$

and

$$\begin{aligned}f_2(1)&=q_2c_1^G+r_2c_2^G-c_2^G=q_2(c_1^G-c_2^G)-p_2c_2^G\\&=\frac{q_2q_1^Gp_2^G}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G}-\frac{p_2q_1^Gq_2^G}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G}=\frac{q_2p_1^Gp_2^G\,(\beta_2^G-\beta_2)}{p_1^Gp_2^G+q_1^Gp_2^G+q_1^Gq_2^G}.\end{aligned}$$

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 $f_1(\alpha)\le0$ for all $\alpha\in[\alpha_1,1]$ and $f_2(\alpha)\le0$ for all $\alpha\in[\alpha_2,1]$. Let $\alpha_0=\max\{\alpha_1,\alpha_2\}$. Then, if $\beta_i<\beta_i^G$, (A8) holds for all $\alpha\in[\alpha_0,1]$.

□

Proof of Theorem 1: The MDP model we introduced in Section 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 A4(c) that $D_i v_\alpha(x_1,x_2) \le c_i^G$. Then,

$$h_\alpha(x_1,x_2) = v_\alpha(x_1,x_2) - v_\alpha(0,0) = \sum_{i=0}^{x_1-1} D_1 v_\alpha(i,0) + \sum_{j=0}^{x_2-1} D_2 v_\alpha(x_1,j) \le x_1 c_1^G + x_2 c_2^G \le (b+1)(c_1^G + c_2^G).$$

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). Using (A7), the average cost optimality equation (3) can be rewritten as

$$h(x_1,x_2) + g = \min_{(a_1,a_2)\in\mathcal{A}(x_1,x_2)} \Big\{ a_1\phi_1^G + a_2\phi_2^G + q_1(x_1-a_1) + \Gamma h(x_1-a_1, x_2-a_2) \Big\}. \tag{A31}$$

For $x + y \le b$,

$$\begin{aligned}
\Gamma h(x,y) &= \sum_{i=0}^{x+y+1}\sum_{j=0}^{x+y+1-i} P(i,j|x,y) \lim_{\alpha\to 1^-} h_\alpha(i,j) = \lim_{\alpha\to 1^-}\left[\sum_{i=0}^{x+y+1}\sum_{j=0}^{x+y+1-i} P(i,j|x,y) h_\alpha(i,j)\right] \\
&= \lim_{\alpha\to 1^-}\left[\sum_{i=0}^{x+y+1}\sum_{j=0}^{x+y+1-i} P(i,j|x,y) v_\alpha(i,j) - \sum_{i=0}^{x+y+1}\sum_{j=0}^{x+y+1-i} P(i,j|x,y) v_\alpha(0,0)\right] \\
&= \lim_{\alpha\to 1^-}\big[\Gamma v_\alpha(x,y) - v_\alpha(0,0)\big].
\end{aligned}$$

Then, we have for $x + y \le b-1$,

$$\begin{aligned}
D_1\Gamma h(x,y) &= \Gamma h(x+1,y) - \Gamma h(x,y) = \lim_{\alpha\to 1^-}\big[\Gamma v_\alpha(x+1,y) - v_\alpha(0,0)\big] - \lim_{\alpha\to 1^-}\big[\Gamma v_\alpha(x,y) - v_\alpha(0,0)\big] \\
&= \lim_{\alpha\to 1^-}\big[\Gamma v_\alpha(x+1,y) - \Gamma v_\alpha(x,y)\big] = \lim_{\alpha\to 1^-} D_1\Gamma v_\alpha(x,y), \\
D_2\Gamma h(x,y) &= \Gamma h(x,y+1) - \Gamma h(x,y) = \lim_{\alpha\to 1^-}\big[\Gamma v_\alpha(x,y+1) - \Gamma v_\alpha(x,y)\big] = \lim_{\alpha\to 1^-} D_2\Gamma v_\alpha(x,y).
\end{aligned}$$

If (A8) holds, then from (A17) and (A18), for any $x + y \le b-1$,

$$D_1\Gamma v_\alpha(x,y) \le \frac{c_1^G}{\alpha} - q_1, \quad D_2\Gamma v_\alpha(x,y) \le \frac{c_2^G}{\alpha}.$$

