Submitted to
manuscript (Please, provide the manuscript number!)

Designing Response Supply Chain Against Bioattacks

David Simchi-Levi
Department of Civil and Environmental Engineering, Institute for Data, Systems, and Society, and Operations Research Center, MIT, Cambridge, MA, dslevi@mit.edu

Nikolaos Trichakis
Operations Research Center and Sloan School of Management, MIT, Cambridge, MA, ntrichakis@mit.edu

Peter Yun Zhang
Institute for Data, Systems, and Society, MIT, Cambridge, MA, pyzhang@mit.edu

We study a prescriptive model for end-to-end design of a supply chain for medical countermeasures (MCM) to defend against bioattacks. We model the defender's MCM inventory prepositioning and dispensing capacity installation decisions, attacker's move, and defender's adjustable shipment decisions, so as to minimize inventory and life loss costs, subject to population survivability targets. We explicitly account for the strategic interaction between defender's and attacker's actions, assuming information transparency. We consider the Affinely Adjustable Robust Counterpart (AARC) to our problem, which enables us to deal with realistic networks comprising millions of nodes. We provide theoretical backing to the AARC performance by proving its optimality under certain conditions. We conduct a high-fidelity case study on the design of a MCM supply chain with millions of nodes to guard against anthrax attacks in the United States. We calibrate our model using data from a wide variety of sources, including literature and field experiments. We produce policy insights that have been long sought after but elusive up until now.

Key words: bioterrorism; supply chain design; robust optimization
History:

1. Introduction

In this paper, we study the problem of designing a medical countermeasures (MCM) supply chain in preparation against bioattacks, *i.e.*, the intentional release of pathogens or biotoxins against humans to cause serious illness and death. Bioattacks have been a rising first-order concern to many countries worldwide in the last 15 years; the U.S. alone, for instance, has poured approximately $60 billion into biodefense preparedness since the 9/11 attacks [33]. Bioattacks are considered a major threat because a minute quantity of pathogens is sufficient to infect humans; furthermore, the appropriate MCMs need to be administered within a short time window to effectively reduce casualty. For example, if bioterrorists were to release *Bacillus anthracis*, *i.e.*, anthrax, over a large city, hundreds of thousands of people could be at risk, and MCMs would need to be administered to them within a few days to have the intended effect [43, 52]. Distributing such large MCM quantities

to the general public within such short time periods poses a considerable operational challenge, which is often compounded by attack detection delays.

In order to be prepared to deliver adequate medication in a timely manner in response to a bioattack, first, MCMs need to be stockpiled at appropriate locations. Some inventories can be stored in central locations to take advantage of deployment flexibility, pooling effects, and scale economies in holding costs, while others can be positioned closer to populous areas, or even predispensed to the intended users in the form of home medical kits, in order to reduce transportation time. Second, for end users to have prompt access to deployed MCMs, target areas need to have adequate dispensing capacity in place, *e.g.*, in the form of publicly accessible dispensing facilities or dedicated courier delivery services. Too little capacity would result in inefficient usage of available inventory and devastating delays; too much would result in wasted resources. Given candidate storage locations, where and how much MCM inventory should then be prepositioned in order to cost-effectively defend against bioattacks? How much dispensing capacity should be installed at target areas? Our goal in this paper is to address these research questions and develop a decision support framework that could guide policy design.

MCM supply chain in the U.S. The Centers for Disease Control and Prevention (CDC) maintains MCM inventories within the Strategic National Stockpile (SNS), which is a critical component of the national health security programs managed by the CDC Office of Public Health Preparedness and Response (PHPR) [45, 46]. The inventory storage and delivery infrastructure of SNS form a crucial line of defense for the nation's public health security.

For many bioattack threats, SNS currently stores MCM inventories in central repositories, while its delivery infrastructure enables deployment of virtually any amount of stored MCMs to almost any major city in the U.S. within 12 to 36 hours, at negligible shipping costs. But centrally stored inventories may not be responsive enough for events such as anthrax attacks. In recent years, CDC has been increasingly aware of alternatives, namely the prepositioning of MCMs to locations closer to the population. Prepositioning has been raised as a first-order issue for SNS, and PHPR policymakers have commissioned several studies to examine different strategies [43, 52]. To date, several facets of the problem have been analyzed in isolation, including location and transportation options, cost of each system component, and health effects of delayed treatment.

However, a comprehensive cost-benefit analysis at the *system level* that *integrates* individual considerations into a practical quantitative decision support tool has been elusive. Our work helps fill in this gap, the importance of which is indeed highlighted in the following excerpt from a report prepared by the Board of Scientific Counselors for PHPR [32, 44]:

> "Both analytic and simulation (experimental) modeling activities should be increased substantially. Modeling will allow [the Division of] SNS to make quantitatively-based decisions on

how much inventory to hold and where to hold it. An end-to-end model capturing the flow of materials in the SNS, as well as costs and logistical and health measures, should begin at the SNS-managed inventory site and go all the way to the point of dispensing to the public. Using such models will reveal bottlenecks, provide cost estimates, and help SNS properly evaluate the costs and consequences of alternative Response Supply Chain configurations."

Herein, we focus on an end-to-end analytic model as described in the quote above. The absence of such integrative models so far attests to the underlying technical challenges: (a) the multiplicity of decisions involved, both strategic and operational; (b) the subtle way prepositioning could influence bioattackers' actions; and (c) the scale of the problem. In particular, on the strategic level, policymakers need to decide how much inventory to store and how much capacity to install among thousands of locations, in anticipation of an attack. On the operational level, in response to an attack, appropriate shipment decisions need to be made to efficiently dispense MCMs based on the specific demand and supply conditions. The *static* inventory positioning and capacity installation decisions need to be made in conjunction with myriads of *contingent* shipment policies. Furthermore, some information about planned responses is already available in the public domain, and prepositioning MCMs, particularly predispensing medical kits, is likely to grant to bioattackers even greater visibility to planned responses. Consequently, bioattackers, because they usually act in a pre-planned way to inflict the greatest damage possible, are more likely to target under-served populated areas, as opposed to ones that have access to abundant stockpiles. That attackers' actions are likely to be adversarial is well recognized by biodefense policymakers, see, *e.g.*, [52], and the intelligence and military operations research literature [6, 19, 26, 27, 34].

