Submitted to *Operations Research*
manuscript (Please, provide the manuscript number!)

Dynamic Distribution of Casualties to Medical Facilities in the Aftermath of a Disaster

Alex F. Mills
Department of Operations and Decision Technologies, Kelley School of Business, Indiana University, Bloomington IN 47405, millsaf@indiana.edu

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

The safe and timely treatment of casualties following a disaster is a challenging task that requires coordination among emergency medical services providers and medical facilities. This challenge is especially serious for casualties who require transportation to facilities where they will be treated. We formulate the problem of dynamically distributing such casualties as a queueing control problem where both transportation and treatment resources are limited. We partially characterize the solution that maximizes the expected total discounted reward, show that in general the optimal solution may be complex and non-intuitive, and propose heuristic policies. Using a simulation study based on data from a national survey of emergent trauma patients, we demonstrate that these policies are both efficient and practical. In particular, we observe that there are situations where proposed heuristics can more than double the expected discounted patient throughput compared to sending patients to the nearest facility or the facility with the shortest queue. Moreover, the proposed dynamic heuristics can be effective even when information on the congestion levels at the medical facilities is only available periodically.

Key words: Dynamic programming/optimal control: applications; Health care, ambulance service: distribution.

1. Introduction

Natural disasters and other mass-casualty events such as those caused by terrorist attacks are a reality of life. Although they are rare, when they happen, these events can affect hundreds to thousands of people and place a huge burden on emergency response systems for hours to days. It is not always possible to prevent such events from happening but the damage they cause may be mitigated to some extent by advanced planning. One way an effective emergency response can make a difference is through providing timely treatment to the patients by carefully distributing them

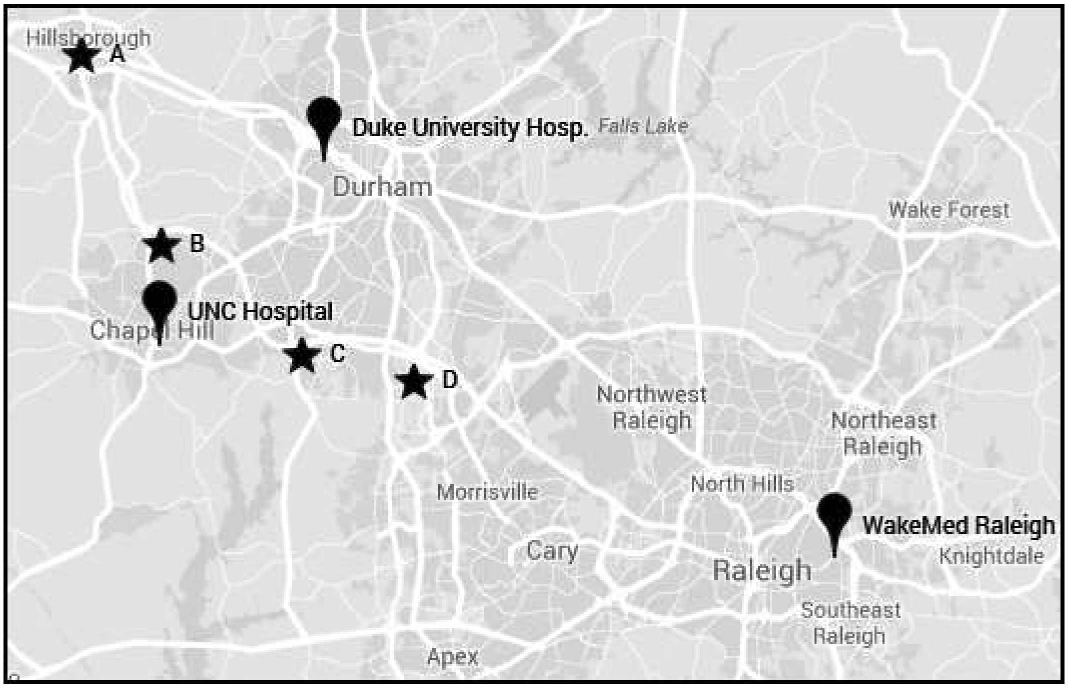
to nearby hospitals. This paper provides some guidelines and insights into how such distribution decisions should be made.

In the aftermath of a disaster, the main goal of emergency response management is to ensure the safety of the affected people and the timely treatment of as many casualties as possible. This requires making several complex decisions under time pressure and sometimes under security and safety concerns: Which hospitals will be involved? How much surge capacity each hospital will need? Which patients should receive priority for transportation by emergency vehicles? Responders also have to re-evaluate these decisions periodically based on these changing conditions, which might be a result of infrastructure damage on the transportation routes, overcrowding at the treatment facilities, etc. One important decision that needs to be updated according to these changing conditions is how to distribute casualties from areas affected by the disaster to medical facilities that are participating in the response effort. In this article, we focus on this dynamic decision making problem, which arises in the aftermath of a large-scale emergency event affecting multiple regions that are geographically separated but close enough to be served by the same group of treatment facilities (e.g., within the same urban area).

To understand the complexity of the problem, consider a hypothetical disaster scenario in the Research Triangle region of North Carolina, depicted in Figure 1. In the figure, the three Level I trauma centers serving the region (namely, UNC Hospital, Duke University Hospital, and WakeMed Raleigh Hospital) are shown at their actual locations, and casualties of the hypothetical disaster are clustered around areas A, B, C, and D, designated by stars on the map. For such a scenario, our decision problem is how to distribute casualties via ambulances from these four incident locations to the three trauma centers based on the changing congestion levels at these medical facilities. There are several factors complicating this decision problem. Because this event involves many casualties, it will not be feasible to rely on a single trauma center that is nearest to each incident location. Note that locations B and C are closest to UNC, while A and D are about the same distance to UNC and Duke. WakeMed is the most distant hospital, and depending on the size of the event, the emergency response management may want to send casualties to WakeMed as well. However, it is not clear how big an event would require some of the casualties to be distributed to WakeMed. Clearly, sending a casualty to WakeMed would not only delay the treatment to that patient due to the long travel time, but also delay other casualties' transportation because transferring a casualty to a distant hospital means that an ambulance will not be available for an extended period of time.

While each type of event is unique in terms of resource requirements and treatments provided, one consensus in the emergency response literature is that the dynamic management of casualty distribution could lead to improved outcomes by incorporating information about the capabilities of treatment facilities and their levels of congestion. In a study of thousands of patients evacuated

Figure 1 Map of a hypothetical disaster scenario in the Research Triangle region of North Carolina, where there are three Level I trauma centers (map data ©2013 Google). Casualties are located in the areas designated by stars.



from an earthquake, Tanaka et al. (1998) suggested that to improve the outcome of the response effort, “disaster officials must know the capabilities and capacity of each area hospital at all times to select appropriate triage and mode of transport for each victim.” Robenshtok et al. (2003) state that in a chemical event, “specific hospitals should be designated to receive only chemical agent casualties.” Hick et al. (2011) suggest that during a nuclear incident, emergency managers should bypass hospitals that are “completely overwhelmed.” Hrdina et al. (2009) note that a disaster involving radiation “will require a wider distribution of patients” than a typical mass-casualty event and thus “a networked system” for assigning patients to medical facilities is an essential part of any model for responding to such an incident. Making decisions dynamically is particularly important in a disaster involving the release of hazardous substances (HAZMAT). Hazardous substances include biological, chemical, or nuclear agents, and they may be released deliberately (as in a terrorist attack) or accidentally (as in a meltdown at a nuclear power plant). HAZMAT events pose an additional challenge because both casualties and responders are at risk from exposure (U.S. Department of Homeland Security 2005). Hence, it is especially important in such events to evacuate casualties away from the incident site and decontaminate them as soon as possible. Although the preferable practice is to establish decontamination as close to the scene of the incident as possible, health care facilities often do not rely on public safety agencies to provide

decontamination services but instead elect to provide those services prior to admission of patients (Hick et al. 2003).

In the aftermath of a disaster, most medical facilities receive two types of patients: *ambulatory patients* who present by their own means and *non-ambulatory patients* with severe injuries that require transportation by emergency vehicles such as ambulances and helicopters. The emergency responders in general have little or no control over the transportation of ambulatory patients. On the other hand, an emergency plan can provide guidelines as to how the non-ambulatory patients should be sent to different hospitals within the area to increase the effectiveness of the response effort. Furthermore, in response to such events, hospitals establish mobile clinics, usually outside the hospital, to provide basic service to ambulatory patients, and thereby create additional capacity for patients who are in more critical conditions (usually those who are non-ambulatory) inside the hospital (Hick et al. 2004). In this article, our goal is to develop dynamic decision rules that prescribe which non-ambulatory casualties should be transported to which treatment facilities by taking into account changing congestion levels at the facilities in addition to the travel times and hospital capabilities. Here, casualties are characterized by their geographical locations, which determine their travel times to receiving facilities.

Our main approach to this problem is to build a queueing model and formulate the underlying control problem as a Markov decision process (see Section 3). Our formulation captures the essential tradeoff in the casualty distribution problem: sending a casualty to a less congested facility farther from the disaster, which may be good for that casualty, will tie up a transportation resource for a longer period of time, which will be detrimental to later casualties. This formulation enables us to obtain analytical results (presented in Section 4) about the types of policies that are expected to perform well. For example, we identify conditions under which casualties should be routed to the least congested facility. The formulation also allows us to identify situations where the optimal policy is not necessarily intuitive. Building on these results, we then employ two methods to develop heuristic policies: the myopic approach and the one-step policy improvement heuristic, details of which can be found in Section 5. The final step in our analysis is a set of numerical experiments and a realistic simulation study, which uses data from the National Hospital Ambulatory Medical Care Survey. These studies, as reported in Section 6, show that our heuristic policies perform well compared with a baseline static and a dynamic policy. Our analysis also provides useful insights into the casualty distribution problem, which we discuss and summarize in Section 7.

Before we present our mathematical and numerical analysis, we review the relevant work within the operations research literature on emergency response and dynamic routing problems in the next section. Proofs of all analytical results presented in this article are provided in an Electronic Companion.

2. Literature Review

Emergency vehicle allocation and routing problems in the aftermath of a mass-casualty event have been studied in three recent articles by Gong and Batta (2007), Jotshi et al. (2009), and Zayas-Cabán et al. (2013). Gong and Batta (2007) consider the problem of determining an initial allocation of ambulances to groups of casualties in a disaster and use deterministic optimization methods to suggest an allocation that minimizes the total time needed to evacuate all casualties. Jotshi et al. (2009) consider the problem of both allocating and routing emergency vehicles in a disaster. Their model is designed to use large amounts of real-time data on medical facilities, casualties, vehicle locations, and roadways in order to apply a heuristic policy for casualty pickup, vehicle routing, and casualty dropoff problems. For each of these problems, the authors recommend a simple heuristic using parameters such as travel times and number of casualties, which are assigned weights by the user. One drawback to this type of heuristic approach is that a considerable amount of "tuning" of the user-specified weights may be needed. In another relevant work, Zayas-Cabán et al. (2013) develop policies to mobilize ambulances from surrounding areas to respond to a mass-casualty incident and to allocate this additional capacity among affected regions. The objective is to clear the system as quickly as possible while keeping the cost of moving vehicles under a certain budget. Our work differs from these three articles in terms of either the main problem considered or the approach taken. In particular, unlike in Gong and Batta (2007) and Zayas-Cabán et al. (2013), we assume that the initial decision of ambulance allocation to disaster sites has been made and the problem is determining how patients will be distributed among the receiving hospitals based on the dynamically changing congestion levels. We believe that the problems of static ambulance allocation to disaster sites and dynamic ambulance distribution to hospitals afterwards are both important in determining the success of the response effort in the aftermath of a disaster. However, taking a mathematical approach to solve these two optimization problems simultaneously would be difficult. One could at best numerically study some simple decision rules for the combined problem, as in Jotshi et al. (2009). In this paper, we focus solely on the dynamic ambulance distribution problem with the goal of developing policies that are based in part on some mathematical analysis.

Dynamic routing of customers or jobs to servers is a well-studied problem in the operations research literature. In many service systems, customers must decide which queue to join, and in other cases, such as a call center, they are routed to an appropriate server upon their arrival. In the case where servers are homogeneous, routing customers to the shortest queue is optimal in many situations; see Whitt (1986) for several examples and counterexamples. When servers are nonidentical but have exponentially distributed service times, a generalization of the shortest-queue

policy is known to be optimal (Hordijk and Koole 1992). A number of articles have developed index policies for routing customers in more complex systems, such as those with heterogeneous servers, impatient customers, or general delay costs—see, e.g., van Mieghem (1995), Mandelbaum and Stolyar (2004), Armony (2005), Glazebrook et al. (2009), Argon et al. (2009). These articles assume that customers are routed to a server instantaneously, either upon their arrival (after which they may wait in a queue at their assigned server), or just before service (after waiting in a single queue). In our problem, this would correspond to ambulances traveling instantaneously to medical facilities, which would strip away an essential component of the problem. Hence, in our formulation, distributing casualties to medical facilities requires acquiring a limited resource (e.g., an ambulance) for a period of time. This feature significantly complicates the decision making problem because the assignment of the current patient to a facility affects the later patients not just by lengthening the queue at one of the facilities, but also by preventing the other patients from using the transportation resource while the current patient is being transported.

There is also a growing literature on the evacuation of treatment facilities in the aftermath of a disaster. Tayfur and Taaffe (2009) use optimization to identify staffing and transportation schedules while evacuating patients from a hospital to a set of receiving hospitals under the objective of cost minimization. Similarly, Bish et al. (2011) employ a mathematical programming approach to study an hospital evacuation problem but with the objective of minimizing risk to the affected patients. Facility evacuations (and specifically hospital evacuations) are amenable to solution by deterministic optimization methods mainly because there is no issue of queueing at the receiving hospitals (at least in the formulations that are considered in the literature). In contrast, queueing of emergent patients at the hospitals is an essential part of the casualty distribution problem, which we consider in this study. For a more detailed review of literature on hospital evacuations and directions for future research, see Childers and Taaffe (2010).

Finally, there is a vast literature on the daily operations of emergency medical services, from which we briefly review several recent articles: some studied the static problem of allocating ambulances (Ingolfsson et al. 2008), others considered dynamic redeployment of ambulances (Maxwell et al. 2010), and others developed dispatching rules to emergency calls (McLay and Mayorga 2013). Unlike in our problem, these problems are all related to daily emergency operations and none of them need to take into account the congestion levels at the hospitals. One exception is the case where hospitals divert ambulances due to high patient loads, which has been studied by Deo and Gurvich (2011) and Allon et al. (2013). Neither Deo and Gurvich (2011) nor Allon et al. (2013) considered the scarcity of ambulances (in particular, its effect on later patients), which is much less of an issue for daily emergencies than for a disaster setting.

3. Dynamic Programming Formulation

In this section, we develop a dynamic programming model to study the main tradeoffs inherent in the problem of distributing non-ambulatory casualties to multiple medical facilities in a resource-constrained environment. In particular, when considering which hospital to send an individual casualty, the alternatives may include a facility that is closest to the incident, one that provides the highest quality care, and one that is the least congested relative to its capacity. For disasters involving many casualties, making this choice is particularly difficult because because sending a casualty to a facility that is best for him or her may result in delaying another casualty's transfer to a hospital. In order to study these tradeoffs, in the dynamic programming formulation we purposely abstract away some of the other elements of the problem (such assumptions are discussed later in this section). By making these abstractions, we are able to generate insights into the fundamental tradeoff inherent to this problem and also develop solution methods that can lead to implementable policies (see Sections 4 and 5). However, we later conduct simulations to test the proposed methods under much more general conditions where the restrictive assumptions are relaxed (see Section 6).

In our formulation, we consider a disaster with multiple casualty locations that are separated geographically and have infinitely many casualties. Assume that each location has a set of transportation resources that can be used to transport the casualties to one of several receiving facilities such as major hospitals (that we also refer to as "stations" throughout the article). The rate at which casualties can be transported from a given location to a given facility is inversely proportional to the travel time between the location and the facility. Once a casualty reaches the facility, he or she waits to receive service (i.e., treatment) from a single server that operates according to the first-come-first-served discipline. After the completion of the service of a casualty, the system earns a "reward." The performance measure of interest is the expected total discounted reward earned by the system. Here, the reward can be a measure of some medical outcome, and it can depend on the facility where the casualty receives service. Thus, the rewards can be used to represent differences in quality or capability among the facilities. By discounting the rewards, we capture the time sensitivity of medical treatments and the benefits gained from timely treatment. In the absence of information about the capabilities of the different facilities, it may be convenient to think of all rewards as being one, in which case the performance measure of interest is simply the discounted throughput of the system.