Then, since $\beta_i < \beta_i^G$ for both $i = 1,2$, we know from Lemma A9 that there exists α_0 such that for $\alpha \in [\alpha_0, 1]$, (A8) holds and thus for such α we have,

$$D_1\Gamma h(x,y) = \lim_{\alpha\to 1^-} D_1\Gamma v_\alpha(x,y) \le \lim_{\alpha\to 1^-} \frac{c_1^G}{\alpha} - q_1 = \phi_1^G - q_1, \tag{A32}$$

$$D_2\Gamma h(x,y) = \lim_{\alpha\to 1^-} D_2\Gamma v_\alpha(x,y) \le \lim_{\alpha\to 1^-} \frac{c_2^G}{\alpha} = \phi_2^G, \tag{A33}$$

where the last equalities follow from (1), (A5), and

$$\lim_{\alpha\to 1^-} c_1^G = \lim_{\alpha\to 1^-} \frac{\alpha q_1^G(1-\alpha r_2^G)}{(1-\alpha r_1^G)(1-\alpha r_2^G)-\alpha^2 p_1^G q_2^G} = \frac{q_1^G(p_2^G+q_2^G)}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} = \phi_1^G,$$
$$\lim_{\alpha\to 1^-} c_2^G = \lim_{\alpha\to 1^-} \frac{\alpha^2 q_1^G q_2^G}{(1-\alpha r_1^G)(1-\alpha r_2^G)-\alpha^2 p_1^G q_2^G} = \frac{q_1^G q_2^G}{p_1^G p_2^G + q_1^G p_2^G + q_1^G q_2^G} = \phi_2^G.$$

If (a_1,a_2) and (a_1+1,a_2) are both feasible actions in state (x_1,x_2), then (A32) implies that

$$(a_1+1)\phi_1^G + a_2\phi_2^G + q_1(x_1-a_1-1) + \Gamma h(x_1-a_1-1, x_2-a_2)$$
$$\geq a_1\phi_1^G + a_2\phi_2^G + q_1(x_1-a_1) + \Gamma h(x_1-a_1, x_2-a_2),$$

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 (A33) implies

$$a_1\phi_1^G + (a_2+1)\phi_2^G + q_1(x_1-a_1) + \Gamma h(x_1-a_1, x_2-a_2-1)$$
$$\geq a_1\phi_1^G + a_2\phi_2^G + q_1(x_1-a_1) + \Gamma h(x_1-a_1, x_2-a_2),$$

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.

□

A4.2. Proof of Theorem 2

From Theorem 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 (A31) as

(i) if $x_1+x_2 \leq b-1$,

$$h(x_1,x_2)+g = q_1x_1 + \Gamma h(x_1,x_2),$$

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

$$h(x_1,x_2)+g = \phi_1^G + q_1(x_1-1) + \Gamma h(x_1-1,x_2),$$

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

$$h(x_1,x_2)+g = \phi_2^G + q_1x_1 + \Gamma h(x_1,x_2-1),$$

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

$$h(x_1,x_2)+g = \min\Big\{\phi_1^G + q_1(x_1-1) + \Gamma h(x_1-1,x_2),\ \phi_2^G + q_1x_1 + \Gamma h(x_1,x_2-1)\Big\}.$$

Let $\bar{\delta}_n(x_1) = \Delta\Gamma h(x_1, n - x_1)$ for $0 \le x_1 \le n \le b-1$. 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

$$\bar{\delta}_{b-1}(x_1 - 1) = \Delta\Gamma h(x_1 - 1, x_2 - 1) = \Gamma h(x_1, x_2 - 1) - \Gamma h(x_1 - 1, x_2) > \phi_1^G - \phi_2^G - q_1. \quad \text{(A34)}$$

As in the proof of Theorem 1, using Theorem 6.4.2 of Sennott (1999) we can write, for $x, y \ge 0, x + y \le b - 1$,

$$\Delta\Gamma h(x,y) = \Gamma h(x+1, y) - \Gamma h(x, y+1)$$
$$= \lim_{\alpha \to 1^-} \left[\Gamma v_\alpha(x+1, y) - \Gamma v_\alpha(x, y+1)\right] = \lim_{\alpha \to 1^-} \Delta\Gamma v_\alpha(x, y).$$