Our methodology. We model the SNS as a network: inventory stockpiles as nodes, and shipment routes as edges. Bioattacks correspond to demand surges at some of the nodes. To model interactions between bioattackers' and policymakers' decisions, we adopt the defender-attacker-defender model (DAD), which can be viewed as a *multistage* robust optimization (RO) model [11, 19]. Specifically, we consider MCM prepositioning as a first-stage static defender decision, bioattacks as second-stage adjustable attacker decisions, and MCM shipments as third-stage adjustable defender decisions. We calibrate the objective and uncertainty set of possible demand scenarios in ways that reflect policymakers' considerations. In particular, we deal with minimizing either inventory holding costs under survivability guarantees for all possible attack scenarios, or worst-case inventory and health costs, where each scenario comprises simultaneous attacks in multiple locations. To preclude overly pessimistic scenarios in which too many locations are attacked, our uncertainty sets include attack budget constraints, which are both standard and crucial in the DAD framework [4].

Unfortunately, there are no scalable solution approaches that perform provably well even for the inventory prepositioning problem, let alone joint optimization of inventory and dispensing capacity

decisions. In particular, DAD models and multistage RO problems of the type we study here are generally intractable [4, 11, 19, 20]. There has been valuable effort in solving reasonably sized problems [5], but they are orders of magnitude smaller than what our problem entails. In the RO literature, a popular heuristic is the so-called Affinely Adjustable Robust Counterpart (AARC), which restricts adjustable decisions to be affine in the uncertainty variables [11]. Although AARCs often lead to tractable formulations, they can perform rather poorly in high-dimensional problems [13]. Worse, all known performance guarantees for AARCs apply to problems with "symmetric" uncertainty sets. The aforementioned attack budget constraints that are necessary in our problem lead to sets that are the "least-symmetric," casting considerable doubt over the performance of AARCs in our setting (see our Literature Review and discussion in §3.2 for more details).

Our contributions. First, for the inventory prepositioning problem, we provide an approximate solution approach, with theoretical backing for its performance, that can deal with network sizes on the order of millions of nodes. In particular, we consider the Affinely Adjustable Robust Counterpart, which reduces to a linear optimization problem (LP) by restricting the adjustable shipment policies to be affine in the demand shocks. We provide theoretical backing to its performance by proving that AARC is in fact optimal under certain conditions. Our work contributes to the RO literature by proving optimality of AARC in a new and important context—see our discussion in Literature Review and in §4. Numerical studies we conducted revealed that AARCs provide near-optimal performance for the inventory prepositioning problem under general settings as well.

Second, we consider the joint problem of inventory prepositioning and dispensing capacities optimization. We show how dispensing capacity, which is costly, can be captured as a first-stage decision and how the third-stage shipment decisions can be reformulated so as to reflect the underlying capacitated dispensing operations. We derive the AARC for this joint problem as an LP and conduct numerical studies to confirm that it still produces near-optimal solutions.

Third, successfully applying AARC to our problem speaks to the potential of its broader impact for other DAD problems. In particular, the AARC approach is applicable to tri-level programs that are nonlinear and with integer variables. In the sense of [4], AARC is suitable for eight out of the nine types of operator models (the only exception is simulation-based operator model that does not have a closed form description at all).

Fourth, we use our model in a thorough, large-scale case study of MCM supply chain design for the SNS to defend against anthrax attacks. Tractability of our approach enables us to perform the first study we are aware of in the literature that is at a realistic, nationwide scale and deals with networks with millions of nodes. We calibrate our model using multiple literature sources that studied different facets of SNS design in isolation. Our integrated framework enables us to derive prescriptive recommendations to the CDC. For example, we find that if CDC wanted to ensure

85% survivability for attacks that simultaneously affected (any) two states, each with at most two cities being attacked, with a detection time of under 48 hours, then the minimum required annual inventory and dispensing capacity budget would be about $210 million; for survivability target of 92%, the budget would increase to $553 million. This increasing marginal cost phenomenon, along with other cost/policy implications, can be explored quantitatively within our model.

1.1. Literature Review

Resource allocation for military applications. For many military defense problems, such as ours, stochastic and priority-list approaches are inappropriate for capturing the strategic interaction between intelligent defender and attacker [3, 26, 42]. Therefore, it is desirable to use a sequential game approach, such as DAD. A main challenge for this approach is tractability. In the past decade, there has been concentrated and valuable effort in developing scalable solution approaches, especially for the interdiction type of problems [3, 5, 10, 20, 34, 37]. Atkinson [10] and Lazzaro [37] show that the state-of-the-art algorithms can solve networks with thousands of nodes. Alderson *et al.* [5] mentions that for the significantly simpler bilevel games, problems with thousands of nodes can be readily solved by decomposition algorithms with commercial solvers.

We differ from the existing DAD papers in two ways. First, we deal with problem sizes on the order of *millions* of nodes and present for them scalable approximate solution approaches backed by performance guarantees. Such drastic increase in problem size would render existing decomposition-based solution approaches insufficient. To deal with high dimensionality, we use the AARC heuristic from RO and contribute back to the theoretical RO literature by proving a novel performance guarantee of AARC (see related discussion below). Second, we provide an integrated modeling of cost, health deprivation, and antibiotic efficacy over time for our application. This in part requires a precise analysis of time-expanded multi-commodity network flow, and how the AARC heuristic is analytically suited for it. We consider these insights to be generalizable since multi-commodity network flow models are considered essential primitives for DAD problems [4].

Another set of military operations research papers [12, 29, 47, 60] also study Stackelberg game problems. For example, Berman and Gavious [12] study the defender-attacker problem as a min-max facility location problem. Zhuang and Bier [60] study the difference between random and fully strategic attackers. While they focus on equilibrium analysis, our paper formulates and solves a large-scale, high-fidelity prescriptive model to support decisions.