We let $\mathcal{R}$ denote the finite set of all transportation resources. Each transportation resource is assigned to a specific location although there may be more than one resource per location. However, without loss of generality, we can assume that each location has one transportation resource: because locations have infinitely many casualties, having more than one resource at a location is the

same as having more than one location with identical traveling time distributions. (Hence, we let $\mathcal{R}$ also denote the set of incident locations.) We define $\mathcal{S}$ as the finite set of stations, each of which has a single server with a dedicated queue. A resource $i \in \mathcal{R}$ can transport casualties to station $j \in \mathcal{S}$ with a travel time that is exponentially distributed with mean $1/\tau_{ij}$, where $0 < \tau_{ij} < \infty$. This travel is assumed to be preemptive, i.e., the transportation resources can be re-routed while en route, and travel back to the location is assumed to be instantaneous. The service times at station j are independent and identically distributed exponential random variables with rate μ_j, where $0 < \mu_j < \infty$. When a casualty finishes service at station j, the system receives a positive and bounded reward r_j. We seek a dynamic policy that maximizes the expected total discounted reward earned by the system.

We denote the state of the system by $\mathbf{X}(t) = (X_1(t), \ldots, X_{|\mathcal{S}|}(t))$, where $X_j(t)$ is the number of casualties at station $j \in \mathcal{S}$ at time $t \geq 0$ and $|\mathcal{S}|$ is the cardinality of $\mathcal{S}$. The state space, which we denote by $\mathcal{Q}$, is equal to $\mathbb{N}^{|\mathcal{S}|}$, where $\mathbb{N}$ is the set of non-negative integers. The expected total discounted reward with discount rate $\alpha > 0$ is given by

$$\mathsf{E}\left[\int_0^\infty e^{-\alpha t} r(\mathbf{X}(t))dt\right],$$

where $r(\mathbf{x}) \equiv \sum_{j\in\mathcal{S}} r_j \mu_j I_j(\mathbf{x})$, and $I_j(\mathbf{x})$ is the indicator function that takes value one if $x_j > 0$ and zero if $x_j = 0$, for $\mathbf{x} \in \mathcal{Q}$.

Before we provide a Markov decision process (MDP) formulation, we would like to discuss some of the assumptions that we had to make in order to characterize optimal policies analytically. First, we assume that each incident location has an ample supply of patients. This is clearly an approximation even though the number of casualties in a disaster is generally large. Second, we assume that each facility operates as a single-server queue. This assumption may be valid in certain situations. For example, in the case of a disaster involving hazardous material, decontamination capacity at the facilities may be limited by the number of trained personnel available. In particular, Hick et al. (2003) note that at small facilities, there may be as few as one two-person team responsible from decontaminating non-ambulatory patients. Furthermore, for a general large-scale event that would place a heavy load on nearby hospitals, a single-server approximation may be reasonable even when there are multiple servers. Another assumption that we had to make is that vehicles have Markovian travel times and become instantaneously available after dropping off a casualty. Clearly, some of these assumptions, which are needed for analytical tractability, are hard to justify in reality. Therefore, in Section 6, we test our methods by means of a simulation study where we relax the assumptions of an ample patient supply, a single server, preemptive transportation decisions, Markovian travel times, and instantaneous return travel. This simulation study provides

evidence that the heuristic policies developed under these assumptions perform well even when these assumptions do not hold.

We next formulate the optimization problem at hand as a Markov decision process. Throughout the paper, we use uniformization (Lippmann 1975) with the finite uniformization constant $\beta \equiv \sum_{j\in\mathcal{S}} \mu_j + \tau$, where $\tau \equiv \sum_{i\in\mathcal{R}} \tau_i$ and $\tau_i \equiv \max_{j\in\mathcal{S}} \tau_{ij}$ for $i \in \mathcal{R}$. In other words, by observing the system only at potential transitions, which occur according to a Poisson process with rate β, we study the discrete-time MDP that is embedded in the continuous process. Without loss of generality, we let $\beta = 1$. We assume that each state transition epoch for the uniformized process is also a decision epoch, where the action to be taken is the assignment of a destination station to each incident location. Because each location can be assigned to any station, there are $|\mathcal{S}|^{|\mathcal{R}|}$ possible actions at each epoch. For this problem, it is sufficient to consider only stationary deterministic policies, i.e., policies that assign a single station to each location at any state, independent of time (see, e.g., Theorem 6.2.10 in Puterman 1994). We denote the set of such policies by $\mathcal{P}$.

Let $V(\mathbf{x})$ denote the maximum expected total discounted reward that can be obtained by a policy in $\mathcal{P}$ when the system starts in state $\mathbf{x} \in \mathcal{Q}$. We are interested in finding a policy in $\mathcal{P}$ that yields $V(\mathbf{x})$. In other words, for any initial state $\mathbf{x} \in \mathcal{Q}$, we want to solve

$$\max_{\pi\in\mathcal{P}} \mathsf{E}\left[\int_0^\infty e^{-\alpha t} r(\mathbf{X}^\pi(t))dt \,\middle|\, \mathbf{X}^\pi(0) = \mathbf{x}\right], \tag{1}$$

where $\mathbf{X}^\pi(t)$ is the system state at time $t \geq 0$ under policy $\pi \in \mathcal{P}$. $V(\mathbf{x})$, which is equal to (1) by definition, must satisfy the following optimality equation:

$$V(\mathbf{x}) = \frac{1}{1+\alpha}\left(r(\mathbf{x}) + \sum_{j\in\mathcal{S}} \mu_j V(\mathbf{x} - I_j(\mathbf{x})\mathbf{e}_j) + \tau V(\mathbf{x}) + \sum_{i\in\mathcal{R}} \max_{j\in\mathcal{S}}\{\tau_{ij} M_j(\mathbf{x})\}\right), \tag{2}$$

where $\mathbf{e}_j$ is a vector of size $|\mathcal{S}|$ having its jth component equal to one and all others equal to zero, and $M_j(\mathbf{x}) \equiv V(\mathbf{x}+\mathbf{e}_j) - V(\mathbf{x})$ is the marginal value of having an additional casualty at station j when the system is in state $\mathbf{x} \in \mathcal{Q}$. The value function has three components: (i) the instantaneous reward $r(\mathbf{x})$, which is the expected reward earned from service completion in the next transition, (ii) the future expected reward $\sum_{j\in\mathcal{S}} \mu_j V(\mathbf{x} - I_j(\mathbf{x})\mathbf{e}_j)$ that will be received if the next transition is a service completion, and (iii) the future expected reward $\tau V(\mathbf{x}) + \sum_{i\in\mathcal{R}} \max_{j\in\mathcal{S}}\{\tau_{ij} M_j(\mathbf{x})\}$ that will be received if the next transition is a transportation completion. In Section 4, we provide some partial characterizations of the policies that satisfy (2).

Although it may not be immediately obvious, the optimization problem defined above is also related to a *make-to-stock inventory problem* with lost sales. In the general make-to-stock inventory problem, a firm produces several different products using a single production resource and finished goods are placed in inventory immediately (see, e.g., Veatch and Wein 1996). If demand arrives

for a stocked-out product, it is either backordered or lost and a cost is incurred. At each decision epoch, the firm must decide which product to produce using the production resource, with the objective of minimizing the long-run-average or total discounted cost. Our model can be viewed as a make-to-stock inventory problem with lost sales and multiple production resources, and without holding costs. Specifically, each transportation resource in our problem corresponds to a production resource, and the queues at the different facilities correspond to inventories of the different products. In Veatch and Wein (1996) and subsequent work on make-to-stock inventory (e.g., Ha 1997), index policies were developed primarily pertaining to the case where the objective is to minimize the long-run average costs. Our model differs from those in the inventory control literature with respect to the the cost structure, as a reward is received upon service completion in our model, whereas in the make-to-stock model, a cost is incurred for holding and when demand is lost or backordered. More importantly, our model has multiple resources with different service rates and capabilities, which is a scenario that, to the best of our knowledge, has never been considered in the literature on make-to-stock inventory, and complicates the analysis substantially.

4. Analytical Results

In this section, we establish several analytical results that provide some insight into the structure of the optimal casualty distribution policy, i.e., the solution to the dynamic programming problem given in (2). To aid in the exposition of the analytical results, we start by defining the concept of one station being "preferable" to another.

Definition 1. For an incident location $k \in \mathcal{R}$ and state $\mathbf{x} \in \mathcal{Q}$, we say that station j is $(\mathbf{x}, k)$-preferable to station l if and only if $\tau_{kj} M_j(\mathbf{x}) \geq \tau_{kl} M_l(\mathbf{x})$, for $j, l \in \mathcal{S}$.

When station j is $(\mathbf{x}, k)$-preferable to station l, it means that sending a casualty from incident k to station j will yield an expected total discounted reward at least as large as sending a casualty from incident k to station l in state $\mathbf{x}$ if an optimal policy is used for all future decisions (see Equation (2)). Thus, whenever there exists a station j that is $(\mathbf{x}, k)$-preferable to station l, we can safely eliminate the possibility of sending a casualty from location k to station l in state $\mathbf{x}$. Similarly, Equation (2) implies that it is optimal to send a casualty from location $k \in \mathcal{R}$ to station $j \in \mathcal{S}$ in state $\mathbf{x} \in \mathcal{Q}$ if and only if station j is $(\mathbf{x}, k)$-preferable to every other station.

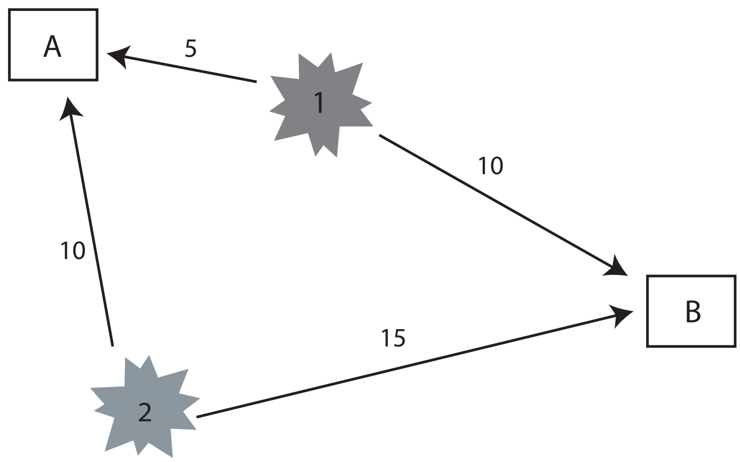
Our first proposition establishes a monotonicity relationship among any pair of two incident locations for any two given stations.

Proposition 1. *Consider any pair of incidents $i, k \in \mathcal{R}$ and any pair of stations $j, l \in \mathcal{S}$, and without loss of generality, label them such that*

$$\frac{\tau_{ij}}{\tau_{il}} \geq \frac{\tau_{kj}}{\tau_{kl}}. \tag{3}$$

For any state $\mathbf{x} \in \mathcal{Q}$, if station j is $(\mathbf{x}, k)$-preferable to station l, then station j is also $(\mathbf{x}, i)$-preferable to station l.

Figure 2 An example with two incidents (locations 1 and 2) and two facilities (A and B). Numbers on the arcs indicate the mean travel times (in minutes) between the incidents and treatment facilities.



In order to demonstrate the intuition behind Proposition 1, consider the example shown in Figure 2, which illustrates a hypothetical simultaneous bombing incident at two locations in an urban city center with two trauma centers. In this example, casualties from location 1 can be transported to facility A twice as fast as to facility B whereas casualties from location 2 can be transported to facility A only one-and-a-half times as fast as to facility B. By Proposition 1, in a given state, if it is optimal for a casualty from location 2 to go to facility A, then it is also optimal for a casualty from location 1 to go to facility A in the same state. An intuitive explanation for this result is that responders at location 1 would prefer to go to facility A versus facility B more strongly than those at location 2, where the difference in travel times between the two facilities is not as large.

Proposition 1 leads to Corollary 1, which helps us eliminate certain *suboptimal* policies from consideration. Note that in this article we use a non-strict definition of suboptimality, i.e., policy π is suboptimal if there exists another policy that performs at least as well as policy π.

Corollary 1. *If* (3) *holds for incident locations $i, k \in \mathcal{R}$ and stations $j, l \in \mathcal{S}$, then a policy that sends a casualty from location k to station j while sending a casualty from location i to station l in any state $\mathbf{x} \in \mathcal{Q}$ is suboptimal. Thus, any policy that assigns location k to station j and location i to station l in the same state can be eliminated from consideration.*

To demonstrate how Corollary 1 can be applied for suboptimal action elimination, consider the example given in Figure 2. For this example, Corollary 1 implies that one of the four possible actions, namely, the action that simultaneously sends a casualty from location 1 to station B and

a casualty from location 2 to station A, is suboptimal regardless of the system state. Hence, by applying the corollary the action space reduces from a set of four actions to a set of three.

In the next proposition, we demonstrate two additional types of monotonicity present in this model. Specifically, Proposition 2 shows that it is always beneficial to have more casualties waiting at each station, and it is always beneficial to have more casualties having completed service than those waiting for service at a station. Besides validating our model, these results are also used in obtaining Proposition 3, which partially characterizes the optimal policy.

Proposition 2. *The following inequalities hold for all states* $\mathbf{x} \in \mathcal{Q}$ *and for all stations* $j \in \mathcal{S}$*.*

$$V(\mathbf{x}+\mathbf{e}_j) \geq V(\mathbf{x}) \tag{4}$$

$$V(\mathbf{x}) + r_j \geq V(\mathbf{x}+\mathbf{e}_j). \tag{5}$$

In Proposition 3, we consider the decision of choosing between two stations for which mean travel times to the incident locations are the same. Such a scenario might happen, for example, when the incidents take place in a rural area and casualties need to be transferred to medical facilities in a distant city. In that case, the medical facilities would be so close to each other with respect to the distance to the casualties that any difference in travel times to different facilities would be negligible.

Proposition 3. *Suppose that for two stations* $i, j \in \mathcal{S}$*, we have* $r_i \geq r_j$*,* $\mu_i \geq \mu_j$*, and* $\tau_{ki} = \tau_{kj}$ *for all* $k \in \mathcal{R}$*. If* $x_i \leq x_j$*, then*

$$V(\mathbf{x}+\mathbf{e}_i) \geq V(\mathbf{x}+\mathbf{e}_j), \tag{6}$$

which implies that station i *is* $(\mathbf{x}, k)$*-preferable to station* j *for all* $k \in \mathcal{R}$ *in every state* $\mathbf{x} \in \mathcal{Q}$ *such that* $x_i \leq x_j$*.*

Proposition 3 states that if there exist two facilities that are equidistant to all locations, then if one has the fastest service, largest reward, and shortest queue, it should be chosen over the other. Proposition 3 also immediately leads to Corollary 2, which gives an "agreeability" condition for optimality.

Corollary 2. *Suppose that* $\tau_{kj} = \tau_k$ *for all* $k \in \mathcal{R}$*,* $j \in \mathcal{S}$*. If one station has the fastest server and the largest reward, then it is optimal to send all casualties to that station whenever that station is the least congested.*

Hordijk and Koole (1992) proved a result similar to Proposition 3 (and thereby to Corollary 2) in the context of a general queue-routing problem. In particular, their Theorem 3.2 finds that under the same arrival rate from sources to stations and under certain additional conditions on the cost

function, a server that has the fastest service and shortest queue should be chosen. However, their result does not imply Proposition 3. First, our reward structure does not meet their conditions on the cost function. Second, Proposition 3 does not require all τ_{ij} to be equal to τ_{il} for all station pairs $j,l \in \mathcal{S}$, but only requires $\tau_{ij'} = \tau_{il'}$ for the two specific stations $j',l' \in \mathcal{S}$ under comparison, unlike the result of Hordijk and Koole (1992).