Since $\beta_i < \beta_i^G$ for $i = 1, 2$, we know from Lemma (A9) that there exists $\alpha_0 < 1$ such that (A8) holds for $\alpha \in [\alpha_0, 1]$. As in the proof or Proposition A2, we can also conclude from Lemma A6 that for $x + y \le b - 1$ and $x > 0$, $\Delta\Gamma v_\alpha(x,y) \ge \Delta\Gamma v_\alpha(x-1, y+1)$ for $\alpha \in [\alpha_0, 1]$. Then, it follows that for $n = x + y$,

$$\bar{\delta}_n(x) = \Delta\Gamma h(x,y) \ge \Delta\Gamma h(x-1, y+1) = \bar{\delta}_n(x-1). \quad \text{(A35)}$$

Thus, for fixed n, $\bar{\delta}_n(x)$ is non-decreasing in x for $0 \le x \le n$.

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 (A35), 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 (A34) 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.

□

A4.3. Proof of Theorem 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 (A5) and (A14), 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 (A5) and (A14) with (1) and (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 $\tilde{f}_k(1) = (\phi_k^G - \phi_k) - \left(\phi_{3-k}^G - \phi_{3-k}\right)$.

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 A9, if $\beta_i < \beta_i^G$ for $i = 1,2$, then there exists $\alpha_0 \in (0,1)$ such that (A8) 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 (6) holds for $k = 1$, then (A8) and (A15) 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 A3 (a). Hence, $a^*(x_1, x_2) = (1, 0)$ according to (A34) 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 (6) holds for $k = 2$, then (A8) and (A16) hold for all $\alpha \in [\bar{\alpha}, 1]$. Then, for all x_1, x_2 such that $x_1 > 0$, $x_2 > 0$ and $x_1 + x_2 = b + 1$,

$$\Delta\Gamma h(x_1 - 1, x_2 - 1) = \lim_{\alpha \to 1^-} \Delta\Gamma v_\alpha(x_1 - 1, x_2 - 1) \le \lim_{\alpha \to 1^-} \frac{c_1^G - c_2^G}{\alpha} - q_1 = \phi_1^G - \phi_2^G - q_1,$$

where the inequality follows from $\Delta\Gamma v_\alpha(x_1 - 1, x_2 - 1) \le \frac{c_1^G - c_2^G}{\alpha}$, which has been established in the proof of Proposition A3 (b). Hence, $a^*(x_1, x_2) = (0, 1)$ according to (A34) for all x_1, x_2 such that $x_1 > 0$, $x_2 > 0$ and $x_1 + x_2 = b + 1$.

□

Appendix A5: Supplements for the simulation study

A5.1. Computation of the death probabilities and expected lengths of stay

Transition probability matrix as described in Figure 2 is

$$P_{ICU} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ q_1 & r_1 & p_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & q_{2L} & r_{2L} & 0 & p_{2L} & 0 & 0 & 0 \\ 0 & q_{2H} & 0 & r_{2H} & p_{2H} & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{3L} & r_{3L} & 0 & p_{3L} & 0 \\ 0 & 0 & 0 & q_{3H} & 0 & r_{3H} & p_{3H} & 0 \\ 0 & 0 & 0 & 0 & 0 & q_4 & r_4 & p_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Using this matrix, we can compute the expected mortality rate (ϕ_i) for $i \in \{1, 2L, 2H, 3L, 3H, 4\}$ by solving the following system of linear equations:

$$[\phi_1, \phi_{2L}, \phi_{2H}, \phi_{3L}, \phi_{3H}, \phi_4]' = [q_1, 0, 0, 0, 0, 0]' + P_{\text{ICU\_submatrix}}[\phi_1, \phi_{2L}, \phi_{2H}, \phi_{3L}, \phi_{3H}, \phi_4]',$$

where

$$P_{\text{ICU\_submatrix}} = \begin{bmatrix} r_1 & p_1 & 0 & 0 & 0 & 0 \\ q_{2L} & r_{2L} & 0 & p_{2L} & 0 & 0 \\ q_{2H} & 0 & r_{2H} & p_{2H} & 0 & 0 \\ 0 & 0 & q_{3L} & r_{3L} & 0 & p_{3L} \\ 0 & 0 & q_{3H} & 0 & r_{3H} & p_{3H} \\ 0 & 0 & 0 & 0 & q_4 & r_4 \end{bmatrix}.$$