Biodefense. Existing biodefense literature does not focus on the interaction between defender and attacker. It generally assumes an abundance of MCM inventories and exogenous demand scenarios [16, 24, 30, 35, 36, 38, 39, 40, 41, 56]. They study, for example, the optimal number and layout of points of dispense [30, 38, 39, 40, 41], the effectiveness of predetermined, exogenous inventory

levels on the treatment of patients after bioattacks and influenza outbreaks [16, 35, 36], and a small number of response strategies for airborne anthrax attacks [24, 56]. Taking MCM supply for granted and assuming attack scenarios to be known is limiting, because in practice, MCMs are shipped from different locations with different lead times (from hours to days) and shipment sizes, resulting in time-dependent flows, and terrorists' decisions are unknown in advance. In contrast, we take an integrated approach and study the entire supply chain within the DAD framework.

Inventory prepositioning. Beyond the biodefense literature, other streams have studied prepositioning problems, including humanitarian logistics [21, 22, 25] and manufacturing [48, 49]. The latter two papers provide a decision support tool to evaluate performance under disruptive risks for a given inventory configuration and a given disruption scenario. Unlike the *descriptive* nature of these models, ours derives *prescriptive* suggestions that incorporate two stages of decision making.

Recent works by Uichanco [53] and Simchi-Levi *et al.* [50] study inventory prepositioning to hedge against capacity disruption and demand uncertainties. Uichanco provides a solution method that scales with the number of vertices in the polytope describing the network structure. Such a method would lead to an intractable formulation in our case because a national biodefense network typically comprises millions of nodes and edges. In contrast, we provide a tractable model that can deal with problems of the size faced by U.S. policymakers. By incorporating dispensing capacity optimization, our model is more general too. In [50], the authors also focus on providing an exact solution method. Not surprisingly, this leads to an intractable formulation too, namely an LP with an exponential number of constraints. They propose a constraint generation algorithm, which they find to work well in practice for bi-partite graphs involving 200 nodes. Their constraint generation algorithm still involves the solution of mixed-integer linear programs in its subroutines. Such approaches could not possibly be employed to tackle networks with millions of nodes, like ours. In contrast, our focus is to provide tractable and scalable (polynomial-time) algorithms, which require the solution only of linear optimization problems. Although not being exact for general network topologies, we back our approach with optimality guarantees as discussed above.

Robust network flow. Atamtürk and Zhang [9] study the computational complexity of two-stage robust network flow under budgeted uncertainty sets and find it to be intractable in general. For a tree structure, they use a dynamic programming (DP) approach to show that it can be solved in polynomial time. Unfortunately, their DP approach is not suitable for our problem. Being developed within a constructive proof to show that robust tree network flow problems are polynomially solvable, the DP approach breaks down and cannot be applied for general, non-tree network flow problems. Furthermore, it is also tailored to deal only with uncapacitated network flow problems and it is unclear if it can be applied to tackle dispensing capacity issues. From a practical standpoint, because the SNS network need not be a tree, and because dispensing capacity is a major concern

[52], the DP approach by Atamtürk and Zhang cannot be applied to tackle the SNS design problem we study in our work. They also provided a cutting plane algorithm that can be applied to solve general (capacitated) networks, but this exact solution algorithm has no polynomial time guarantee and is not practical for large instances, *e.g.*, the ones we solved in Section 6. Our goal is different. Instead of focusing on computational complexity, we study a tractable solution method for general networks involving a rich set of decisions (inventory, capacity, shipment, and dispensing over time), and provide strong theoretical and numerical backing for its performance. Our solution method enables us to apply it to a practically relevant policy design problem, demonstrated in a case study involving millions of nodes and considering important practical decisions and constraints.

Affinely Adjustable Robust Optimization. An introduction to this topic is included in Ben-tal *et al.* [11]. There have only been a handful of results that show AARC heuristics to be optimal under special circumstances [7, 14, 15, 31]. Our work contributes to this stream by providing an optimality proof under a new context. Bertsimas *et al.* [15] studies one-dimensional problems, *e.g.*, inventory problems that involve a single stocking level, while in Delage and Ardestani-Jaafari's work [7], the authors deal with newsvendor type problems under demand uncertainty. We differ from these papers by studying multidimensional problems, *i.e.*, shipment decisions in a supply chain network. The papers by Iancu *et al.* [31] and Bertsimas *et al.* [14] deal with families of problems, where for AARC optimality (or "good" performance) they require the uncertainty sets to be lattices, or possess a "symmetric" structure. In our model, in order to preclude overly conservative solutions, we rely on simplex-type budget constraints that invalidate both the lattice and the symmetric structures. Put differently, we prove optimality for a problem that is according to [14] in some sense the "least conducive" to the AARC optimality conditions set forth in the literature so far.

2. The Inventory Prepositioning Problem

We begin by considering the MCM inventory prepositioning problem, assuming for now that sufficient dispensing capacity is installed. For exposition purposes, we use SNS to exemplify our discussion, but our model and solution heuristic generalize to any network. The SNS network is designed to protect the public susceptible to bioattacks in a set of geographic locations, *e.g.*, densely populated towns or neighborhoods, which we shall refer to as *demand* locations. The demand locations are split into administrative divisions according to a hierarchical structure, *e.g.*, boroughs, municipalities, provinces, states, etc. There are $L+1$ levels in the hierarchy, indexed by $\ell = 0, \ldots, L$, whereby each division at the lowest (Lth) level comprises precisely one demand location. Divisions at some intermediate ℓth level comprise subsets of divisions at the lower $(\ell+1)$th level in a nested fashion; the highest (0th) level includes a single division, *e.g.*, federal/national level. For a division i at the ℓth level, let $\mathcal{P}(i)$ be its parent division at the $(\ell-1)$th level, $\mathcal{C}(i)$ the set of (children)