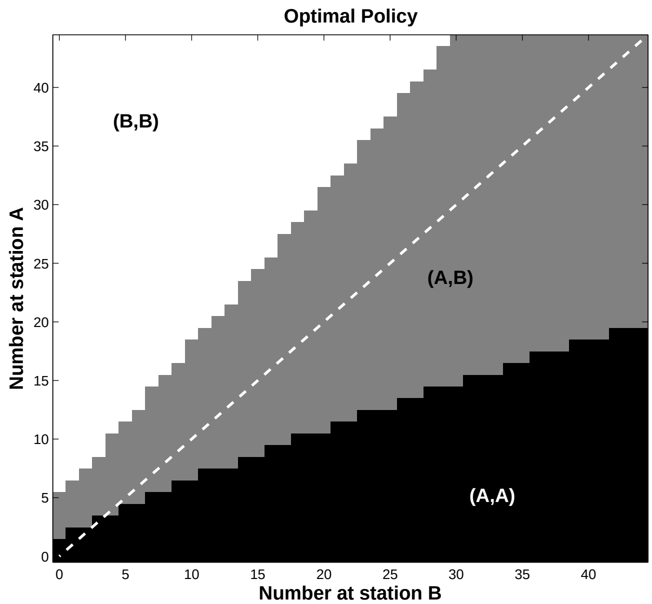
Corollary 2 states that if for each location, the travel times do not change depending on the choice of the service facility and if one facility has the fastest service and largest reward, then that facility can be declared the "top choice" facility in the sense that anytime that facility has the shortest queue, it should be the choice for all locations. This result is intuitive from the perspective of an individual casualty, who would prefer to go to a facility which is the fastest, best in capability, and least congested. Following a similar intuition, one might expect that a facility would also be declared the "top choice" if it is the *closest* station to every incident location in addition to being the fastest and having the largest reward. Interestingly, however, this is not always the case as we demonstrate with the following example.

Example 1. Consider again the example given in Figure 2, where $\tau_{1A} = 12$, $\tau_{1B} = \tau_{2A} = 6$, and $\tau_{2B} = 4$, all in casualties per hour. Suppose that $\mu_A = 15$ per hour, $\mu_B = 12$ per hour, $r_A = r_B = 1$, and $\alpha = 0.7$. We then obtain the structure for the optimal policy, which is presented in Figure 3, using the policy iteration algorithm and truncating the state space so that each station has a capacity of 45 casualties. We know from Corollary 1 that sending a casualty from location 1 to station B and a casualty from location 2 to station A is suboptimal. Hence, in Figure 3 we observe only three regions. An interesting observation from this plot is that even though station A is superior in every dimension to station B, it is not necessarily optimal to send casualties from all locations to station A, even when it has the shortest queue (see, e.g., the gray region below the dashed line). We explain this seemingly counterintuitive result as follows. An optimal policy aims to prevent facilities being idle to the extent possible. Therefore, because one can fill the queue at station A relatively "quickly," when station A has sufficiently many casualties in its queue to keep it busy for a while, one can "afford" to devote some of the transportation resources to taking casualties to station B and thereby reduce the risk that station B becomes idle.

When mean travel times change with both the station and location, we can prove the existence of a "top choice" station if there is a single incident location. However, in this case, our definition of a "top choice" station is different: a top choice station here is the preferred station when it is completely *empty*.

Proposition 4. *Suppose that we have a single incident location labeled as location 1, and that there exists a station (that we label, without loss of generality, as station 1) so that* $\tau_{11} r_1 \mu_1 \geq \tau_{1j} r_j \mu_j$

Figure 3 The structure of the optimal policy for the example shown in Figure 1, under the assumptions that $\mu_A = 15$ per hour, $\mu_B = 12$ per hour, $r_A = r_B = 1$, and $\alpha = 0.7$. For each of the three regions, the first [second] element of the ordered pair is the optimal station for location 1 [2]. Dashed line indicates states with equal queue length.



and $\tau_{11} r_1 \geq \tau_{1j} r_j$ for all $j \in \mathcal{S}$. Then it is optimal to send a casualty to station 1 at every decision epoch for which $x_1 = 0$.

Corollary 3. *Suppose that we have a single incident location and that one station has the largest reward, fastest server, and fastest transportation rate. Then, it is optimal to send a casualty to that station whenever it is empty.*

In this section, we have provided some partial characteristics for the optimal policy. In particular, we have seen that under certain conditions, the optimal policy could simply send a casualty to the least congested station. However, as Example 1 demonstrates, the optimal policy may have a structure that is not immediately obvious, even in the relatively small case of two locations and two stations, where one station appears to be superior to the other in every aspect. Furthermore, due to the curse of dimensionality present in MDPs, it would be extremely time consuming to solve (2) when $\mathcal{R}$ is large (e.g., when a large-scale disaster affects many locations) and/or $\mathcal{S}$ is large. Motivated by the difficult nature of finding the optimal policy, we devote the remainder of this article to developing heuristic methods for the casualty distribution problem and studying their performance through numerical and simulation studies.

5. Heuristic Policies

In this section, we use two methods to develop heuristics that can perform close to the optimal solution. The *myopic approach* chooses an action that maximizes the expected reward for the casualty to be dispatched at the current decision epoch. On the other hand, the *one-step policy improvement approach* requires first setting a static policy and then applying a single step of the policy improvement algorithm. We discuss the details of these two methods in Sections 5.1 and 5.2, and we propose their extensions to the case where the state information is available only at certain points in time in Section 5.3. We report the performance of our heuristic policies from a numerical experiment and a realistic simulation study in Section 6.

5.1. Myopic policy

To obtain a myopic policy, we consider each location separately and at each decision epoch, we determine the station that would yield the largest expected discounted reward for the current casualty to be transported from each location. Specifically, recall that in the uniformized formulation given in Section 3, transitions occur at rate $\beta = 1$. Therefore, if the current casualty at location i is sent to station j, the casualty reaches station j with probability τ_{ij}, in which case he or she joins the queue and becomes the (x_j+1)st casualty. Then, the service completion time of this casualty will be an Erlang random variable with x_j+1 phases and mean $(x_j+1)\mu_j^{-1}$. Hence, the expected discounted reward for the current casualty is $(1+\alpha)^{-1}r_j\left(\frac{\mu_j}{\mu_j+\alpha}\right)^{x_j+1}$, if that casualty reaches station j. On the other hand, with probability $1-\tau_{ij}$, the next transition is not the arrival of the casualty from location i to station j, in which case the expected reward is zero. Therefore, the expected discounted reward earned from choosing to send the current casualty at location i to station j is $(1+\alpha)^{-1}\tau_{ij}r_j\left(\frac{\mu_j}{\mu_j+\alpha}\right)^{x_j+1}$. Because $(1+\alpha)^{-1}$ is positive and does not depend on i or j, we use the index

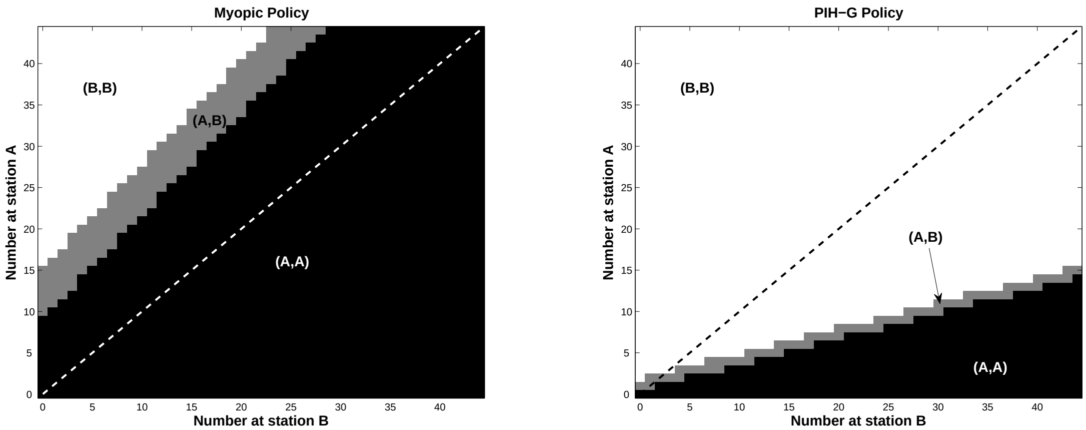
$$\tau_{ij}r_j\left(\frac{\mu_j}{\mu_j+\alpha}\right)^{x_j+1} \tag{7}$$

to propose a myopic policy. Specifically, at each decision epoch, our myopic policy sends a casualty from location $i \in \mathcal{R}$ to station $j \in \mathcal{S}$ that has the largest index (7). Note that this myopic policy results in a set of linear switching curves for each location. In other words, for location $i \in \mathcal{R}$, station $j \in \mathcal{S}$ should be chosen at state $x \in \mathcal{Q}$ if and only if

$$x_j \le a_{kj}x_k + c_{ijk} \quad \forall k \in \mathcal{S}, \tag{8}$$

where a_{kj} and c_{ijk} are constants that depend on system parameters (see Table EC.1 in the Electronic Companion).

Figure 4 (a) Myopic policy and (b) PIH-G policy for the example whose optimal policy is shown in Figure 3 .



(a) (b)

Figure 4(a) shows the structure of the myopic policy for the example whose optimal policy is given in Figure 3. The myopic policy has the advantage of being simple to calculate. Moreover, it can be shown that it shares the same structural properties satisfied by the optimal policy as given in Propositions 1, 3, and 4.

5.2. Policy improvement heuristics

In this section, we develop heuristics based on applying a single step of the policy improvement algorithm starting from a static (state-independent) policy. It is usually much easier to find an optimal static policy (or at least a static policy that performs well) than to find an optimal dynamic policy. Furthermore, the policy that results from applying a single step of the policy improvement algorithm is guaranteed either to be optimal or to perform strictly better than the initial policy (see Proposition 6.4.1 in Puterman 1994). Hence, one can expect to obtain a dynamic policy that performs well by initializing the policy iteration algorithm with a good static policy. This approach would be also easy to apply when the resulting dynamic policy is an index policy with a closed-form expression, which does not necessitate application of the policy iteration step for every instance of the problem. Several studies from the queueing control literature successfully implemented this approach to develop near-optimal policies (see, e.g., Krishnan 1990, Ansell et al. 2003, and Argon et al. 2009).

In this study, we use *Bernoulli splitting* to obtain an initial static policy under which casualties from location $i \in \mathcal{R}$ are sent to station $j \in \mathcal{S}$ independently with probability $\rho_{ij} \geq 0$ at each decision epoch, where $\sum_{j\in\mathcal{S}} \rho_{ij} = 1$. Under such a policy, casualties from location i arrive to station j according to a Poisson process with rate $\rho_{ij}\tau_{ij}$. Then, by defining $\lambda_j \equiv \sum_{i\in\mathcal{R}} \rho_{ij}\tau_{ij}$ and $\rho_j \equiv \lambda_j/\mu_j$,

one can see that each station acts as an M/M/1 queue with arrival rate λ_j, service rate μ_j, and traffic intensity ρ_j under a Bernoulli-splitting policy.

Let Γ be the set of all Bernoulli-splitting policies, which are completely characterized by probabilities $\{\rho_{ij}, i \in \mathcal{R}, j \in \mathcal{S}\}$. In order to apply one step of the policy improvement algorithm, we need to determine the value function $V^\gamma(\mathbf{x})$ associated with a given policy $\gamma \in \Gamma$, which is defined as the expected total discounted reward obtained under policy γ when the initial state is $\mathbf{x} \in \mathcal{Q}$.

PROPOSITION 5. *Let $\gamma \in \Gamma$ be a policy having probabilities $\{\rho_{ij}, i \in \mathcal{R}, j \in \mathcal{S}\}$. The value function $V^\gamma(\mathbf{x})$ associated with γ is given by*

$$V^\gamma(\mathbf{x}) = \sum_{j \in \mathcal{S}} \left[\frac{\mu_j r_j}{\alpha} - \frac{2\mu_j r_j}{\lambda_j + \alpha - \mu_j + \eta_j} \left(\frac{\lambda_j + \mu_j + \alpha - \eta_j}{2\lambda_j} \right)^{x_j} \right], \tag{9}$$

where $\eta_j \equiv \sqrt{(\mu_j + \lambda_j + \alpha)^2 - 4\lambda_j \mu_j}$.

The value function given by (9) is separable by station because each station is a separate M/M/1 queue under any given Bernoulli-splitting policy. By using (9), we can compute the marginal value function $M_j^\gamma(\mathbf{x}) \equiv V^\gamma(\mathbf{x} + \mathbf{e}_j) - V^\gamma(\mathbf{x})$ associated with policy $\gamma \in \Gamma$ as

$$M_j^\gamma(\mathbf{x}) = \frac{\mu_j r_j}{\lambda_j} \left(\frac{\mu_j - \lambda_j + \alpha - \eta_j}{\mu_j - \lambda_j - \alpha - \eta_j} \right) \left(\frac{\mu_j + \lambda_j + \alpha - \eta_j}{2\lambda_j} \right)^{x_j}. \tag{10}$$

Applying one step of the policy improvement algorithm corresponds to using $M_j^\gamma(\mathbf{x})$ in place of $M_j(\mathbf{x})$ in (2). Namely, the one-step policy improvement heuristic (PIH) that uses static policy $\gamma \in \Gamma$ sends casualties to the station with the largest index $\tau_{ij} M_j^\gamma(\mathbf{x})$. Similar to the myopic heuristic, the PIH will consist of a set of linear switching curves because of the exponential structure of its index; see (10). In the remainder of this section, we focus on two different ways of choosing a Bernoulli-splitting policy γ for use in the policy improvement heuristic.

5.2.1. Optimal Bernoulli-splitting policy Because applying a step of the policy improvement algorithm results in a dynamic policy that performs at least as well as the initial static policy, intuitively we would expect that the one-step policy improvement heuristic should work well if we initialize it with the optimal Bernoulli-splitting policy, i.e., the policy that yields the largest expected total discounted reward within Γ. Finding this optimal static policy (denoted by ST-O) within Γ requires maximizing $V^\gamma(\mathbf{x})$, which is given by (9), subject to $\sum_{j \in \mathcal{S}} \rho_{ij} = 1, \forall i \in \mathcal{R}$, and $\rho_{ij} \geq 0, \forall i \in \mathcal{R}$ and $j \in \mathcal{S}$. Unfortunately, the objective function is nonlinear in ρ_{ij} and hence the problem is difficult to solve except for cases with a small number of locations and stations, where it can be solved numerically. In our numerical experiments presented in Section 6, we report the performance of the policy improvement heuristic with optimal Bernoulli-splitting policy (denoted by PIH-O) for cases with a small number of locations and stations.

5.2.2. Fluid approximation As an alternative to determining the optimal static policy for the original model, we consider finding the optimal static policy for a fluid approximation of the problem. A fluid approximation is one in which the discrete entities are treated as continuous, and transportation of casualties to the stations and subsequent service thereof is equivalent to the flow of a fluid. Recall that the inverse of τ_{ij} is defined to be the mean travel time to station j from location i. Therefore, in a fluid approximation, the fluid flowing from location i to station j first arrives at time $1/\tau_{ij}$, flows at rate $\rho_{ij}\tau_{ij}$ units of fluid per time unit, and continuously earns a reward of r_j per unit of fluid. When this reward is discounted to time zero at discount factor α, the total reward earned by fluid flowing from location i to station j becomes $\int_{1/\tau_{ij}}^{\infty} \rho_{ij}\tau_{ij}r_j e^{-\alpha t}dt = \rho_{ij}\tau_{ij}r_j e^{-\frac{\alpha}{\tau_{ij}}}/\alpha$, assuming that the stations are initially empty and idle. One can then find the optimal static policy for the fluid approximation (denoted by ST-F) by solving the following linear program:

$$\max \quad \sum_{i\in\mathcal{R}}\sum_{j\in\mathcal{S}} \tau_{ij} r_j e^{-\frac{\alpha}{\tau_{ij}}} \rho_{ij} \tag{F}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{R}} \rho_{ij}\tau_{ij} \le \mu_j, \quad \forall j\in\mathcal{S} \tag{11}$$

$$\sum_{j\in\mathcal{S}} \rho_{ij} \le 1, \quad \forall i\in\mathcal{R} \tag{12}$$

$$\rho_{ij} \ge 0, \quad \forall i\in\mathcal{R}, j\in\mathcal{S}. \tag{13}$$

Constraint set (11) guarantees that the fluid does not arrive to station j at a total rate faster than μ_j. A solution where fluid arrives at a total rate faster than μ_j would actually be feasible, but it could not increase the expected total discounted reward because the fluid can never flow out of station j (and thus earn a reward) at a rate faster than μ_j. In constraint set (12), we allow the total probability for a given incident location to be less than one to guarantee that (F) has a feasible solution in cases where transportation resources have a high capacity relative to treatment facilities. Because having arrivals to a station at a rate faster than the station's service rate would have no benefit in the fluid formulation, it is possible that in the optimal solution to (F), not all transportation capacity is used.