The solution can then be found to be

$$[\phi_1, \phi_{2L}, \phi_{2H}, \phi_{3L}, \phi_{3H}, \phi_4]' = (I - P_{\text{ICU\_submatrix}})^{-1} \cdot [q_1, 0, 0, 0, 0, 0]'. \tag{A36}$$

Similarly, we can compute the expected length-of-stay in ICU (L_i) for $i \in \{1, 2L, 2H, 3L, 3H, 4\}$ by solving

$$[L_1, L_{2L}, L_{2H}, L_{3L}, L_{3H}, L_4]' = [1,1,1,1,1,1]' + P_{\text{ICU\_submatrix}} \cdot [L_1, L_{2L}, L_{2H}, L_{3L}, L_{3H}, L_4]',$$

from which we find

$$[L_1, L_{2L}, L_{2H}, L_{3L}, L_{3H}, L_4]' = (I - P_{\text{ICU\_submatrix}})^{-1} \cdot [1,1,1,1,1,1]'. \tag{A37}$$

A5.2. Patient arrival process

Figure A1 shows the number of Influenza A patients reported by public health laboratories in the US (over time) obtained from CDC FluView Interactive webpage (`https://www.cdc.gov/flu/weekly/fluviewinteractive.htm`).

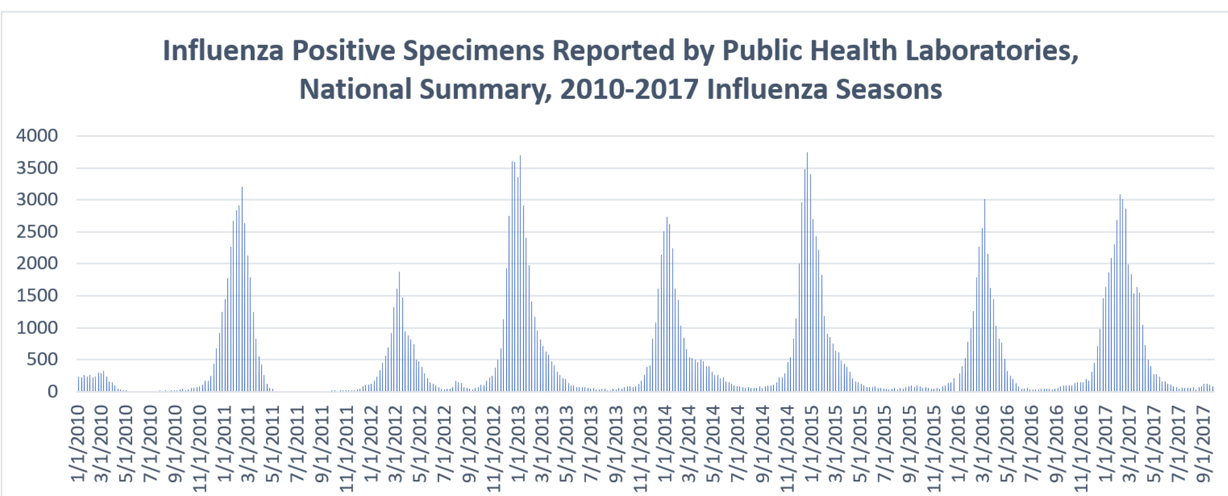

Figure A1 Influenza positive specimens reported by public health laboratories from 2010 to 2017.

A5.3. Computation of aggregated transition probabilities

We assume that the decision maker's estimates for the expected dying probabilities in ICU and GW, and ICU expected length of stay for aggregated stages $A1$ and $A2$ satisfy the following:

$$\theta_{A1} = \sum_{i \in A1} \theta_i, \phi_{A1} = \frac{\sum_{i \in A1} \theta_i \phi_i}{\theta_{A1}}, \phi_{A1}^G = \frac{\sum_{i \in A1} \theta_i \phi_i^G}{\theta_{A1}}, L_{A1} = \frac{\sum_{i \in A1} \theta_i L_i}{\theta_{A1}};$$
$$\theta_{A2} = \sum_{i \in A2} \theta_i, \phi_{A2} = \frac{\sum_{i \in A2} \theta_i \phi_i}{\theta_{A2}}, \phi_{A2}^G = \frac{\sum_{i \in A2} \theta_i \phi_i^G}{\theta_{A2}}, L_{A2} = \frac{\sum_{i \in A2} \theta_i L_i}{\theta_{A2}}.$$