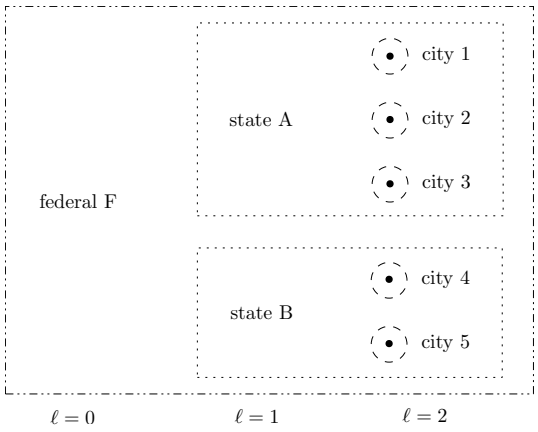

Figure 1 **Five demand locations (cities) split into three administrative divisions, federal $\ell = 0$, state $\ell = 1$ and city $\ell = 2$. Note that $\mathcal{P}(1) = \mathcal{P}(2) = \mathcal{P}(3) = \mathrm{A}$ and $\mathcal{P}(4) = \mathcal{P}(5) = \mathrm{B}$; $\mathcal{C}(\mathrm{A}) = \mathcal{D}(\mathrm{A}) = \{1, 2, 3\}$ and $\mathcal{C}(\mathrm{B}) = \mathcal{D}(\mathrm{B}) = \{4, 5\}$; $\mathcal{D}(\mathrm{F}) = \{1, 2, 3, 4, 5\}$.**

divisions that it includes at the $(\ell+1)$th level, and $\mathcal{D}(i)$ the set of demand locations that it includes, *i.e.*,

$$\mathcal{D}(i) = \{\text{demand location } j : \underbrace{\mathcal{P}(\mathcal{P}(\ldots\mathcal{P}(\mathcal{P}}_{L-\ell \text{ times}}(j))\ldots)) = i\}.$$

Figure 1 includes an illustrative example, where the demand locations are cities, split under state and federal divisions.

To serve the demand locations, each administrative division at the $0, \ldots, L-1$ levels maintains a *stockpile* of appropriate medical countermeasures. In particular, in case of a bioattack, MCMs would typically need to be shipped from the—usually remotely located—stockpiles to the demand locations, where they are subsequently dispensed to the affected population.

Let the demand locations and stockpiles correspond to nodes in a directed graph (V, E), which we refer to as demand and stockpile nodes, respectively. The subset of demand nodes is denoted with V_D. We index the root node (stockpile node at the highest level division) with 0. The edges in E connect stockpile nodes with demand nodes. Specifically, $(i, j) \in E$ if and only if inventory can be shipped from i to $j \in V_D$ once an attack occurs. Shipments have negligible costs and are not subject to capacity constraints for the purposes of the SNS network.[1] However, there is a fixed lead time τ_{ij} for shipping inventory from i to j for all $(i, j) \in E$. Note that even though a "hierarchical inventory network" like the SNS resembles a tree structure, and the demand and stockpile node sets are usually disjoint, our model and solution approach are more general and can be applied to any network (V, E), accommodating demand and stockpile on any node.

We focus on a particular bioattack threat, in anticipation of which a single MCM type is stored. Inventory of this MCM can be prepositioned at any node $i \in V$. If i is a demand node, the stored

[1] Unlike shipping MCMs, dispensing is often capacitated—a situation we address in §5.

inventory corresponds to predispensed medical kits. Let $\mathbf{x} \in \mathbb{R}^{|V|}$ denote the amounts of stored inventories allocated at the nodes. The set of all feasible inventory allocations, $X \subseteq \mathbb{R}^{|V|}$, is assumed to be a polytope and could include inventory related constraints, *e.g.*, non-negativity constraints, budget constraints subject to inventory setup cost and variable cost. A unit of inventory is adequate to treat precisely one individual. Associated with storing inventory at any node are purchasing, replenishment, maintenance costs, etc. We refer to these costs simply as *holding* costs, and denote the per-unit costs with $\mathbf{h} \in \mathbb{R}^{|V|}$. That is, storing x_i MCM units at node $i \in V$ costs $h_i x_i$.

When an attack takes place, a subset of demand nodes are affected, and parts of their populations are in need of treatment with the MCMs. We also say that division i is affected, if any of its children demand nodes $\mathcal{D}(i)$ is affected. For example, if city 1 is attacked in Figure 1, we say that it is affected and so is state A. Let $\mathbf{d} \in \mathbb{R}^{|V_D|}$ be the (a priori unknown) vector with the realized number of affected individuals in each of the demand nodes. Once $\mathbf{d}$ is revealed, inventory stored at the affected demand nodes is available immediately. In case of shortages, inventory from other stockpile nodes can be shipped to satisfy the demands, as long as there exist edges between them. For each edge $(i,j) \in E$, let f_{ij} denote the amount of inventory shipped from i to satisfy demand at node j. Note that shipment decisions are contingent on the realized demand, which we sometimes make explicit by writing them as $f_{ij}(\mathbf{d})$. Inventory shipped from i would be made available for treatment at node j only after τ_{ij} time units. Delays in treatment of affected population could lead to lower probability of survival. We use $\rho_{ij} \in [0,1]$ to denote the *survival probability* (or survivability) of an individual in node j if treated with inventory shipped from node i, and $\overline{\rho} \in [0,1]$ if left untreated. For example, suppose that node j is affected and 50% of the affected individuals are treated with inventory prepositioned at that node, 30% are treated with inventory shipped from node i, and 20% are left untreated. The *average* survivability would then be $0.5\rho_{jj} + 0.3\rho_{ij} + 0.2\overline{\rho}$.

With respect to the possible attack scenarios, CDC considers ones involving simultaneous attacks to multiple locations across the country [32]. In view of this, we consider attack scenarios (or demand vectors) where, for each administrative division i, at most Γ_i of its children divisions $\mathcal{C}(i)$ are affected. We assume Γ_i's to be non-negative integers and refer to them as the *attack scale* parameters. For example, if the administration levels are states and cities, $\Gamma_i = 2$ for state i would mean that we consider attack scenarios with at most 2 cities in state i being affected.

In case demand location i is affected, the maximum number of individuals in need for treatment at that location is $\hat{d}_i$.[2] A further probabilistic characterization of the possible demand vectors $\mathbf{d}$

[2] The maximum number of affected individuals for a bioattack is estimated by policymakers based on census considerations, *e.g.*, population density, transmission and contagion characteristics, *e.g.*, the required quantity of spores to carry out an attack and their ability to remain aloft and travel further for aerosol attacks, and others. The estimation usually relies on experimental data, see for example our discussion in Section 6.

appears to be, unfortunately, prohibitive for a variety of reasons. One relates to limited historical data, given that bioattacks have so far been rarely encountered in practice. A second, and more important, reason is the nature of terrorist attacks being *adversarial* and *endogenous*, instead of purely random and exogenous. In particular, unlike natural disasters, for instance, bioattacks are pre-meditated and often carefully planned in order to maximize the damage inflicted. The choice of which locations to attack then, which essentially drives the realization of $\mathbf{d}$, is likely to be influenced by the inventory decisions, given that prepositioned inventory (or lack thereof) is partially visible to the public. Both issues of endogeneity and the adversarial choice of the attacked areas are indeed considered and acknowledged by policymakers [52] and academics studying the DAD framework [3, 5, 6, 20, 23, 34, 37]. In the next section, we consider a suitable model for $\mathbf{d}$ that addresses these issues, instead of postulating a probabilistic description.