Formulation (F) is a continuous relaxation of the generalized assignment problem (see, e.g., Öncan 2007). If the problem is relaxed by removing constraints (11), which would be a good approximation when stations are fast compared to transportation resources, then the resulting relaxation could be solved optimally via greedy choice. One can also use the idea of greedy choice to obtain a feasible solution to (F) as follows. Initialize by setting $\rho_{ij} = 0$ for all $i \in \mathcal{R}, j \in \mathcal{S}$. Next, make a list of all station-location pairs $(i,j), i\in\mathcal{R}, j\in\mathcal{S}$ in a descending order of $\tau_{ij} r_j e^{-\frac{\alpha}{\tau_{ij}}}$. Then, proceeding in order through the list, assign the largest possible non-negative value to ρ_{ij} such that constraints (11) and (12) are satisfied. Stop when each station-location pair in the list has been

considered once. We call this procedure Algorithm 1, which is stated more formally in the Electronic Companion together with a result (Proposition 7) that shows that Algorithm 1 returns a feasible solution to (F). We denote the resulting "greedy" policy by ST-G. In our numerical experiments, we will report the results of the policy improvement heuristic using both the probability assignment from the exact solution to (F) (denoted by PIH-F) and from the approximate solution based on Algorithm 1 (denoted by PIH-G).

Figure 4(b) shows the structure of the PIH-G policy for the same example whose optimal policy is shown in Figure 3. By examining Figures 3 and 4, we see that in this instance the myopic policy seems to do a better job approximating the optimal policy for location 1 while PIH-G seems to do a better job of approximating the policy for location 2. Note that both the myopic policy and the PIH-G heuristic correctly ruled out the action of sending a location 1 casualty to facility A and a location 2 casualty to facility B in the same decision epoch. Indeed, it is easy to show that both the myopic policy and the PIH policies share the structural property of optimal policies given in Proposition 1.

5.3. Modified heuristics for intermittent status updates

Until now, we have assumed that state information $\{x_j, j \in \mathcal{S}\}$ is available at every decision epoch. While the type of information needed (namely, the number of casualties in the system at each facility) is simple, it may not be easy to obtain this information continuously in the aftermath of a disaster. After such an event, communication between the decision maker and the medical facilities may be disrupted due to outages in telecommunication equipment or due to the chaotic nature of such an event. Nevertheless, the decision maker should be able to keep making decisions dynamically with the most up-to-date information in hand. In this section, we describe how our heuristic policies can be modified for the case where state information is available only intermittently.

In order to use the myopic policy or the policy improvement heuristic with intermittent status updates, the decision maker must keep track of the following information: (i) the state of each facility at the last status update $\{x_j, j \in \mathcal{S}\}$, (ii) the number of casualties distributed to each facility since the last status update, which we denote by $\{y_j, j \in \mathcal{S}\}$, and (iii) the amount of time since the last status update, which we denote by $s \geq 0$. The decision maker then can approximate the current state of each facility to be $(x_j + y_j - X_j(s))^+$, where $X_j(s)$ is a Poisson random variable with mean $\mu_j s$, and $z^+ \equiv \max\{z, 0\}$ for any real number z. Note that $X_j(s)$ is the number of potential departures at facility $j \in \mathcal{S}$ during a time period of length s, and is an approximation for the number of departures from facility j. As long as x_j is not too close to zero and s is not too large, we would expect this approximation to work well.

Using the above approximation of the state information, we modify the myopic policy such that from each location i, we send a casualty to station j with the largest index

$$\tau_{ij} r_j \mathsf{E}\left[\left(\frac{\mu_j}{\mu_j+\alpha}\right)^{\left(x_j+y_j-X_j(s)\right)^+ +1}\right].$$

Similarly, we modify the policy improvement heuristic for any given $\{\rho_{ij}, i\in\mathcal{R}, j\in\mathcal{S}\}$, such that from each location i, we send a casualty to station j with the largest index

$$\tau_{ij}\frac{\mu_j r_j}{\lambda_j}\left(\frac{\mu_j-\lambda_j+\alpha-\eta_j}{\mu_j-\lambda_j-\alpha-\eta_j}\right)\mathsf{E}\left[\left(\frac{\mu_j+\lambda_j+\alpha-\eta_j}{2\lambda_j}\right)^{\left(x_j+y_j-X_j(s)\right)^+}\right].$$

By using the approximation that $X_j(s)$ is a Poisson random variable, we can rewrite these two indices as

$$\tau_{ij} r_j \left[b_1 F(x_j+y_j;t) + b_2(1-F(x_j+y_j;s\mu_j))\right], \tag{14}$$

where $F(\cdot;\omega)$ denotes the cumulative distribution function of a Poisson random variable with mean ω. The specific values of b_1, b_2, and t under the myopic and PIH policies are given in Table EC.2 in the Electronic Companion.

6. Numerical Results

In our numerical analysis, we first compared the heuristic policies developed in Section 5 to the optimal solution of the formulation described in Section 3. Later, we conducted a simulation study by means of which we tested our heuristic policies under more realistic conditions than those of our mathematical model described in Section 3.

6.1. Numerical experiments: comparison with the optimal policy

In order to test the performance of our heuristics as a solution to the optimization problem (1), we conducted a set of numerical experiments with 1,000 randomly generated instances, each with two locations and two facilities. Using a small number of locations and facilities enables us to find the optimal solution by solving (1) numerically. Each random instance has service rates (μ_j) and transportation rates (τ_{ij}) chosen uniformly in $[1,10]$. All rewards are set to one, which means that the performance measure is the expected discounted throughput.

For each instance, we calculated the optimal policy using the value iteration algorithm under a truncated state space with 625 states. We then calculated the expected discounted throughputs for the optimal policy, myopic policy, and the three variations of the policy improvement heuristic, i.e., PIH-F, PIH-G, and PIH-O, assuming that all stations are initially empty and all transportation resources are idle. For comparison, we also computed the performance of the three static policies

discussed in Section 5.2, namely, ST-F, ST-G, and ST-O policies. We repeated each experiment with $\alpha \in \{0.1, 0.7, 2.0\}$. The discount factor $\alpha = 0.7$ corresponds to a discount of approximately 50% of the reward per unit time. The other discount factors $\alpha = 0.1$ and $\alpha = 2.0$ are chosen to provide a comparison under (relatively) light and heavy discounting to represent less or more urgent conditions, respectively.

The results from these numerical experiments are shown in Table 1. The entries in the third through eighth columns summarize the distribution of the percentage of the optimal discounted throughput attained by each heuristic policy based on 1,000 instances. Specifically, entries in the third through eight columns provide the minimum performance (Min), first quartile (Q1), median (Q2), third quartile (Q3), maximum (Max), and mean in the given order.

Table 1 shows that all dynamic heuristics perform well in terms of the mean performance, achieving an average of at least 92% of the value of the optimal policy. However, there is a clear tradeoff between simplicity and performance when we consider the worst-case performances. More specifically, the myopic policy has the simplest index to calculate, but performs poorly in a small number of instances. On the other hand, PIH performs very well under all three initial static policies for all performance statistics, being essentially indistinguishable from the optimal dynamic policy in the mean performance. It is especially important to note that starting from the greedy or fluid optimal policies provide a similar performance compared with starting from the optimal static policy, because calculating the optimal static policy is practical only for problems with a relatively small number of locations and stations.

Another observation from Table 1 is that the myopic heuristic is the most sensitive heuristic to the discount factor α among all dynamic policies considered. Although the myopic heuristic performs comparably with the other dynamic heuristics when α is large, its performance (especially the worst-case performance) degrades with smaller α. This can be explained by the fact that when α is small the problem approaches a problem of maximizing the long-run average reward, and hence, the decisions in future epochs (which the myopic policy does not take into account) become relatively more important. On the other hand, static policies appear to perform notably better for smaller α. The performances of all policy improvement heuristics are extremely good regardless of the value of α, which is another reason why one might prefer these policies.

While we have shown that our heuristics produce numerical results that are very close to the optimal policy, these numerical experiments alone are not sufficient to draw conclusions about the heuristics' applicability. In particular, it is important to consider scenarios with more than two incident locations and two facilities, and where the restrictive modeling assumptions do not hold. Adding more locations and facilities quickly makes the MDP intractable due to the curse of dimensionality, and relaxing the assumptions (such as the distributional assumption on travel

Table 1 Performance of heuristic policies as a percentage of optimal discounted throughput.

α	Policy	Min	Q1	Q2	Q3	Max*	Mean
0.1	ST-F	51	78	85	88	91	81
	ST-G	48	75	85	88	91	81
	ST-O	51	79	92	97	99	87
	Myopic	30	94	99	100	100	92
	PIH-F	89	100	100	100	102	99
	PIH-G	95	100	100	100	102	100
	PIH-O	95	99	100	100	102	99
0.7	ST-F	49	66	72	75	80	70
	ST-G	46	65	71	75	80	69
	ST-O	52	72	81	87	94	79
	Myopic	86	97	99	100	100	98
	PIH-F	87	99	100	100	102	99
	PIH-G	91	100	100	100	102	100
	PIH-O	91	98	100	100	102	99
2.0	ST-F	33	59	63	67	73	62
	ST-G	33	57	63	66	73	61
	ST-O	51	66	73	78	87	72
	Myopic	92	98	99	100	100	98
	PIH-F	88	99	100	100	102	99
	PIH-G	90	99	100	100	102	99
	PIH-O	90	98	100	100	102	99

* Because we used state-space truncation to find the optimal policy, some of the heuristics performed better than the so-called optimal solution in certain cases.

times) makes it impossible to find the optimal performance as well as the performance of heuristic policies numerically. Therefore, we use simulation to examine the performance of our heuristic policies in a more realistic environment in Section 6.2.

6.2. Simulation experiments

We use a simulation study to test the performances of our policies in the case where some of the assumptions made in our MDP formulation are relaxed. In particular, we relax the assumptions that ambulance travel times and treatment times are exponentially distributed, the return time for ambulances to the locations is negligible, trips can be preempted, and the number of casualties is infinite. Furthermore, in this simulation study, we consider the case where stations have multiple servers, which might be a more realistic representation of servers (such as beds) at a treatment facility. Finally, we repeated the simulation study both when the information on the state of facilities are available at any given point in time and when the state information is updated intermittently.

In order to construct a simulation study that is geographically realistic, similar to the situation depicted in Figure 1, we suppose that a trauma-related mass-casualty incident has occurred in an urban area that covers 100 square miles (ten miles by ten miles), resulting in $n \in \{3, 4, 5\}$ incident locations with 50 casualties each. Each location has two transportation resources, which can travel to the medical facilities at an average speed of 40 mph. There are three Level I or Level II trauma centers, two of which are located within the urban area and one of which is located 15 to 30 miles outside the area. Travel between points occurs according to Manhattan distance (i.e., distances are calculated using the L^1-norm), and the travel time has a lognormal distribution. This distributional assumption is based on the empirical observation that the lognormal distribution best represents travel times of ambulances in Ingolfsson et al. (2008).

To determine a distribution for emergency department (ED) treatment times, we used the 2010 National Hospital Ambulatory Medical Care Survey (NHAMCS) (National Center for Health Statistics 2010), from which we obtained a sample of patients arriving to emergency departments with traumatic injuries (sample size of 10,754). Using this data, we calculated the length of stay for each patient by taking the length of the visit (determined from arrival and departure times) and subtracting the waiting time. Patients with total ED times exceeding one day (less than one percent of the observations) were excluded. The mean ED length of stay of approximately 124 minutes for this data is consistent with the existing medical literature on surge capacity, which assumes an ED length of stay of approximately 2.5 hours (Bayram et al. 2010). Using SAS software, we fit the length of stay to the exponential distribution and to three other distributions that are commonly used to model the time taken to perform a task (Law 2007); namely, Weibull, Gamma, and Lognormal. For each distribution, parameters were estimated via the maximum likelihood estimation method and the distributions were compared with the empirical data using the Kolmogorov-Smirnov and Anderson-Darling tests. The Lognormal distribution with parameters $\mu = 4.32$ and $\sigma = 1.00$ was chosen as the best fit based on the smallest Kolmogorov-Smirnov statistic (4.55 versus 5.00 for Gamma, which is the next best fit) and the smallest Anderson-Darling statistic (65.18 versus 93.98 for Gamma).

In our simulation study, we considered four different policies: a baseline static policy, where all casualties are sent to the *nearest facility*; a baseline dynamic policy, where all casualties are sent to the facility that has the *shortest queue* (in expected waiting time) at the time they begin transportation; the myopic policy (see Section 5.1); and the PIH-G policy (see Section 5.2). Myopic policy and PIH-G were chosen as candidate policies because they strike a balance between simplicity and effectiveness, and thus of the policies we have developed, have the maximum potential for implementation. "Nearest-facility" is a policy that is used in practice in daily emergencies as well as in major events. For example, when sarin gas was released in 15 Tokyo subway stations in 1995,

providers were unable to get information about hospital availability due to poor communication, and most patients were taken to the nearest hospital (Okumura et al. 1998). While it is not clear whether the "shortest-queue" policy has been used in practice, intuitively it is a natural alternative for cases where providers are concerned about bypassing busy hospitals due to congestion; see, e.g., Hick et al. (2011). Furthermore, it is simple and is proven to be optimal under certain conditions (see, e.g., Proposition 3).

6.2.1. A randomized study In the first part of our simulation experiments, 50 scenarios are generated randomly to observe the average performance of our heuristics. More specifically, in this randomized study, locations of the incidents and the facilities within the urban area are generated randomly for each scenario using a two-dimensional Uniform random variate, while the distance from the suburban facility to the center of the urban area is generated using a Uniform random variate. These scenarios also differ in the service level of the facilities and their capacities. In particular, the generated trauma centers are either Level I (with probability 0.58), which have a median of 39 ED beds, or Level II (with probability 0.42), which have a median of 25 ED beds. These probabilities for Level I and II trauma centers are obtained from a national survey by MacKenzie and Hoyt (2003), while the numbers of beds are taken from Rivara et al. (2006). Because Rivara et al. (2006) presented only the median number of beds, we added variability by generating the number of beds for each trauma center according to a Uniform random variate with range 10, centered about the median. In general, utilization of ED beds at a trauma center is quite high (MacKenzie and Hoyt 2003), and beds for patients from a disaster would become available over a period of time as patients in the ED would be moved to create surge capacity (Hick et al. 2004, Peleg and Kellermann 2009). Therefore, we assumed that the ED beds became available at a constant rate over a period of two hours after the time of the incident as existing patients are discharged or moved to other facilities. In order to apply our heuristics to the case with multiple beds, we set $\mu_j = B_j/m$, where B_j is the number of beds at facility j and m is the mean length of stay.

All simulations were run in MATLAB software using 100 replications and common random numbers across policies. In Table 2, we report the minimum, mean, and maximum percentage improvement in discounted throughput by the myopic and PIH-G policies over the two baseline policies across the 50 scenarios in columns 4 through 6 and 9 through 11. We also report the number of scenarios (out of 50) where each heuristic provides a better performance (at a significance level of 0.05) and the number of scenarios where each heuristic provides a worse performance (at a significance level of 0.05) in columns 7, 8, 12, and 13.