The aggregated-stage transition probabilities can then be estimated from $\phi_{A1}, \phi_{A2}, L_{A1}, L_{A2}$ by solving the following system of equations:

$$\begin{bmatrix} 1, & 1, & 1 \\ 1, & \phi_{A1}, & \phi_{A2} \\ 0, & L_{A1}, & L_{A2} \end{bmatrix} \cdot \begin{bmatrix} q_{A1}, & q_{A2} \\ r_{A1}, & r_{A2} \\ p_{A1}, & p_{A2} \end{bmatrix} = \begin{bmatrix} 1, & 1 \\ \phi_{A1}, & \phi_{A2} \\ L_{A1} - 1, & L_{A2} - 1 \end{bmatrix},$$

and thus,

$$\begin{bmatrix} q_{A1}, & q_{A2} \\ r_{A1}, & r_{A2} \\ p_{A1}, & p_{A2} \end{bmatrix} = \begin{bmatrix} 1, & 1, & 1 \\ 1, & \phi_{A1}, & \phi_{A2} \\ 0, & L_{A1}, & L_{A2} \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1, & 1 \\ \phi_{A1}, & \phi_{A2} \\ L_{A1} - 1, & L_{A2} - 1 \end{bmatrix}.$$

A5.4. Table EC.1

See Table EC.1 on the next page.

A5.5. Simulation results on patient length of stay

Figures A2, A3, and A4 below report the mean length-of-stay under every policy π (over the 100 replications) for each scenario and a 95% confidence interval for the mean. The figures provide these confidence intervals along with the box plots, where we also indicate the 1st and 3rd quantiles, the minimum, and the maximum values.

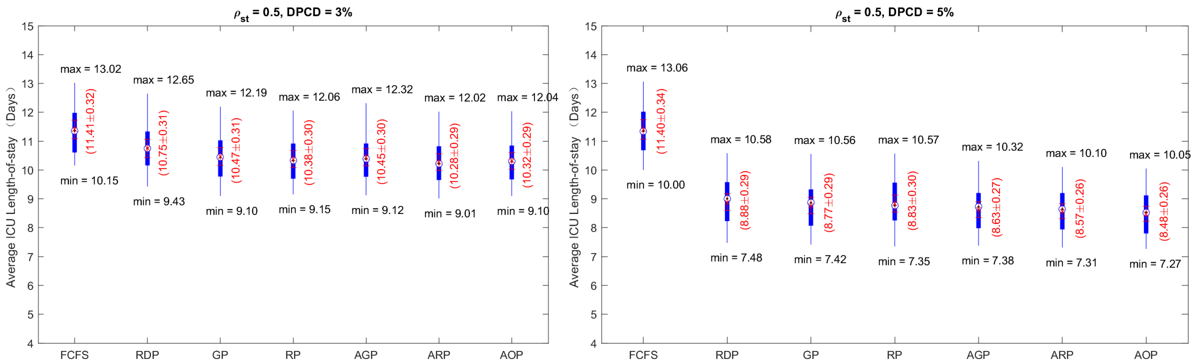

Figure A2 Average length of stay in the ICU for each scenario with $\rho_{\text{st}} = 0.5$, where average is taken over 100 replications.