In general, more (less) inventory in the system leads, on the one hand, to higher (lower) holding costs and, on the other hand, to fewer (more) potential casualties. Therefore, CDC needs to balance inventory costs and the costs of life loss. We consider two possible ways to navigate this tradeoff, leading to two problem formulations:

1. Policymakers explicitly quantify the cost of life lost to be b monetary units. The problem is then to select an inventory allocation so as to minimize holding costs plus worst-case life loss costs under all possible attack scenarios. We refer to this formulation as *Life Loss Cost* (LLC).
2. Policymakers specify survivability targets: let $1-\epsilon_i$ be the target for average survivability in demand node i, for some $\epsilon_i \in [0,1]$, which we refer to as the *survivability target* parameters. The problem is then to select an inventory allocation so as to minimize holding costs, while providing the survivability guarantees implied by the specified targets in all demand locations, under all possible attack scenarios. We refer to this formulation as *Life Loss Guarantee* (LLG).

Our model is well-suited for policy decision-making purposes. In particular, the attack scale parameters $\{\Gamma_i : i \in V\}$ enable policymakers to specify the severity of attacks they want to hedge against, both in terms of *magnitude*, since higher values of Γ for lower level nodes translate into more areas affected, *e.g.*, cities, and in terms of *complexity*, since higher values of Γ for higher level nodes translate into the affected areas being dispersed among more divisions, *e.g.*, states. Furthermore, the life loss cost b in the (LLC) formulation, or the survivability target parameters $\{\epsilon_i : i \in V_D\}$ in the (LLG) formulation, reflect the policymakers' aversion to casualties and insufficient coverage. Based on the policymakers' selections, the model prescribes appropriate inventory prepositioning strategies and elicits their minimum required costs, which allows for tradeoff analyses.

3. Formulation and Solution Approach

For both (LLC) and (LLG), we formulate the inventory prepositioning problem as a DAD problem, or multistage robust optimization problem. In particular, at the first stage, the defender decides on

the prepositioned inventory allocation $\mathbf{x} \in X$. At the second stage, the attacker chooses an attack scenario or demand vector $\mathbf{d}$ from the set of scenarios compatible with the specified attack scale parameters, denoted by $U \subset \mathbb{R}^{|V_D|}$, so as to minimize expected survivors. We also refer to U as the demand uncertainty set. At the third stage, after the realized demand vector $\mathbf{d} \in U$ is revealed, the defender decides on the shipment strategy $\{f_{ij}(\mathbf{d}) : (i,j) \in E\}$. Formally, we model (LLC) as

$$\text{(LLC):} \quad \min_{\mathbf{x},\mathbf{f}(\cdot),\mathbf{s}(\cdot)} \quad \mathbf{h}^T\mathbf{x} \quad + \quad \max_{\mathbf{d}\in U} \quad b\sum_{i\in V_D}\left((1-\overline{\rho})s_i(\mathbf{d}) + \sum_{j:(j,i)\in E}(1-\rho_{ji})f_{ji}(\mathbf{d})\right) \tag{1a}$$

$$\text{subject to} \quad x_i \geq \sum_{j:(i,j)\in E} f_{ij}(\mathbf{d}), \quad \forall i \in V, \forall \mathbf{d} \in U \tag{1b}$$

$$s_i(\mathbf{d}) + \sum_{j:(j,i)\in E} f_{ji}(\mathbf{d}) = d_i, \quad \forall i \in V_D, \forall \mathbf{d} \in U \tag{1c}$$

$$s_i(\mathbf{d}) \geq 0, \quad \forall i \in V_D, \forall \mathbf{d} \in U \tag{1d}$$

$$f_{ij}(\mathbf{d}) \geq 0, \quad \forall (i,j) \in E, \forall \mathbf{d} \in U \tag{1e}$$

$$\mathbf{x} \in X. \tag{1f}$$

The auxiliary variables $\{s_i(\cdot) : i \in V_D\}$ capture demand shortages, *i.e.*, the number of affected individuals left untreated at each node. Note that both the shipment decision variables $\mathbf{f}(\cdot)$ and the shortage variables $\mathbf{s}(\cdot)$ are adjustable, contingent on the realized demand vector $\mathbf{d}$. The objective is to minimize inventory holding costs plus worst-case life loss costs. Constraint (1b) is a node capacity constraint: the total amount of inventory shipped from node i should be less than the amount stockpiled at i. Constraint (1c) defines the demand shortage variable s_i for each node i, and d_i is the ith element of $\mathbf{d}$.

Similarly, we formulate (LLG) as follows:

$$\text{(LLG):} \quad \min_{\mathbf{x},\mathbf{f}(\cdot),\mathbf{s}(\cdot)} \quad \mathbf{h}^T\mathbf{x} \tag{2a}$$

$$\text{subject to} \quad \overline{\rho}s_i(\mathbf{d}) + \sum_{j:(j,i)\in E} \rho_{ji}f_{ji}(\mathbf{d}) \geq (1-\epsilon_i)d_i, \quad \forall i \in V_D, \forall \mathbf{d} \in U \tag{2b}$$

$$\text{(1b), (1c), (1d), (1e), (1f).} \tag{2c}$$

Compared to (LLC), the (LLG) formulation has an objective of minimizing holding costs, while ensuring that the average survivability at each node is higher than the target set, as reflected by the added constraint (2b). Throughout this paper we assume that $\rho_{ji} \geq \overline{\rho}$ for all $(j,i) \in E$, *i.e.*, treatment increases survivability; and $\overline{\rho} \leq 1 - \epsilon_i$ for all $i \in V_D$, otherwise (LLG) reduces to a trivial problem with optimal cost of zero. Note that we impose a survivability target on a node level instead of in aggregation—this provides better control and could ensure equitable population protection across regions. We note later in Theorem 1 that (LLG) is a special case of (LLC).