The results of the randomized simulation study show that even when the assumptions of the MDP model are violated, the heuristics that resulted from our analysis were still valuable. Both

Table 2 Simulation performance over 50 randomly generated scenarios with n incident locations.

			vs. Nearest Facility					vs. Shortest Queue				
α	n	Policy	Min	Mean	Max	Better	Worse	Min	Mean	Max	Better	Worse
0.1	3	Myopic	0%	4%	55%	14	0	7%	36%	64%	50	0
	3	PIH-G	-9%	2%	51%	16	22	13%	33%	65%	50	0
	4	Myopic	0%	6%	40%	23	0	-22%	27%	67%	43	5
	4	PIH-G	-12%	7%	41%	24	16	8%	27%	56%	50	0
	5	Myopic	-1%	8%	69%	25	0	-17%	23%	59%	46	3
	5	PIH-G	-12%	8%	93%	25	16	1%	21%	53%	49	0
0.7	3	Myopic	-6%	18%	111%	24	0	51%	145%	290%	50	0
	3	PIH-G	-41%	-9%	95%	8	31	33%	87%	234%	50	0
	4	Myopic	-3%	29%	113%	35	0	53%	132%	242%	50	0
	4	PIH-G	-44%	-5%	64%	16	27	19%	65%	128%	50	0
	5	Myopic	1%	25%	144%	39	0	26%	104%	216%	50	0
	5	PIH-G	-45%	-5%	117%	10	28	8%	52%	111%	50	0
2.0	3	Myopic	-13%	19%	165%	17	0	76%	235%	497%	50	0
	3	PIH-G	-37%	-3%	125%	6	19	53%	170%	425%	50	0
	4	Myopic	-8%	26%	117%	20	0	85%	204%	429%	50	0
	4	PIH-G	-39%	3%	85%	15	21	63%	142%	275%	50	0
	5	Myopic	-8%	19%	128%	20	0	51%	163%	350%	50	0
	5	PIH-G	-35%	1%	125%	6	17	44%	119%	236%	50	0

heuristics consistently outperform the policy of choosing the hospital with the shortest queue. This is due to the fact that choosing the shortest queue ignores the fact that the transportation resource becomes unavailable for an extended period of time when the hospital with the shortest queue is far away. In many scenarios, both heuristics were also better than choosing the nearest hospital, since choosing the nearest hospital completely ignores congestion. However, performance of the policy improvement heuristic (PIH-G) deteriorated compared with choosing the nearest facility, especially when the discount rate is high. This can be explained by our numerical observation that the PIH-G policy tends to choose distant facilities more often, which sometimes turns out to be a poor choice for future patients in the simulation, where such a decision is taken non-preemptively. The myopic policy, which is not the best performer in the numerical results presented in Section 6.1, performs better in the simulation study. When coupled with the fact that the myopic policy index is also simple to calculate, the myopic policy appears to be a strong candidate for practical application.

We next repeated the above simulation study with state information updates only every 15 minutes, using the modified heuristics developed in Section 5.3. To simulate intermittent state updates, we maintained three vector-valued state variables: the true number of patients at each

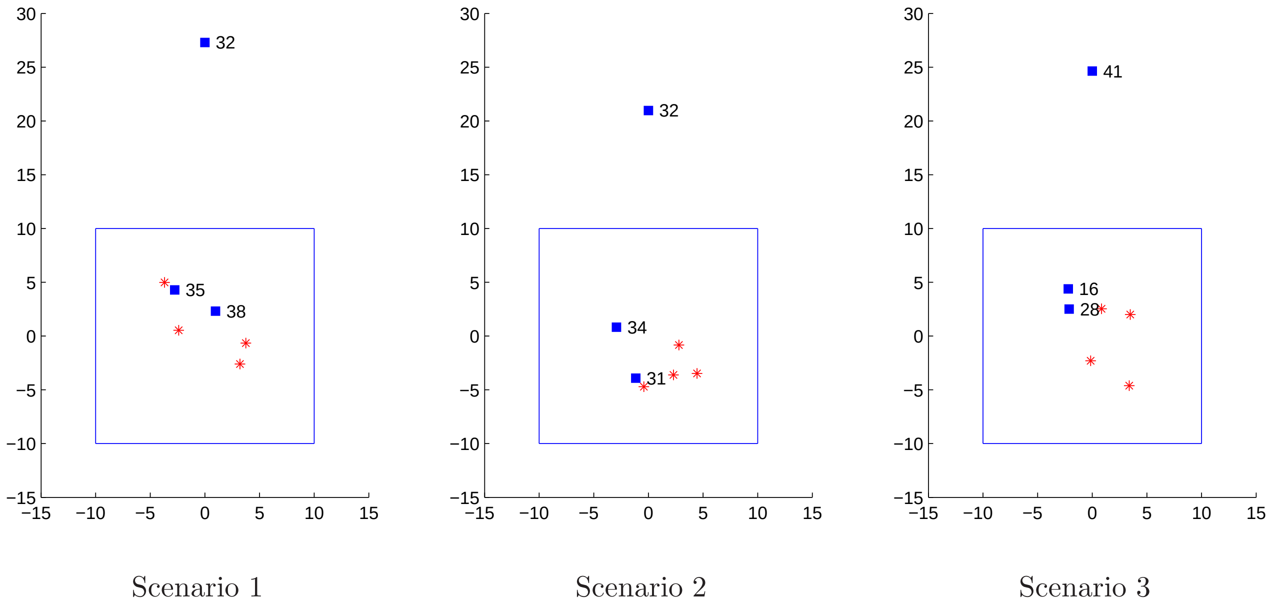
Table 3 Simulation performance over 50 randomly generated scenarios with state updates every 15 minutes.

			vs. Nearest Facility					vs. Shortest Queue				
α	n	Policy	Min	Mean	Max	Better	Worse	Min	Mean	Max	Better	Worse
0.1	3	Myopic	-1%	3%	50%	8	0	7%	34%	62%	50	0
	3	PIH-G	-8%	3%	56%	17	19	16%	35%	68%	50	0
	4	Myopic	-1%	4%	31%	22	0	-22%	25%	66%	43	5
	4	PIH-G	-9%	8%	41%	26	14	9%	29%	62%	50	0
	5	Myopic	-1%	7%	58%	22	0	-21%	21%	59%	44	3
	5	PIH-G	-9%	9%	93%	27	12	3%	23%	53%	50	0
0.7	3	Myopic	-5%	16%	99%	20	0	42%	140%	261%	50	0
	3	PIH-G	-32%	-2%	100%	10	27	38%	101%	235%	50	0
	4	Myopic	-2%	28%	107%	33	0	47%	129%	249%	50	0
	4	PIH-G	-32%	2%	75%	17	25	23%	79%	153%	50	0
	5	Myopic	1%	23%	135%	34	0	21%	101%	218%	50	0
	5	PIH-G	-33%	1%	116%	15	25	9%	62%	126%	50	0
2.0	3	Myopic	-11%	18%	164%	17	0	83%	231%	457%	50	0
	3	PIH-G	-29%	7%	150%	13	8	62%	198%	436%	50	0
	4	Myopic	-4%	25%	114%	21	0	66%	203%	440%	50	0
	4	PIH-G	-27%	12%	97%	16	12	62%	168%	353%	50	0
	5	Myopic	-3%	18%	120%	17	0	42%	161%	374%	50	0
	5	PIH-G	-20%	9%	121%	8	4	39%	137%	291%	50	0

facility, which was used internally in the simulation; the observed number of patients at each facility, which was updated only every 15 minutes by retrieving the true value; and the number of patients sent to each facility since the last update. The state information updates were simulated in the same way for all three dynamic heuristics, including the shortest-queue policy, i.e., for the shortest-queue policy, we chose the queue with shortest expected queue length $\mathsf{E}\left[(x_j + y_j - X_j(s))^+/\mu_j\right]$, where x_j, y_j, and $X_j(s)$ are defined in Section 5.3. The results of these simulation runs are presented in Table 3, where the column headers are the same as in Table 2. The change in performance improvement was not substantially worse than in the case with real-time state information (in fact, the performance of PIH-G improved somewhat in some scenarios), suggesting that the modified myopic and PIH-G policies can be still effective in situations where real-time state information is not available.

6.2.2. Closer examination of three scenarios The randomized design of the study presented in Section 6.2.1 obscures insights into the characteristics of scenarios where different heuristics perform particularly well or particularly poorly. In order to more clearly understand how the different heuristics could be effectively used, we now examine three representative scenarios, which

Figure 5 (x, y)-coordinates of the locations of hospitals and incidents in three scenarios used in the in-depth examination of the simulation study. Stars indicate incident locations, while boxes indicate hospitals, with total bed capacity listed next to each hospital.



are shown in Figure 5, from the randomized study with four geographically separated incidents. In Figure 5, stars indicate the locations of incidents, while boxes indicate the locations of the trauma centers, with total bed capacity listed next to each trauma center. All other system parameters and distributions are the same as in Section 6.2.1. The performance results for these three scenarios are shown in Table 4.

Scenario 1 presents a relatively "easy" decision problem, where two of the incidents are located nearest the hospital with 35 beds and two of the incidents are located nearest the hospital with 38 beds. In this case, the myopic policy performs comparably with sending patients to the nearest facility. The policy improvement heuristic performs worse than the nearest-facility policy, and the shortest-queue policy performs worst of all. Intuitively, one can see that in this scenario, capacity is not severely limited, the incidents are close to the facilities, and the nearest-facility policy results in a balanced load to each of the two city facilities. Therefore, there is not much to gain by using a dynamic policy, and it is even possible to lose by using a poorly constructed dynamic policy (such as the shortest-queue policy, which will unnecessarily utilize the suburban facility).

In Scenario 2, although the facility capacities are similar to Scenario 1, the load is no longer balanced by sending patients to the nearest facility, as all incidents are nearest to the facility with 31 beds. In this case, the two proposed dynamic policies provide substantial improvement, increasing the throughput by about $15\% - 50\%$ for myopic compared to using the nearest-facility policy and by about 35% to well over 100% for myopic compared to using the shortest-queue policy.

Table 4 95% confidence intervals on the discounted throughput under the myopic, PIH-G, shortest-queue, and nearest-facility policies for the scenarios given in Figure 5.

Scenario	α	Myopic	PIH-G	Shortest Queue	Nearest Facility
1	0.1	121.9 ± 0.6	110.2 ± 0.6	74.7 ± 0.6	121.9 ± 0.6
	0.7	19.1 ± 0.4	11.1 ± 0.3	5.5 ± 0.2	18.1 ± 0.4
	2.0	2.2 ± 0.1	1.6 ± 0.1	0.5 ± 0.1	2.4 ± 0.2
2	0.1	107.1 ± 0.8	105.8 ± 0.7	79.2 ± 0.6	93.1 ± 1.0
	0.7	15.3 ± 0.4	10.1 ± 0.3	6.1 ± 0.2	9.8 ± 0.4
	2.0	1.8 ± 0.1	1.5 ± 0.1	0.6 ± 0.1	1.2 ± 0.1
3	0.1	98.9 ± 0.4	91.5 ± 0.6	75.7 ± 0.5	52.8 ± 1.1
	0.7	12.4 ± 0.3	8.2 ± 0.3	5.9 ± 0.2	4.4 ± 0.2
	2.0	1.5 ± 0.1	1.3 ± 0.1	0.6 ± 0.1	0.5 ± 0.1

Finally, Scenario 3 has an unbalanced geographical distribution (similar to Scenario 2), but with far lower capacity at the urban facilities. As one might expect, the proposed dynamic policies again provide substantial improvement over the baseline policies (in the case of heavier discounting at a rate of 0.7 or 2.0, the improvement was over 100%), but now the shortest-queue policy outperforms the nearest-facility policy. Intuitively, in Scenario 3 one can see that because of the lower capacity, judicious use of the suburban facility will be necessary to ensure an acceptable throughput of patients. Thus, in this scenario, we will see the largest benefit from using a dynamic policy as compared to sending patients to the nearest facility.

Several insights emerge from this closer examination of specific scenarios. When incidents are geographically dispersed and nearby facilities have sufficient capacity, sending patients to the nearest facility achieves a reasonable throughput. On the other hand, when patients are clustered around a single facility and this facility does not have sufficient capacity, then a dynamic policy ensures that the closest facility does not become overwhelmed, but the dynamic policy must be constructed carefully. In particular, a dynamic policy that over-utilizes distant facilities (such as the shortest-queue policy used as a baseline) may do more harm than good by lengthening travel times.

7. Conclusion

In the aftermath of a disaster, distributing casualties to medical facilities where they can receive care is a major operational challenge, especially when transportation and treatment resources are limited, and there are multiple incident sites that are geographically separated. In this study, we developed a queueing model to study the casualty distribution problem, which led to several analytical and numerical results that could be used in designing emergency response plans.

One of our most important findings is that incorporating the census information about medical facilities into decision making can improve the outcome of a response effort substantially compared

to simply distributing casualties to the geographically nearest facility. We have observed that it is especially crucial to use policies that take into account changing hospital congestion levels during disasters where hospital capacities are severely limited and incident sites are geographically "unbalanced" with respect to the location of the facilities (e.g., a single facility is the nearest facility for most or all incident locations). However, designing a simple yet effective dynamic policy that takes into account the state of medical facilities is not a trivial task. A "good" dynamic policy should take into account the delay the patient will encounter if sent to a given medical facility, the length of time that an ambulance will be used in traveling to the designated facility, the quality of service that would be given at that facility, and the effects of decisions on future patients. For example, we have observed that simply sending patients to a facility that is reasonably close to the incident and has the shortest queue may perform terribly. The good news is that the dynamic policies that we proposed in this study perform very well in the scenarios considered in our simulation study. In particular, our myopic policy, which is a simple index policy that makes allocations independently for each incident location, has shown a strong performance throughout our entire numerical and simulation experiments.

There are several ways that our heuristic policies (especially the myopic policy) can be used. First, they can be incorporated into training exercises to help emergency response planners better understand their capability to respond to specific disaster scenarios. Our heuristics can be also used to develop some specific rules of thumbs that could be incorporated into emergency response plans. For example, one can use our myopic heuristic to determine the types of events for which the emergency management should start including distant hospitals that do not typically serve the region the incidents took place. Finally, we believe that our myopic policy is simple enough to be implemented in a decision support tool that could assist in the response effort in real time whenever information on the hospital congestion levels can be obtained on a regular basis.

Acknowledgments

The authors thank James E. Winslow, MD, Medical Director for the North Carolina Office of Emergency Medical Services, for providing valuable advice and perspective in undertaking this research. This work was partially supported by the National Science Foundation under Grant Numbers CMMI–0927607 and CMMI–1234212.

References

Allon, G., S. Deo, W. Lin. 2013. The impact of size and occupancy of hospital on the extent of ambulance diversion: theory and evidence. *Operations Research* **61**(3) 544–562.

Ansell, P.S., K.D. Glazebrook, C. Kirkbride. 2003. Generalised "join the shortest queue" policies for the dynamic routing of jobs to multi-class queues. *Journal of the Operational Research Society* **54** 379-389.

Argon, N.T., L. Ding, K. D. Glazebrook, S. Ziya. 2009. Dynamic routing of customers with general delay costs in a multiserver queueing system. *Probability in the Engineering and Informational Sciences* **23**(2) 175–203.

Armony, M. 2005. Dynamic routing in large-scale service systems with heterogeneous servers. *Queueing Systems* **51**(3-4) 287–329.

Bayram, J. D., S. Zuabi, I. Subbarao. 2010. Disaster metrics: Quantitative benchmarking of hospital surge capacity in trauma-related multiple casualty events. *Disaster Medicine and Public Health Preparedness* **5**(2) 117–124.

Bish, D., E. Agca, R. Glick. 2011. Decision support for hospital evacuation and emergency response. *Annals of Operations Research* doi:10.1007/s10479-011-0943-y.

Brandeau, M. L., J. H. McCoy, N. Hupert, J-E. Holty, D. M. Bravata. 2009. Recommendations for modeling disaster responses in public health and medicine: A position paper the society for medical decision making. *Medical Decision Making* **29** 438–460.

Childers, A., K. Taaffe. 2010. Healthcare facility evacuations: Lessons learned, research activity, and the need for engineering contributions. *Journal of Healthcare Engineering* **1**(1) 125–140.

Deo, S., I. Gurvich. 2011. Centralized vs. decentralized ambulance diversion: A network perspective. *Management Science* **57**(7) 1300–1319.

Glazebrook, K. D., C. Kirkbride, J. Ouenniche. 2009. Index policies for the admission control and routing of impatient customers to heterogeneous service stations. *Operations Research* **57**(4) 975–989.

Gong, Q., R. Batta. 2007. Allocation and reallocation of ambulances to casualty clusters in a disaster relief operation. *IIE Transactions* **39**(1) 27–39.

Ha, A. Y. 1997. Optimal dynamic scheduling policy for a Make-To-Stock production system. *Operations Research* **45**(1) 42–53.

Hick, J. L., D. Hanfling, J. L. Burstein, et al. 2004. Health care facility and community strategies for patient care surge capacity. *Annals of Emergency Medicine* **44**(3) 253–261.

Hick, J. L., P. Penn, D. Hanfling, et al. 2003. Establishing and training health care facility decontamination teams. *Annals of Emergency Medicine* **42**(3).