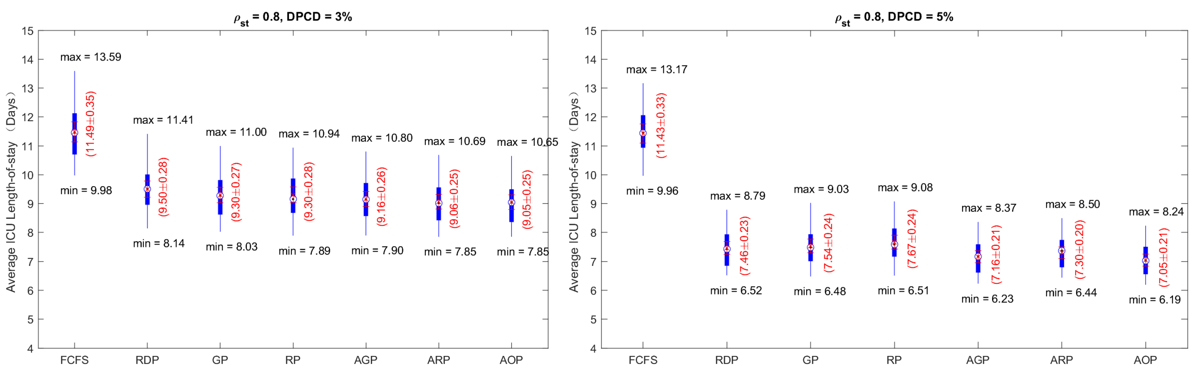

Figure A3 Average length of stay in the ICU for each scenario with $\rho_{\text{st}} = 0.8$, where average is taken over 100 replications.

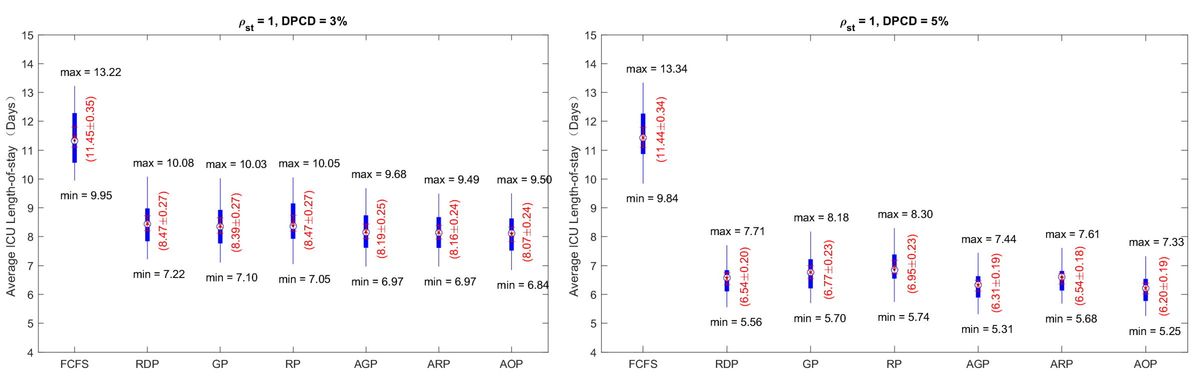

Figure A4 Average length of stay in the ICU for each scenario with $\rho_{\text{st}} = 1$, where average is taken over 100 replications.

Table EC.1 Fraction of time Condition (5) is violated (as explained in Section 6.3) in each scenario of the simulation study

Scenario	DPCD = 0.3%			DPCD = 0.5%		
Transition	$\rho = 0.5$	$\rho = 0.8$	$\rho = 1$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 1$
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	1	1	1	1	1	1
4	0.135	0.262	0.317	0.206	0.282	0.313
5	1	1	1	1	1	1
6	0	0	0	0	0	0
7	0	0.012	0.067	0.06	0.131	0.167
8	0	0	0	0	0	0.004
9	1	1	1	1	1	1
10	1	1	1	1	1	1
11	0	0	0	0	0	0
12	0	0	0	0	0	0.03
13	1	1	1	1	1	1
14	0.115	0.238	0.302	0.198	0.27	0.306
15	1	1	1	1	1	1
16	0	0	0	0	0	0
17	0.012	0.135	0.194	0.131	0.206	0.242
18	0.052	0.175	0.234	0.159	0.23	0.266
19	1	1	1	1	1	1
20	0.02	0.139	0.198	0.135	0.21	0.246
21	0.286	1	1	0.298	1	1
22	0.06	0.187	0.246	0.159	0.238	0.274
23	0	0.028	0.087	0.067	0.143	0.175
24	1	1	1	1	1	1
25	1	1	1	1	1	1
26	1	1	1	1	1	1
27	1	1	1	1	1	1
28	0.179	0.302	1	0.23	0.306	0.996
29	0	0	0	0	0	0
30	1	1	1	1	1	1