3.1. Set of Attack Scenarios

We provide a formulation for the set of attack scenarios under consideration. In particular, recall that, given some attack scale parameters $\{\Gamma_i : i \in V\}$, a possible scenario involves demand nodes affected so that, for each administrative division i, no more than Γ_i of its children divisions $\mathcal{C}(i)$ are affected. Consider

$$U := \left\{ \mathbf{d} \in \mathbb{R}^{|V_D|} : \boldsymbol{\xi} \in \mathbb{R}^{|V|},\, d_i = \hat{d}_i \xi_i \,\forall i \in V_D,\, 0 \le \xi_i \le \xi_{\mathcal{P}(i)} \,\forall i \in V \setminus \{0\},\, \xi_0 = 1,\, \sum_{j \in \mathcal{C}(i)} \xi_j \le \Gamma_i \xi_i \,\forall i \in V \right\},$$

where ξ_i indicates whether location i is attacked, and $\hat{d}_i$ is the maximum number of people affected when location i is attacked. As we will show, only the extreme points of U will be relevant for the optimal solution of (LLC) and (LLG). Therefore, to argue that the polytope U represents the set of attack scenarios we are interested in, it suffices to show that its extreme points precisely correspond to the attack scenarios under consideration. Indeed, as we will also show, the auxiliary variables $\boldsymbol{\xi}$ are binary at the extreme points of U. The constraint $\xi_i \le \xi_{\mathcal{P}(i)}$ enforces all parent divisions of an affected node to also be affected, *i.e.*, $\xi_j = 1$ for all j such that $i \in \mathcal{D}(j)$ and i is affected. Finally, constraint $\sum_{j \in \mathcal{C}(i)} \xi_j \le \Gamma_i \xi_i$ ensures that, if i is affected, at most Γ_i of its children divisions are affected, as we required. As a side note, it can be readily seen that the set U also includes all remaining attack scenarios that are compatible with the attack scale parameters, but have a number of affected individuals $d_i \le \hat{d}_i$. To facilitate exposition, we also let

$$\Xi := \left\{ \boldsymbol{\xi} \in \mathbb{R}^{|V|} : 0 \le \xi_i \le \xi_{\mathcal{P}(i)} \,\forall i \in V \setminus \{0\},\, \xi_0 = 1,\, \sum_{j \in \mathcal{C}(i)} \xi_j \le \Gamma_i \xi_i \,\forall i \in V \right\}.$$

We can then simply express $U = \left\{ \mathbf{d} \in \mathbb{R}^{|V_D|} : d_i = \hat{d}_i \xi_i \,\forall i \in V_D,\, \boldsymbol{\xi} \in \Xi \right\}$.

3.2. Solution Approach

By involving both static, $\mathbf{x}$, and recourse decisions, $\mathbf{f}(\cdot)$, $\mathbf{s}(\cdot)$, formulations (LLC) and (LLG) fall into the class of so-called multi-stage adjustable robust optimization problems (ARO). Problems in that class are, in general, computationally intractable [11], because they require the optimization over functions, or policies, instead of vectors, and this makes them infinite-dimensional problems. Specifically, $\mathbf{f}(\cdot)$ and $\mathbf{s}(\cdot)$ are policies that could take different values contingent on the uncertain parameters' realization (namely, the demand $\mathbf{d}$).

Most popular techniques to deal with ARO problems in the literature are heuristics and constraint/column generation methods. One popular heuristic, which we also adopt to tackle our problem, is to limit attention to policies restricted to depend affinely on the uncertain parameters, which are often referred to as *Affine Policies (AP)*, as opposed to *Fully-Adjustable Policies (FP)*.

This restriction often enables tractability, see, *e.g.*, [11]. In our setting, the affinely adjustable robust counterpart of (LLG), for example, is

$$\text{(ALLG):} \min_{\mathbf{x},\mathbf{F},\mathbf{S}} \quad \mathbf{h}^T\mathbf{x} \tag{3a}$$

$$\text{subject to} \quad \overline{\rho}(\mathbf{S}_i^T\mathbf{d}+S_i^0)+\sum_{j:(j,i)\in E}\rho_{ji}(\mathbf{F}_{ji}^T\mathbf{d}+F_{ji}^0)\geq(1-\epsilon_i)d_i, \quad \forall i\in V_D,\,\forall\mathbf{d}\in U \tag{3b}$$

$$x_i\geq\sum_{j:(i,j)\in E}(\mathbf{F}_{ij}^T\mathbf{d}+F_{ij}^0), \quad \forall i\in V,\,\forall\mathbf{d}\in U \tag{3c}$$

$$(\mathbf{S}_i^T\mathbf{d}+S_i^0)+\sum_{j:(j,i)\in E}(\mathbf{F}_{ji}^T\mathbf{d}+F_{ji}^0)=d_i, \quad \forall i\in V_D,\,\forall\mathbf{d}\in U \tag{3d}$$

$$\sum_{j:(j,i)\in E}(\mathbf{F}_{ji}^T\mathbf{d}+F_{ji}^0)\leq d_i, \quad \forall i\in V_D,\,\forall\mathbf{d}\in U \tag{3e}$$

$$\mathbf{F}_{ji}^T\mathbf{d}+F_{ji}^0\geq 0, \quad \forall(i,j)\in E,\,\forall\mathbf{d}\in U \tag{3f}$$

$$\mathbf{x}\in X, \tag{3g}$$

where $\mathbf{F}_{ij}\in\mathbb{R}^{|V_D|}$, $F_{ij}^0\in\mathbb{R}$ for all $(i,j)\in E$, and $\mathbf{S}_i\in\mathbb{R}^{|V_D|}$, $S_i^0\in\mathbb{R}$ for all $i\in V_D$ are vectors of decision variables corresponding to the affine policies' coefficients. The affinely adjustable version of (LLC) is of similar form. We can reformulate (ALLC) and (ALLG) using standard robust optimization techniques (see Corollary 1.3.5. in [11]) to obtain linear optimization problems. Importantly, the resulting linear optimization problems are polynomial in size of the original inputs, enabling tractability and scalability.