Hick, J. L., D. M. Weinstock, C. N. Coleman, et al. 2011. Health care system planning for and response to a nuclear detonation. *Disaster Medicine and Public Health Preparedness* **5**(Supplement 1) S73–88.

Hordijk, A., G. Koole. 1992. On the assignment of customers to parallel queues. *Probability in the Engineering and Informational Sciences* **6**(4) 495–511.

Hrdina, C. M., C. N. Coleman, S. Bogucki, et al. 2009. The "RTR" medical response system for nuclear and radiological mass-casualty incidents: A functional TRiage-TReatment-TRansport medical response model. *Prehospital and Disaster Medicine* **24**(3) 167–178.

Ingolfsson, A., S. Budge, E. Erkut. 2008. Optimal ambulance location with random delays and travel times. *Health Care Management Science* **11**(3) 262–274.

Jotshi, A., Q. Gong, R. Batta. 2009. Dispatching and routing of emergency vehicles in disaster mitigation using data fusion. *Socio-Economic Planning Sciences* **43**(1) 1–24.

Krishnan, K.R. 1990. Joining the right queue: a state-dependent decision rule. *IEEE Transactions on Automatic Control* **35** 104-108.

Law, A.M. 2007. *Simulation Modeling and Analysis*. McGraw Hill.

Lippmann, S. 1975. Applying a new device in the optimization of exponential queueing systems. *Operations Research* **23**(4) 687–710.

MacKenzie, K. J., D. B. Hoyt. 2003. National inventory of hospital trauma centers. *Journal of the American Medical Association* **289**(12) 1515–1522.

Mandelbaum, A., A. L. Stolyar. 2004. Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized cμ-rule. *Operations Research* **52**(6) 836–855.

Maxwell, M. S., M. Restrepo, S. G. Henderson, H. Topaloglu. 2010. Approximate dynamic programming for ambulance redeployment. *INFORMS Journal on Computing* **22**(2) 266–281.

McLay, L. A., M. E. Mayorga. 2013. A dispatching model for server-to-customer systems that balances efficiency and equity. *Manufacturing & Service Operations Management* **15**(2) 205–220.

National Center for Health Statistics. 2010. 2010 NHAMCS micro-data file documentation. Tech. rep., U.S. Centers for Disease Control and Prevention. Available at ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS/doc2010.pdf. Accessed on November 9, 2012.

Okumura, T., K. Suzuki, A. Fukuda, et al. 1998. The tokyo subway sarin attack: Disaster management, part 1: Community emergency response. *Academic Emergency Medicine* **5**(6) 613-617.

Öncan, T. 2007. A survey of the generalized assignment problem and its applications. *INFOR* **45**(3) 123–141.

Peleg, K., A. L. Kellermann. 2009. Enhancing hospital surge capacity for mass casualty events. *Journal of the American Medical Association* **302**(5) 565–567.

Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley.

Rivara, F. P., A. B. Nathens, G. J. Jurkovich, R. V. Maier. 2006. Do trauma centers have the capacity to respond to disasters? *The Journal of Trauma: Injury, Infection, and Critical Care* **61**(4) 949–953.

Robenshtok, E., A. Eisenkraft, S. Luria. 2003. Preparedness of local and national health systems for a chemical terrorist attack. J. Shemer, Y. Shoenfeld, eds., *Terror and Medicine*.

Ross, S. 1983. *Introduction to Stochastic Dynamic Programming*. Academic Press.

Tanaka, H., A. Iwai, J. Oda, et al. 1998. Overview of evacuation and transport of patients following the 1995 Hanshin-Awaji earthquake. *The Journal of Emergency Medicine* **16**(3) 439–444.

Tayfur, E., K. Taaffe. 2009. A model for allocating resources during hospital evacuations. *Computers & Industrial Engineering* **57**(4) 1313–1323.

U.S. Department of Homeland Security. 2005. *National Planning Scenarios*. Available at http://media. washingtonpost.com/wp-srv/nation/nationalsecurity/earlywarning/NationalPlanningScenariosApril2005.pdf. Accessed on November 9, 2012.

van Mieghem, J. A. 1995. Dynamic scheduling with convex delay costs: The generalized c|mu rule. *The Annals of Applied Probability* **5**(3) 809–833.

Veatch, M. H., L. M. Wein. 1996. Scheduling a make-to-stock queue: Index policies and hedging points. *Operations Research* **44**(4) 634–647.

Whitt, W. 1986. Deciding which queue to join: Some counterexamples. *Operations Research* **34**(1) 55–62.

Zayas-Cabán, G., M. E. Lewis, M. Olson, S. Schmitz. 2013. Emergency medical service allocation in response to large scale events. *IIE Transactions on Healthcare Systems Engineering* **3**(1) 57–68.

Electronic Companion

In this electronic companion, we provide the proofs of analytical results presented in the main paper. We also present Algorithm 1, Proposition 7 and its proof, and Tables EC.1 and EC.2, which are referred to in the main paper.

Proof of Proposition 1. For $\mathbf{x} \in \mathcal{Q}$, if station j is $(\mathbf{x}, k)$-preferable to station l, then

$$\tau_{kj} M_j(\mathbf{x}) \geq \tau_{kl} M_l(\mathbf{x}). \tag{EC.1}$$

By the condition that

$$\frac{\tau_{ij}}{\tau_{il}} \geq \frac{\tau_{kj}}{\tau_{kl}},$$

or equivalently,

$$\tau_{kl} \geq \frac{\tau_{kj}\tau_{il}}{\tau_{ij}},$$

we have

$$\tau_{kl} M_l(\mathbf{x}) \geq \frac{\tau_{kj}\tau_{il}}{\tau_{ij}} M_l(\mathbf{x}), \tag{EC.2}$$

because $M_l(\mathbf{x})$ is non-negative (a fact whose proof we postpone until Proposition 2). By combining (EC.1) and (EC.2), we conclude that

$$\tau_{ij} M_j(\mathbf{x}) \geq \tau_{il} M_l(\mathbf{x}),$$

which implies that station j is (x, i)-preferable to station l.

To prove Propositions 2 and 3, we will make use of a finite-horizon version of the main dynamic program.

Definition EC.1 (Finite-horizon dynamic program). We define $V^n(\mathbf{x})$ to be the maximum expected reward earned when the system will run for n additional epochs, each epoch occurring according to a Poisson process with rate β. We have

$$V^{n+1}(\mathbf{x}) = \frac{1}{1+\alpha}\left(r(\mathbf{x}) + \sum_{j \in \mathcal{S}} \mu_j V^n(\mathbf{x} - I_j(\mathbf{x})\mathbf{e}_j) + \tau V^n(\mathbf{x}) + \sum_{i \in \mathcal{R}} \max_{j \in \mathcal{S}} \{\tau_{ij}(V^n(\mathbf{x} + \mathbf{e}_j) - V^n(\mathbf{x}))\} \right), \tag{EC.3}$$

for $n \geq 0$ and $\mathbf{x} \in \mathcal{Q}$, and we assume that $V^0(\mathbf{x}) = 0$.

We now state a known result, which allows us to make conclusions about the structure of $V(\cdot)$ by using a limiting property of $V^n(\cdot)$.

Lemma EC.1 **(Chapter 2.3, Proposition 3.1, Ross (1983))**. *For any bounded $V^0(\mathbf{x})$, we have $V^n(\mathbf{x}) \to V(\mathbf{x})$ uniformly as $n \to \infty$.*

Proof of Proposition 2. By Lemma EC.1 and our assumption that $V^0(\mathbf{x})=0$ for all $x \in \mathcal{Q}$, it is sufficient to show that the following inequalities hold for all values of $n \geq 0$ and $\mathbf{x} \in \mathcal{Q}$:

$$V^n(\mathbf{x}+\mathbf{e}_j) \geq V^n(\mathbf{x}), \quad \forall j \in \mathcal{S}. \tag{EC.4}$$

$$V^n(\mathbf{x}) + r_j \geq V^n(\mathbf{x}+\mathbf{e}_j), \quad \forall j \in \mathcal{S}. \tag{EC.5}$$

We first prove (EC.4) by induction. The base case is that $V^0(\mathbf{x}+\mathbf{e}_j) \geq V^0(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{Q}$, which is trivial because $V^0(\mathbf{x})=0$ for all $\mathbf{x} \in \mathcal{Q}$. We now show that if $V^n(\mathbf{x}+\mathbf{e}_j) \geq V^n(\mathbf{x})$ for all $j \in S$ and for all x then the same will hold for $n+1$. From (EC.3), we have

$$\begin{aligned}
(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) &= r(\mathbf{x}+\mathbf{e}_j) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k) \\
&\quad + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(\mathbf{x}+\mathbf{e}_j)\} \\
&\geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(\mathbf{x}+\mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(\mathbf{x})\} \\
&= (1+\alpha) V^n(\mathbf{x}),
\end{aligned}$$

where the inequality follows from the fact that $r(\mathbf{x}+\mathbf{e}_j) \geq r(\mathbf{x})$ for all $j \in \mathcal{S}$ and for all $\mathbf{x} \in \mathcal{Q}$, the fact that $\tau_l \geq \tau_{lk}$ for all $l \in \mathcal{R}$ and $k \in \mathcal{S}$, and the inductive hypothesis. Note that in the above inequality, we apply the inductive hypothesis to the terms in the first summation—namely, $V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k)$—for every case where $j \neq k$ or $j = k$ and $x_j > 0$. In the case where $j = k$ and $x_j = 0$, $V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k)$ is equivalent to $V^n(\mathbf{x}+\mathbf{e}_j - \mathbf{e}_j)$, which is $V^n(\mathbf{x})$; moreover, $V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k)$ is also equivalent to $V^n(\mathbf{x})$. Hence, applying the inductive hypothesis to this term is not necessary.

We now prove (EC.5) by induction. The base case states that $V^0(\mathbf{x}) + r_j \geq V^0(\mathbf{x}+\mathbf{e}_j)$ for all $j \in \mathcal{S}$ and for all x, which is trivial because $V^0(\mathbf{x}) = 0$ for all x and $r_j > 0$. We now show that if $V^n(\mathbf{x}) + r_j \geq V^n(\mathbf{x}+\mathbf{e}_j)$ for all $j \in \mathcal{S}$ and for all x then the same will hold for $n+1$.

$$\begin{aligned}
(1+\alpha)(V^{n+1}(\mathbf{x}) + r_j) &= r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{j\}} \mu_k V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) + \mu_j V^n(\mathbf{x} - I_j(\mathbf{x})\mathbf{e}_j) + \tau V^n(\mathbf{x}) \\
&\quad + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_k) - V^n(\mathbf{x}))\} + (1+\alpha) r_j \\
&\geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{j\}} \mu_k V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) + \mu_j V^n(\mathbf{x} - I_j(\mathbf{x})\mathbf{e}_j) \\
&\quad + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(\mathbf{x}+\mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(\mathbf{x})\} + r_j,
\end{aligned}$$

since $\alpha > 0$. Using the fact that $\beta = \sum_{j\in\mathcal{S}} \mu_j + \tau = 1$, we distribute r_j to each term and obtain

$$(1+\alpha)(V^{n+1}(\mathbf{x})+r_j) \geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{j\}} \mu_k(V^n(\mathbf{x}-I_k(\mathbf{x})\mathbf{e}_k)+r_j) + \mu_j(V^n(\mathbf{x}-I_j(\mathbf{x})\mathbf{e}_j)+r_j)$$
$$+\sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_k)+r_j)+(\tau_l-\tau_{lk})(V^n(\mathbf{x})+r_j)\}.$$

We now apply the inductive hypothesis to every term except one, as well as the fact that $I_k(\mathbf{x}+\mathbf{e}_j) = I_k(\mathbf{x})$ for all $k \neq j$, to obtain

$$(1+\alpha)(V^{n+1}(\mathbf{x})+r_j) \geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{j\}} \mu_k(V^n(x+\mathbf{e}_j-I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k)) + \mu_j(V^n(x-I_j(\mathbf{x})\mathbf{e}_j)+r_j)$$
$$+\sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k))+(\tau_l-\tau_{lk})(V^n(\mathbf{x}+\mathbf{e}_j))\}. \tag{EC.6}$$

Now, for the third term on the right-hand side of (EC.6), we apply the inductive hypothesis *only* in the case where $x_j > 0$, in which case $I_j(\mathbf{x}+\mathbf{e}_j) = I_j(\mathbf{x})$, and we obtain

$$(1+\alpha)(V^{n+1}(\mathbf{x})+r_j) \geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{j\}} \mu_k(V^n(\mathbf{x}+\mathbf{e}_j-I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k))$$
$$+\mu_j(V^n(\mathbf{x}+\mathbf{e}_j-I_j(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_j)+r_j(I_j(\mathbf{x}+\mathbf{e}_j)-I_j(\mathbf{x})))$$
$$+\sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k))+(\tau_l-\tau_{lk})(V^n(\mathbf{x}+\mathbf{e}_j))\}. \tag{EC.7}$$

Note that in the case where $x_j = 0$, (EC.7) is simply a re-writing of (EC.6), since in that case, $V^n(\mathbf{x}+\mathbf{e}_j-I_j(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_j) = V^n(\mathbf{x}-I_j(\mathbf{x})\mathbf{e}_j)$ and $I_j(\mathbf{x}+\mathbf{e}_j)-I_j(\mathbf{x}) = 1$.

Finally, applying the fact that that $r(\mathbf{x}+\mathbf{e}_j) = r(\mathbf{x}) + \mu_j r_j(I_j(\mathbf{x}+\mathbf{e}_j)-I_j(\mathbf{x}))$, we end up with

$$(1+\alpha)(V^{n+1}(\mathbf{x})+r_j) \geq r(\mathbf{x}+\mathbf{e}_j) + \sum_{k\in S} \mu_k V^n(\mathbf{x}+\mathbf{e}_j-\mathbf{e}_k)$$
$$+\sum_{l\in\mathcal{R}} \max_{k\in S}\{\tau_{lk}V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k)+(\tau_l-\tau_{lk})V^n(\mathbf{x}+\mathbf{e}_j)\}$$
$$=(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j).$$

Proof of Proposition 3. The proof is trivial when $i = j$. Assume that $i, j \in \mathcal{S}$ and $i \neq j$. By Lemma EC.1 and our assumption that $V^0(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathcal{Q}$, it is sufficient to show that for any two stations $i, j \in \mathcal{S}$ such that $r_i \geq r_j$, $\mu_i \geq \mu_j$, and $\tau_{ki} = \tau_{kj}$ for all $k \in \mathcal{R}$, if $x_i \leq x_j$, then

$$V^n(\mathbf{x}+\mathbf{e}_i) \geq V^n(\mathbf{x}+\mathbf{e}_j), \tag{EC.8}$$

$$V^n(T_{ij}\mathbf{x}) \geq V^n(\mathbf{x}), \tag{EC.9}$$

for all $n \geq 0$, where T_{ij} is the linear transformation that swaps the ith element of x with the jth element of x.

We prove (EC.8) and (EC.9) by induction on n. The base case requires that the inequalities hold for $V^0(\cdot)$, which is trivial because $V^0(\cdot)$ is always zero. Suppose that (EC.8) and (EC.9) hold for for some arbitrary $n \geq 0$. We will proceed by showing that when $x_i \leq x_j$, the same inequalities hold for $V^{n+1}(\cdot)$.

Inductive step for (EC.8). Case 1: $0 < x_i < x_j$.