On the flipside, tractability of the AP heuristic to deal with ARO problems often comes at the cost of suboptimal solutions. In fact, for some AROs, the suboptimality gap between the objective value under AP and the objective value of the original formulation under FP can grow indefinitely with the dimension of the problem [13]. As we pointed out in the literature review, a handful of papers have recently identified conditions under which AP are indeed optimal. These conditions require the absence of simplex-type constraints in the uncertainty set. In our model, however, the simplex-type constraints $\sum_j\xi_j\leq\Gamma_i\xi_i$ are essential to preclude excessively conservative demand scenarios, where an arbitrary number of nodes are affected. Consequently, given the current state of affairs in the robust optimization literature, the performance of AP heuristics (ALLC) and (ALLG) remains questionable. The next section is devoted to providing evidence that (ALLC) and (ALLG) are indeed likely to produce near-optimal solutions for the original (LLC) and (LLG) formulations.

4. Optimality of Affine Shipment Policies

We provide evidence that affine policies are near-optimal for our problem formulations (LLC) and (LLG). First, and more importantly, we analytically show that under additional assumptions (on the survivability parameters and the network structure), AP are indeed optimal. Second, we conduct numerical studies illustrating that the suboptimality gap remains small for instances that violate the additional assumptions guaranteeing optimality.

4.1. Optimality Result

Consider the following assumptions.

ASSUMPTION 1. *The survival probabilities under treatment are all equal, i.e.,* $\rho_{ij} = \rho \ \forall (i,j) \in E$.

ASSUMPTION 2. *Stockpiles serve all demand nodes in their division, and only demand nodes in their division, i.e.,* $(i,j) \in E \iff j \in \mathcal{D}(i)$.

Assumption 1 requires that the difference in shipment times between stockpile and demand nodes bears no effect on the survivability of treated individuals. Note that this assumption will be violated for biothreats with incubation periods shorter than the shipment times, *e.g.*, for nerve-agents that require treatment within minutes or hours after an attack [52]. However, it will be satisfied for biothreats with longer incubation periods, *e.g.*, for anthrax attacks that are detected early (see also our discussion in Section 6). Assumption 2 requires that stockpiles, which are maintained by administrative divisions, serve affected locations only within their division. In the U.S., this would mean that the stockpiles maintained by states are reserved for usage by their residents. In emergency situations, however, neighboring states, for example, could also provide assistance and this assumption might not hold—we explore this further in the next section.

Let $z^\star_{\text{LLC}}$, $z^\star_{\text{LLG}}$, $z^\star_{\text{ALLC}}$, $z^\star_{\text{ALLG}}$ be the optimal values of formulations (LLC), (LLG), (ALLC), (ALLG) respectively. We have the following result.

THEOREM 1. *Under Assumptions 1 and 2, affine policies are optimal for (LLG), i.e.,*

$$z^\star_{\text{LLG}} = z^\star_{\text{ALLG}}.$$

Furthermore, if $\rho = 1$*, affine policies are optimal for (LLC), i.e.,*

$$z^\star_{\text{LLC}} = z^\star_{\text{ALLC}}.$$

To illustrate the applicability of our methodology beyond the scope of MCM inventory prepositioning, we prove our result in a more general model. In particular, for any node we consider shipments to any of its children nodes, not just its demand (leaf) nodes. Furthermore, we allow demand to occur at any node in V, not just the (leaf) nodes in V_D. We formalize this generalization within the proof of Theorem 1, which can be found in Appendix A, and show that it subsumes problems (LLC) and (LLG) as special cases. We discuss alternative applications in Section 7.

4.2. Performance of Affine Policies for the General Case

To quantify the performance of the AP heuristic for more general cases, we conducted two numerical studies, in which we sequentially relaxed the optimality-guaranteeing Assumptions 1 and 2. For

brevity, we present our studies for the (LLC) formulation—our studies on the (LLG) formulation yielded quantitatively similar results.

We measured the AP heuristic's performance via its suboptimality gap, defined as the relative difference between the heuristic's optimal cost, $z^\star_{\text{ALLC}}$, and the true optimal cost, $z^\star_{\text{LLC}}$, that is $(z^\star_{\text{ALLC}} - z^\star_{\text{LLC}})/z^\star_{\text{LLC}}$. Because the underlying problem is intractable, it is impossible to compute $z^\star_{\text{LLC}}$ for large instances. To quantify the gap then, we rely on the method in [51] to generate a valid lower bound on $z^\star_{\text{LLC}}$. Using this bound instead of the true optimal cost enables us to obtain an *upper bound* on the suboptimality gap. Despite such conservativeness, we found the performance of the AP heuristic to be strong across a wide range of parameter values: median suboptimality gaps were less than 1.5%, and, importantly, they did not grow with the problem's scale.

In Study 1, we relax Assumption 1 and allow the survival probabilities to vary. At a high level, we created 5000 instances of (LLC). For each instance, we generated a random tree-style graph with varying numbers of nodes. Inventory cost, antibiotic efficacy, and attack scale parameters were also randomly sampled—see Table 1 for the sampling ranges, and Appendix B for more details.

Parameter	Sampling Range
Approximate number of nodes	$\{100, 200, 500, 1000, 2000, 5000, 10000\}$
Number of levels	$\{2, 3, 4, 5\}$
Inventory cost	$[0, 0.2] \times b$
Efficacy	$[0, 1]$
Attack scale	$[0, 0.2] \times \lvert C_i \rvert$ for each node i
Demand ($\hat{d}_i$)	[0,1000] for each node i

Table 1 Parameter setup for Study 1: b is the demand loss cost, and $|C_i|$ is the number of children for node i.

In Study 2, we relax Assumption 2 and introduce edges that violate the tree structure. In particular, we considered the same setup as in Study 1, and then for every pair of nodes that were not already connected and resided in adjacent levels in the graph, we added an edge between them with some probability, p_{arc}, so as to obtain "non-tree" graphs. We sampled p_{arc} in $[0, 0.01]$, resulting in graphs that had up to about 200% more edges than the tree-style graphs in Study 1.