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) = r(\mathbf{x}+\mathbf{e}_j) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k) + \tau V^n(\mathbf{x}+\mathbf{e}_j) \\ + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k) - V^n(\mathbf{x}+\mathbf{e}_j))\}.\end{aligned}$$

Rearranging terms, we obtain

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) = r(\mathbf{x}+\mathbf{e}_j) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k) \\ + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk} V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k) + (\tau_l - \tau_{lk})V^n(\mathbf{x}+\mathbf{e}_j)\}.\end{aligned} \tag{EC.10}$$

We next use the fact that $r(\mathbf{x}+\mathbf{e}_i) = r(\mathbf{x}+\mathbf{e}_j)$ (because $x_i > 0$ and $x_j > 0$) and then apply (EC.8) to the second and third terms on the right-hand side of (EC.10). For the ith and jth summand within the second term, we are using the fact that $x_i < x_j$ to apply (EC.8) for the states $\mathbf{x}-\mathbf{e}_i$ and $\mathbf{x}-\mathbf{e}_j$, respectively. For the second term, we also use the fact that $0 < x_i < x_j$ to conclude that $I_k(\mathbf{x}+\mathbf{e}_i) = I_k(\mathbf{x}+\mathbf{e}_j)$ for all $k\in\mathcal{S}$. For the third term, we apply the inductive hypothesis to every term inside the max operator, further using the facts that $x_i < x_j$ and $\tau_l \geq \tau_{lk}$ for all $l\in\mathcal{R}$ and $k\in\mathcal{S}$, to apply the inductive hypothesis in state $\mathbf{x}+\mathbf{e}_i$. We conclude that

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) \leq r(\mathbf{x}+\mathbf{e}_i) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}+\mathbf{e}_i - I_k(\mathbf{x}+\mathbf{e}_i)\mathbf{e}_k) \\ + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk} V^n(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_k) + (\tau_l - \tau_{lk})V^n(\mathbf{x}+\mathbf{e}_i)\}.\end{aligned}$$

Finally, rearranging the terms on the right-hand side of the above inequality results in

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) &\leq r(\mathbf{x}+\mathbf{e}_i) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}+\mathbf{e}_i - I_k(\mathbf{x}+\mathbf{e}_i)\mathbf{e}_k) + \tau V^n(\mathbf{x}+\mathbf{e}_i) \\ &\quad + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_k) - V^n(\mathbf{x}+\mathbf{e}_i))\} \\ &= (1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_i).\end{aligned}$$

Case 2: $0 = x_i < x_j$. This case is similar to the previous one except that we must treat the μ_i term separately.

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) = r(\mathbf{x}+\mathbf{e}_j) + \mu_i V^n(\mathbf{x}+\mathbf{e}_j) + \sum_{k\in\mathcal{S}\setminus\{i\}} \mu_k V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k) \\ + \tau V^n(\mathbf{x}+\mathbf{e}_j) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k) - V^n(\mathbf{x}+\mathbf{e}_j))\}.\end{aligned}$$

Because $r_j \le r_i$,

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) \le r(\mathbf{x}+\mathbf{e}_j) &+ \mu_i r_i + \mu_i(V^n(\mathbf{x}+\mathbf{e}_j) - r_j)\\ &+ \sum_{k\in\mathcal{S}\setminus\{i\}} \mu_k V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k)\\ &+ \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k) + (\tau_l - \tau_{lk})V^n(\mathbf{x}+\mathbf{e}_j)\},\end{aligned}$$

and by applying (EC.5), the above inequality implies that

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) \le r(\mathbf{x}+\mathbf{e}_j) + \mu_i r_i + \mu_i V^n(\mathbf{x}) &+ \sum_{k\in\mathcal{S}\setminus\{i\}} \mu_k V^n(\mathbf{x}+\mathbf{e}_j - I_k(\mathbf{x}+\mathbf{e}_j)\mathbf{e}_k)\\ &+ \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}V^n(\mathbf{x}+\mathbf{e}_j+\mathbf{e}_k) + (\tau_l - \tau_{lk})V^n(\mathbf{x}+\mathbf{e}_j)\}.\end{aligned}$$

Now, by noting that $r(\mathbf{x}+\mathbf{e}_i) = \mu_i r_i + r(\mathbf{x}+\mathbf{e}_j)$ (since $x_i = 0$ and $x_j > 0$) and then applying (EC.8) to every V^n term, which can be done because $x_j > x_i$, we have

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) \le r(\mathbf{x}+\mathbf{e}_i) &+ \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}+\mathbf{e}_i - I_k(\mathbf{x}+\mathbf{e}_i)\mathbf{e}_k)\\ &+ \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}V^n(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_k) + (\tau_l - \tau_{lk})V^n(\mathbf{x}+\mathbf{e}_i)\},\end{aligned}$$

which can be rearranged to give

$$\begin{aligned}(1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_j) &\le r(\mathbf{x}+\mathbf{e}_i) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}+\mathbf{e}_i - I_k(\mathbf{x}+\mathbf{e}_i)\mathbf{e}_k) + \tau V^n(\mathbf{x}+\mathbf{e}_i)\\ &\quad + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_k) - V^n(\mathbf{x}+\mathbf{e}_i))\}\\ &= (1+\alpha)V^{n+1}(\mathbf{x}+\mathbf{e}_i).\end{aligned}$$

Case 3: $x_i = x_j$. In this case, (EC.9) implies (EC.8).

Inductive step for (EC.9). Case 1: $x_j > x_i > 0$.

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) = r(T_{ij}\mathbf{x}) &+ \sum_{k\in S} \mu_k V^n(T_{ij}\mathbf{x} - I_k(T_{ij}\mathbf{x})\mathbf{e}_k)\\ &+ \tau V^n(T_{ij}\mathbf{x}) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}(V^n(T_{ij}\mathbf{x}+\mathbf{e}_k) - V^n(T_{ij}\mathbf{x}))\}.\end{aligned} \tag{EC.11}$$

Noting that $x_i > 0$ and $x_j > 0$, we rearrange terms to obtain

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) \ge r(T_{ij}\mathbf{x}) &+ \sum_{k\in\mathcal{S}\setminus\{i,j\}} \mu_k V^n(T_{ij}\mathbf{x} - I_k(T_{ij}\mathbf{x})\mathbf{e}_k) + \mu_i V^n(T_{ij}\mathbf{x} - \mathbf{e}_i)\\ &+ \mu_j V^n(T_{ij}\mathbf{x} - \mathbf{e}_j) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}V^n(T_{ij}\mathbf{x}+\mathbf{e}_k) + (\tau_l - \tau_{lk})V^n(T_{ij}\mathbf{x})\}.\end{aligned}$$

Next we apply the inductive hypothesis on (EC.9) for n. Because $x_i < x_j$ and using the fact that $r(T_{ij}\mathbf{x}) = r(\mathbf{x})$ (because both x_i and x_j are positive),

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) &\geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{i,j\}} \mu_k V^n(\mathbf{x}-I_k(\mathbf{x})\mathbf{e}_k) + \mu_i V^n(\mathbf{x}-\mathbf{e}_j) \\ &\quad + \mu_j V^n(\mathbf{x}-\mathbf{e}_i) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}V^n(T_{ij}\mathbf{x}+\mathbf{e}_k) + (\tau_l-\tau_{lk})V^n(T_{ij}\mathbf{x})\} \\ &= r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{i,j\}} \mu_k V^n(\mathbf{x}-I_k(\mathbf{x})\mathbf{e}_k) + \mu_j V^n(\mathbf{x}-\mathbf{e}_j) + \mu_j V^n(\mathbf{x}-\mathbf{e}_i) \\ &\quad + (\mu_i-\mu_j)V^n(\mathbf{x}-\mathbf{e}_j) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}V^n(T_{ij}\mathbf{x}+\mathbf{e}_k) + (\tau_l-\tau_{lk})V^n(T_{ij}\mathbf{x})\}.\end{aligned}$$

Now, applying the inductive hypothesis (that (EC.8) holds for n for the state $\mathbf{x}-\mathbf{e}_j-\mathbf{e}_i$) and the condition that $\mu_i \geq \mu_j$, we obtain

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) &\geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{i,j\}} \mu_k V^n(\mathbf{x}-I_k(\mathbf{x})\mathbf{e}_k) + \mu_j V^n(\mathbf{x}-\mathbf{e}_j) + \mu_j V^n(\mathbf{x}-\mathbf{e}_i) \\ &\quad + (\mu_i-\mu_j)V^n(\mathbf{x}-\mathbf{e}_i) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}}\{\tau_{lk}V^n(T_{ij}\mathbf{x}+\mathbf{e}_k) + (\tau_l-\tau_{lk})V^n(T_{ij}\mathbf{x})\}.\end{aligned} \tag{EC.12}$$

Finally, we examine the terms inside the max operator. For $k \notin \{i,j\}$,

$$\tau_{lk}V^n(T_{ij}\mathbf{x}+\mathbf{e}_k) + (\tau_l-\tau_{lk})V^n(T_{ij}\mathbf{x}) \geq \tau_{lk}V^n(\mathbf{x}+\mathbf{e}_k) + (\tau_l-\tau_{lk})V^n(\mathbf{x})$$

by directly applying the inductive hypothesis and the fact that $\tau_l \geq \tau_{lk}$ for all $l \in \mathcal{R}, k \in \mathcal{S}$. For $k = i$, we have

$$\begin{aligned}\tau_{li}V^n(T_{ij}\mathbf{x}+\mathbf{e}_i) + (\tau_l-\tau_{li})V^n(T_{ij}\mathbf{x}) &\geq \tau_{li}V^n(\mathbf{x}+\mathbf{e}_j) + (\tau_l-\tau_{li})V^n(\mathbf{x}) \\ &= \tau_{lj}V^n(\mathbf{x}+\mathbf{e}_j) + (\tau_l-\tau_{lj})V^n(\mathbf{x})\end{aligned}$$

and for $k = j$, we have

$$\begin{aligned}\tau_{lj}V^n(T_{ij}\mathbf{x}+\mathbf{e}_j) + (\tau_l-\tau_{lj})V^n(T_{ij}\mathbf{x}) &\geq \tau_{lj}V^n(\mathbf{x}+\mathbf{e}_i) + (\tau_l-\tau_{lj})V^n(\mathbf{x}) \\ &= \tau_{li}V^n(\mathbf{x}+\mathbf{e}_i) + (\tau_l-\tau_{li})V^n(\mathbf{x}),\end{aligned}$$

because we have assumed that $\tau_{lj} = \tau_{li}$. Thus, for every term inside $\max_{k\in\mathcal{S}}\{\tau_{lk}V^n(\mathbf{x}+\mathbf{e}_k) + (\tau_l - \tau_{lk})V^n(\mathbf{x})\}$, there is a corresponding term inside the max operator in (EC.12), which is at least as large. Hence, from (EC.12), we conclude that

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) &\geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x}-I_k(\mathbf{x})\mathbf{e}_k) + \tau V^n(\mathbf{x}) + \sum_{l\in\mathcal{R}} \max_{k\in S}\{\tau_{lk}(V^n(\mathbf{x}+\mathbf{e}_k) - V^n(\mathbf{x}))\} \\ &= (1+\alpha)V^n(\mathbf{x}).\end{aligned}$$

Case 2: $x_j > x_i = 0$. Rearranging terms in (EC.11) and noting that $x_i = 0$ and $x_j > 0$ yields

$$(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) \geq r(T_{ij}\mathbf{x}) + \sum_{k\in\mathcal{S}\setminus\{i,j\}} \mu_k V^n(T_{ij}\mathbf{x} - I_k(T_{ij}\mathbf{x})\mathbf{e}_k) + \mu_i V^n(T_{ij}\mathbf{x} - \mathbf{e}_i) + \mu_j V^n(T_{ij}\mathbf{x})$$
$$+ \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(T_{ij}\mathbf{x} + \mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(T_{ij}\mathbf{x})\}.$$

Now, applying the inductive hypothesis on (EC.9) for n and using the fact that $r(T_{ij}\mathbf{x}) = r(\mathbf{x}) + r_i\mu_i - r_j\mu_j$ since $x_j > x_i = 0$,

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) &\geq r(\mathbf{x}) + r_i\mu_i - r_j\mu_j + \sum_{k\in\mathcal{S}\setminus\{i,j\}} \mu_k V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) + \mu_i V^n(\mathbf{x} - \mathbf{e}_j) \\ &\quad + \mu_j V^n(\mathbf{x}) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(T_{ij}\mathbf{x} + \mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(T_{ij}\mathbf{x})\} \\ &= r(\mathbf{x}) + r_i\mu_i - r_j\mu_j + \sum_{k\in\mathcal{S}} \mu_k V^n(x - I_k(\mathbf{x})\mathbf{e}_k) \\ &\quad + (\mu_i - \mu_j)\left(V^n(\mathbf{x} - \mathbf{e}_j) - V^n(x)\right) \\ &\quad + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(T_{ij}\mathbf{x} + \mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(T_{ij}\mathbf{x})\}.\end{aligned}$$

Since (EC.5) holds and $\mu_i \geq \mu_j$, we have

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) &\geq r(\mathbf{x}) + r_i\mu_i - r_j\mu_j + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) \\ &\quad + (\mu_i - \mu_j)(-r_j) + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(T_{ij}\mathbf{x} + \mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(T_{ij}\mathbf{x})\},\end{aligned}$$

and because $r_i \geq r_j$,

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) &\geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) \\ &\quad + \sum_{l\in\mathcal{R}} \max_{k\in\mathcal{S}} \{\tau_{lk} V^n(T_{ij}\mathbf{x} + \mathbf{e}_k) + (\tau_l - \tau_{lk}) V^n(T_{ij}\mathbf{x})\}.\end{aligned}$$

Finally, by the same argument used in case 1 for the terms inside the max operator, we conclude that

$$\begin{aligned}(1+\alpha)V^{n+1}(T_{ij}\mathbf{x}) &\geq r(\mathbf{x}) + \sum_{k\in\mathcal{S}} \mu_k V^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) + \tau V^n(\mathbf{x}) + \sum_{l\in\mathcal{R}} \max_{k\in S} \{\tau_{lk}(V^n(\mathbf{x} + \mathbf{e}_k) - V^n(\mathbf{x}))\} \\ &= (1+\alpha)V^{n+1}(\mathbf{x}).\end{aligned}$$

Case 3: $x_i = x_j$. In this case, (EC.9) holds with equality because swapping the ith and jth elements of x results in no change to x.

Proof of Proposition 4. By Lemma EC.1 and our assumption that $V^0(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathcal{Q}$, it is sufficient to show that for all states $\mathbf{x} \in \mathcal{Q}$ such that $x_1 = 0$, and for all $j \in \mathcal{S}$ and $n \geq 0$,

$$\tau_{11} M_1^n(\mathbf{x}) \geq \tau_{1j} M_j^n(\mathbf{x}), \tag{EC.13}$$

where $M_j^n(\mathbf{x}) \equiv V^n(\mathbf{x}+\mathbf{e}_j) - V^n(\mathbf{x})$. We show that (EC.13) holds by induction on n. The base case with $n=0$ is trivial because $M_j^0(\mathbf{x})=0$ for all $j \in \mathcal{S}, \mathbf{x} \in \mathcal{Q}$. We now show that if (EC.13) holds for n then the same will hold for $n+1$.

Before beginning, for compactness of notation, note that we can write $(1+\alpha)M_i^{n+1}(\mathbf{x})$ as

$$\begin{aligned} r_i\mu_i(1-I_i(\mathbf{x})) + \mu_i I_i(\mathbf{x}) M_i^n(\mathbf{x}-\mathbf{e}_i) + \sum_{k\in\mathcal{S}\setminus\{i\}} \mu_k M_i^n(\mathbf{x} - I_k(\mathbf{x})\mathbf{e}_k) + \tau M_i^n(\mathbf{x}) \\ + \max_{k\in\mathcal{S}}\{\tau_{lk} M_k^n(\mathbf{x}+\mathbf{e}_i)\} - \max_{k\in\mathcal{S}}\{\tau_{lk} M_k^n(\mathbf{x})\}, \end{aligned}$$

for $i \in \mathcal{S}$.

Case 1: $x_j = 0$, for $j \in \mathcal{S}\setminus\{1\}$.

$$\begin{aligned} &(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x}) - \tau_{1j}M_j^{n+1}(\mathbf{x})\right) = \\ &\tau_{11}\left[r_1\mu_1 + (\mu_j+\tau)M_1^n(\mathbf{x}) + \sum_{i\in\mathcal{S}\setminus\{1,j\}} \mu_i M_1^n(\mathbf{x}-I_i(\mathbf{x})\mathbf{e}_i) + \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+e_1)\} - \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right] \\ &-\tau_{1j}\left[r_j\mu_j + (\mu_1+\tau)M_j^n(\mathbf{x}) + \sum_{i\in\mathcal{S}\setminus\{1,j\}} \mu_i M_j^n(\mathbf{x}-I_i(\mathbf{x})\mathbf{e}_i) + \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+\mathbf{e}_j)\} - \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right]. \end{aligned}$$

Now, using the inductive hypothesis on each of the terms in the summations, we obtain

$$\begin{aligned} &(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x}) - \tau_{1j}M_j^{n+1}(\mathbf{x})\right) \\ &\quad \geq \tau_{11}\left[r_1\mu_1 + (\mu_j+\tau)M_1^n(\mathbf{x}) + \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+e_1)\} - \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right] \\ &\quad -\tau_{1j}\left[r_j\mu_j + (\mu_1+\tau)M_j^n(\mathbf{x}) + \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+\mathbf{e}_j)\} - \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right]. \end{aligned}$$

Again using the inductive hypothesis, we have

$$\begin{aligned} &(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x}) - \tau_{1j}M_j^{n+1}(\mathbf{x})\right) \\ &\quad \geq \tau_{11}\left[r_1\mu_1 + (\mu_j+\tau)M_1^n(\mathbf{x}) + \max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+e_1)\} - \tau_{11}M_1^n(\mathbf{x})\right] \\ &\quad -\tau_{1j}\left[r_j\mu_j + (\mu_1+\tau)M_j^n(\mathbf{x}) + \tau_{11}M_1^n(\mathbf{x}+\mathbf{e}_j) - \tau_{11}M_1^n(\mathbf{x})\}\right] \\ &\quad \geq \tau_{11}\left[r_1\mu_1 + \mu_j M_1^n(\mathbf{x}) + \tau_{1j}M_j^n(\mathbf{x}+e_1) + (\tau-\tau_{11})M_1^n(\mathbf{x})\right] \\ &\quad -\tau_{1j}\left[r_j\mu_j + (\mu_1+\tau)M_j^n(\mathbf{x}) + \tau_{11}M_1^n(\mathbf{x}+\mathbf{e}_j) - \tau_{11}M_1^n(\mathbf{x})\right] \\ &\quad = \tau_{11}\left[r_1\mu_1 + (\mu_j+\tau-\tau_{11})M_1^n(\mathbf{x}) + \tau_{1j}(V^n(\mathbf{x}+e_1+\mathbf{e}_j) - V^n(\mathbf{x}+e_1))\right] \\ &\quad -\tau_{1j}\left[r_j\mu_j + (\mu_1+\tau)(V^n(\mathbf{x}+\mathbf{e}_j) - V^n(\mathbf{x}))\right. \\ &\quad \left. +\tau_{11}(V^n(\mathbf{x}+e_1+\mathbf{e}_j) - V^n(\mathbf{x}+\mathbf{e}_j)) - \tau_{11}(V^n(\mathbf{x}+e_1) - V^n(\mathbf{x}))\right] \\ &\quad = \tau_{11}\left[r_1\mu_1 + \mu_j M_1^n(\mathbf{x}) + (\tau-\tau_{11})M_1^n(\mathbf{x})\right] \\ &\quad -\tau_{1j}\left[r_j\mu_j + (\mu_1+\tau-\tau_{11})(V^n(\mathbf{x}+\mathbf{e}_j) - V^n(\mathbf{x}))\right] \\ &\quad = \tau_{11}r_1\mu_1 - \tau_{1j}r_j\mu_j + \tau_{11}(\mu_j+\tau-\tau_{11})M_1^n(\mathbf{x}) - \tau_{1j}(\mu_1+\tau-\tau_{11})M_j^n(\mathbf{x}). \end{aligned}$$

By using the inductive hypothesis directly, we obtain

$$(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x})-\tau_{1j}M_j^{n+1}(\mathbf{x})\right)\geq\tau_{11}r_1\mu_1-\tau_{1j}r_j\mu_j+(\mu_j-\mu_1)\tau_{1j}M_j^n(\mathbf{x}).\tag{EC.14}$$

Now, if $\mu_j\geq\mu_1$, the right-hand side of (EC.14) is non-negative by (EC.4) and because in the statement of the proposition, $\tau_{11}r_1\mu_1\geq\tau_{1j}r_j\mu_j$ for all $j\in\mathcal{S}$. Otherwise, note that $\tau_{1j}(\mu_j-\mu_1)<0$, and hence using (EC.5) we conclude that

$$\begin{aligned}(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x})-\tau_{1j}M_j^{n+1}(\mathbf{x})\right)\geq&\tau_{11}r_1\mu_1-\tau_{1j}r_j\mu_j+\tau_{1j}(\mu_j-\mu_1)r_j\\=&\tau_{11}r_1\mu_1-\tau_{1j}r_j\mu_1\geq0\end{aligned}$$

by the assumption that $\tau_{11}r_1\geq\tau_{1j}r_j$.

Case 2: $x_j>0$, for $j\in\mathcal{S}\backslash\{1\}$.

$$\begin{aligned}&(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x})-\tau_{1j}M_j^{n+1}(\mathbf{x})\right)=\\&\tau_{11}\left[r_1\mu_1+\mu_jM_1^n(\mathbf{x}-\mathbf{e}_j)+\tau M_1^n(\mathbf{x})+\sum_{i\in\mathcal{S}\backslash\{1,j\}}\mu_iM_1^n(\mathbf{x}-I_i(\mathbf{x})\mathbf{e}_i)+\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+e_1)\}-\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right]\\&-\tau_{1j}\left[(\mu_1+\tau)M_j^n(\mathbf{x})+\sum_{i\in\mathcal{S}\backslash\{1,j\}}\mu_iM_j^n(\mathbf{x}-I_i(\mathbf{x})\mathbf{e}_i)+\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+\mathbf{e}_j)\}-\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right].\end{aligned}$$

Now, using the inductive hypothesis on each of the terms in the summations, we obtain

$$\begin{aligned}&(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x})-\tau_{1j}M_j^{n+1}(\mathbf{x})\right)\\&\quad\geq\tau_{11}\left[r_1\mu_1+\mu_jM_1^n(\mathbf{x}-\mathbf{e}_j)+\tau M_1^n(\mathbf{x})+\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+e_1)\}-\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right]\\&\quad-\tau_{1j}\left[(\mu_1+\tau)M_j^n(\mathbf{x})+\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+\mathbf{e}_j)\}-\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x})\}\right].\end{aligned}$$

Again using the inductive hypothesis, we have

$$\begin{aligned}&(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x})-\tau_{1j}M_j^{n+1}(\mathbf{x})\right)\\&\quad\geq\tau_{11}\left[r_1\mu_1+\mu_jM_1^n(\mathbf{x}-\mathbf{e}_j)+\tau M_1^n(\mathbf{x})+\max_{i\in\mathcal{S}}\{\tau_{1i}M_i^n(\mathbf{x}+e_1)\}-\tau_{11}M_1^n(\mathbf{x})\right]\\&\quad-\tau_{1j}\left[(\mu_1+\tau)M_j^n(\mathbf{x})+\tau_{11}M_1^n(\mathbf{x}+\mathbf{e}_j)-\tau_{11}M_1^n(\mathbf{x})\}\right]\\&\geq\tau_{11}\left[r_1\mu_1+\mu_jM_1^n(\mathbf{x}-\mathbf{e}_j)+\tau M_1^n(\mathbf{x})+\tau_{1j}M_j^n(\mathbf{x}+e_1)-\tau_{11}M_1^n(\mathbf{x})\right]\\&\quad-\tau_{1j}\left[(\mu_1+\tau)M_j^n(\mathbf{x})+\tau_{11}M_1^n(\mathbf{x}+\mathbf{e}_j)-\tau_{11}M_1^n(\mathbf{x})\right]\\&=\tau_{11}\left[r_1\mu_1+\mu_jM_1^n(\mathbf{x}-\mathbf{e}_j)+\tau_{1j}M_j^n(\mathbf{x}+e_1)+(\tau-\tau_{11})M_1^n(\mathbf{x})\right]\\&\quad-\tau_{1j}\left[(\mu_1+\tau)M_j^n(\mathbf{x})+\tau_{11}M_1^n(\mathbf{x}+\mathbf{e}_j)-\tau_{11}M_1^n(\mathbf{x})\right]\\&=\tau_{11}\left[r_1\mu_1+\mu_jM_1^n(\mathbf{x}-\mathbf{e}_j)+\tau_{1j}(V^n(\mathbf{x}+e_1+\mathbf{e}_j)-V^n(\mathbf{x}+e_1))+(\tau-\tau_{11})M_1^n(\mathbf{x})\right]\end{aligned}$$

$$
\begin{aligned}
&-\tau_{1j}\left[(\mu_1+\tau)(V^n(\mathbf{x}+\mathbf{e}_j)-V^n(\mathbf{x}))+\tau_{11}(V^n(\mathbf{x}+e_1+\mathbf{e}_j)-V^n(\mathbf{x}+\mathbf{e}_j))\right.\\
&\left.-\tau_{11}(V^n(\mathbf{x}+e_1)-V^n(\mathbf{x}))\right]\\
=&\tau_{11}\left[r_1\mu_1+\mu_j M_1^n(\mathbf{x}-\mathbf{e}_j)+(\tau-\tau_{11})M_1^n(\mathbf{x})\right]\\
&-\tau_{1j}(\mu_1+\tau-\tau_{11})M_j^n(\mathbf{x})\\
=&\tau_{11}r_1\mu_1+\tau_{11}\mu_j M_1^n(\mathbf{x}-\mathbf{e}_j)-\tau_{1j}\mu_1 M_j^n(\mathbf{x})+(\tau-\tau_{11})(\tau_{11}M_1^n(\mathbf{x})-\tau_{1j}M_j^n(\mathbf{x}))\\
\geq&\tau_{11}r_1\mu_1+\tau_{11}\mu_j M_1^n(\mathbf{x}-\mathbf{e}_j)-\tau_{1j}\mu_1 M_j^n(\mathbf{x}),
\end{aligned}
$$

by using the inductive argument and the fact that $\tau \geq \tau_{11}$. By applying (EC.5), we obtain

$$(1+\alpha)\left(\tau_{11}M_1^{n+1}(\mathbf{x})-\tau_{1j}M_j^{n+1}(\mathbf{x})\right)\geq \tau_{11}r_1\mu_1+\tau_{11}\mu_j M_1^n(\mathbf{x}-\mathbf{e}_j)-\tau_{1j}r_j\mu_1,$$

which is non-negative by (EC.4) and the condition that $\tau_{11}r_1 \geq \tau_{1j}r_j$.

Proof of Proposition 5. Since the decisions in the static policy are random, the expected total discounted reward is separable by station. That is, we can write $V_\gamma(\mathbf{x})=\sum_{j\in S}W_j^\gamma(x_j)$, where $W_j^\gamma(x_j)$ is the expected total discounted reward for a specific station $j\in\mathcal{S}$ with arrivals at rate λ_j and departures at rate μ_j, when there are x_j casualties waiting. For ease of exposition, we suppress the subscript j (corresponding to the station) and the superscript γ (corresponding to the static policy) everywhere they appear in this proof. By uniformizing the process corresponding to the station with uniformization constant $\lambda+\mu$, we can study the embedded discrete-time Markov chain by observing the queue at the station only at transitions, i.e., arrivals and service completions. Arrivals occur with probability $\lambda/(\lambda+\mu)$ and service completions occur with probability $\mu/(\lambda+\mu)$. By incorporating the discount factor α, we can define $W(x)$ according to the following recursion:

$$W(x)=\frac{\lambda+\mu}{\lambda+\mu+\alpha}\left(\frac{\lambda}{\lambda+\mu}W(x+1)+\frac{\mu}{\lambda+\mu}(r+W(x-1))\right),$$

for $x\geq 1$, or in other words,

$$\lambda W(x)-(\lambda+\mu+\alpha)W(x-1)+\mu W(x-2)=-\mu r, \tag{EC.15}$$

for $x\geq 2$. The boundary condition is

$$\lambda W(1)-(\lambda+\alpha)W(0)=0, \tag{EC.16}$$

because there are no service completions when there are zero casualties at the station.

By using the standard method for solving a nonhomogeneous difference equation with a boundary condition, we obtained

$$W(x)=\frac{\mu r}{\alpha}-\frac{2\mu r}{\lambda+\alpha-\mu+\sqrt{(\lambda+\mu+\alpha)^2-4\lambda\mu}}\left(\frac{\lambda+\mu+\alpha-\sqrt{(\lambda+\mu+\alpha)^2-4\lambda\mu}}{2\lambda}\right)^x. \tag{EC.17}$$

By summing the above expression over all stations, we obtain (9). Using algebraic manipulation it is straightforward to show that if $W(x)$ is given by the expression in (EC.17), then it satisfies (EC.15) and (EC.16).

PROPOSITION EC.1. *Algorithm 1 returns a feasible solution to* (F)*.*

Proof. We will prove the proposition by showing that at the end of each iteration through the second **for** loop, constraints (11), (12), and (13) are satisfied. When the **for** loop is initialized for $k=1$, $\rho_{ij}=0$ for all $i\in\mathcal{R}, j\in\mathcal{S}$ and thus all three constraints are clearly satisfied at the beginning of the first iteration.

Now, we show that if the three constraints are satisfied at the beginning of any iteration of the **for** loop, then they will be satisfied at the end of the same iteration. During the iteration, only one ρ_{ij} can be changed, namely the one in the kth position of the list. Since (11) and (12) are satisfied at the beginning of the iteration, both $1-\sum_{l\in\mathcal{S}}\rho_{il}$ and $\mu_j/\tau_{ij}-\sum_{k\in\mathcal{R}}\tau_{kj}\rho_{kj}/\tau_{ij}$ are non-negative, and hence ρ_{ij} will also be non-negative at the end of the iteration. Moreover, because the new value assigned to ρ_{ij} is at most $1-\sum_{l\in\mathcal{S}}\rho_{il}$, we maintain $\sum_{l\in\mathcal{S}}\rho_{il}\le 1$ and (12) is satisfied. An identical argument can be made for (11).

Since all three constraints are satisfied initially, we have shown all three constraints will be satisfied after the first iteration, and therefore after every subsequent iteration. Thus, the values $\{\rho_{ij}\}$ returned by Algorithm 1 are feasible to (F).

Algorithm 1 Greedy algorithm for obtaining a feasible solution to (F).

function GREEDY($\{\tau_{ij}\},\{\mu_j\},\{r_j\},\alpha$)
    for all $i\in\mathcal{R}, j\in\mathcal{S}$ **do** $\rho_{ij}\leftarrow 0$
    end for
    $list\leftarrow\{(i,j), i\in\mathcal{R}, j\in\mathcal{S}\}$
    SortDescending($list, \tau_{ij}r_j\exp(-\alpha/\tau_{ij})$)
    for $k=1$ to Length($list$) **do**
        $(i,j)\leftarrow list[k]$
        $\rho_{ij}\leftarrow\min\left\{1-\sum_{l\in\mathcal{S}}\rho_{il},\, \mu_j/\tau_{ij}-\sum_{k\in\mathcal{R}}\tau_{kj}\rho_{kj}/\tau_{ij}\right\}$
    end for
    return $\{\rho_{ij}\}$
end function

Table EC.1 Values of coefficients a_{kj} and c_{ijk} in Equation (8).

a_{kj}	$\dfrac{\log\left(\frac{\mu_k}{\mu_k+\alpha}\right)}{\log\left(\frac{\mu_j}{\mu_j+\alpha}\right)}$
c_{ijk}	$\dfrac{\log\left(\frac{\tau_{ik} r_k \mu_k (\mu_j+\alpha)}{\tau_{ij} r_j \mu_j (\mu_k+\alpha)}\right)}{\log\left(\frac{\mu_j}{\mu_j+\alpha}\right)}$

Table EC.2 Specific parameter values for Equation (14) under myopic and PIH policies.

	b_1	b_2	t
Myopic	$e^{s\alpha}\left(\dfrac{\mu_j}{\mu_j+\alpha}\right)^{x_j+y_j+1}$	$\dfrac{\mu_j}{\mu_j+\alpha}$	$s(\mu_j+\alpha)$
PIH	$\dfrac{\mu_j \exp\left(\frac{s\mu_j(\lambda_j+\eta_j-\mu_j-\alpha)}{\mu_j+\lambda_j+\alpha-\eta_j}\right)\left(\frac{\mu_j+\lambda_j+\alpha-\eta_j}{2\lambda_j}\right)^{x_j+y_j}}{\lambda_j\left(\frac{\mu_j-\lambda_j-\alpha-\eta_j}{\mu_j-\lambda_j+\alpha-\eta_j}\right)}$	$\left(\dfrac{\mu_j}{\lambda_j}\right)\left(\dfrac{\mu_j-\lambda_j+\alpha-\eta_j}{\mu_j-\lambda_j-\alpha-\eta_j}\right)$	$\dfrac{2s\lambda_j\mu_j}{\mu_j+\lambda_j+\alpha-\eta_j}$