Using a high-performance computing cluster, we ran the AP heuristic and the exact solution method in [51] for each instance. Each run was given a 4-hour time limit, 16 GB of memory, and 2 CPUs (2.1 GHz each). Across both Studies 1 & 2, AP successfully solved all 10000 instances, with an average run time of 6 seconds. Although the exact solution method managed to solve only 2% of the instances, it generated valid lower bounds for 91% of them, with an average run time of 14116 seconds. The generated lower bounds were then used for our comparisons.

Results. Table 2 provides statistics of the upper bounds on the suboptimality gaps we obtained by using the generated valid lower bounds on the optimal costs. Qualitatively, our analysis reveals that

Min	1st Quantile	Median	Mean	3rd Quantile	Max
0.0%	0.4%	1.4%	1.9%	3.0%	9.9%

Table 2 Statistics of upper bounds on suboptimality gaps for AP heuristic for Studies 1 & 2.

the suboptimality gap is small and does not have strong correlation with any of the parameters we varied, including, among others, the number of nodes and the attack severity. In particular, the studies suggest that the suboptimality gap does not grow with the network size. This is a key consideration given that we cannot tractably characterize the gap for even larger networks due to limitations of the exact solution method. More details are included in Appendix B.

5. Optimizing Dispensing Capacity

In this section, we relax the assumption of sufficient dispensing capacity being pre-installed, and jointly optimize inventory prepositioning and dispensing capacity decisions. By dispensing capacity, we refer to the ability of the appropriate local authorities to distribute MCMs to the general public at demand nodes subsequent to a bioattack—technically, it corresponds to the MCM delivery rate from so-called Receive, Stage and Store facilities to end users. Such joint optimization is essential: too much inventory and not enough dispensing capacity would result in inventory buildup while the population suffers from dispensing delays; too little inventory and too much capacity would result in idle workers and under-utilized resources.

Limited dispensing capacity could introduce further delays before the affected population has access to MCMs, which in turn could affect their efficacy. Unlike delays due to transportation lead times, which are independent of the MCM quantity shipped, delays due to limited dispensing capacity are dependent on the MCM quantity being dispensed. This creates the need to introduce a time dimension explicitly in our model.

We illustrate our approach using the (LLG) formulation. Subsequent to an attack, consider T discrete time periods, each with duration δ, indexed by $t = 1, 2, \ldots, T$. Let u_j be the dispensing capacity in MCM units per time period at demand node j. In view of this capacity, a shipment from some node i, although it would still arrive at j after τ_{ij} time periods, might not be immediately available to the population at its entirety. To capture this, let f_{ijt} be the MCM amount shipped from i and dispensed at j at time t. To measure antibiotic efficacy, let ρ_t be the survival probability if treatment is received t time periods after the attack. We use δ small enough so that the probability of survival remains approximately constant and equal to ρ_t within time period $[(t-1)\delta, t\delta]$.

Let c_j be the cost per unit capacity installation (*e.g.*, training of staff, preparation and maintenance of dispensing facilities), and let p_j be the cost per unit of MCM dispensed after an attack. The joint inventory prepositioning and dispensing capacity optimization problem can be cast as:

$$\text{(LLGC):} \quad \min_{\mathbf{x},\,\mathbf{u},\,\mathbf{f}(\cdot),\,\mathbf{s}(\cdot)} \; \max_{\mathbf{d}\in U} \quad \mathbf{h}^T\mathbf{x} + \mathbf{c}^T\mathbf{u} + \mathbf{p}^T\mathbf{f}(\mathbf{d}) \tag{4a}$$

$$\text{subject to} \quad x_i \geq \sum_{j:(i,j)\in E} \sum_{t=1}^{T} f_{ijt}(\mathbf{d}), \quad \forall i \in V, \forall \mathbf{d} \in U \tag{4b}$$

$$s_j(\mathbf{d}) + \sum_{i:(i,j)\in E} \sum_{t=1}^{T} f_{ijt}(\mathbf{d}) = d_j, \quad \forall j \in V_D, \forall \mathbf{d} \in U \tag{4c}$$

$$\overline{\rho} s_j(\mathbf{d}) + \sum_{i:(i,j)\in E} \sum_{t=1}^{T} \rho_t f_{ijt}(\mathbf{d}) \geq (1-\epsilon_j) d_j, \quad \forall j \in V_D, \forall \mathbf{d} \in U \tag{4d}$$

$$\sum_{i:(i,j)\in E} f_{ijt}(\mathbf{d}) \leq u_j, \quad \forall j \in V_D, \forall \mathbf{d} \in U, \forall t \in [T] \tag{4e}$$

$$f_{ijt}(\mathbf{d}) = 0, \quad \forall (i,j) \in E, \forall \mathbf{d} \in U, \forall t \in [\tau_{ij} - 1] \tag{4f}$$

$$f_{ij}(\mathbf{d}) \geq 0, \quad \forall (i,j) \in E, \forall \mathbf{d} \in U \tag{4g}$$

$$s_i(\mathbf{d}) \geq 0, \quad \forall i \in V_D, \mathbf{d} \in U \tag{4h}$$

$$\mathbf{x} \in X, \mathbf{u} \geq 0. \tag{4i}$$

Constraint (4e) ensures that no more than u_j MCM units are dispensed at demand node j across all origin nodes i, at each time period t. Constraint (4f) ensures that no MCMs shipped from node i can be dispensed at node j before the associated shipment lead time τ_{ij}. All other constraints are similar in spirit to the ones we have in (LLG).

By restricting the flow and demand shortage variables to be affine in demand vector $\mathbf{d}$, the robust counterpart of (LLGC) remains a linear optimization problem. Thus, we are able to solve large problem instances that match the fidelity and scale of a national biodefense network.

Finally, note that introducing a time dimension here can be thought of as simply introducing additional edges in the set E. Therefore, beyond adding capacity variables and accounting for costs, the only essential way that (LLGC) differs from (LLG) is the inclusion of capacity constraints (4e). In the Appendix B, we present a numerical study similar to the ones in §4.2, in which we quantify the AP heuristic's suboptimality gap for problems including such capacity variables, costs, and constraints. As before, we find performance to be strong, with a median suboptimality gap less than 1%.

6. Strategic National Stockpile Supply Chain

In this section, we apply our work in a case study on SNS design for aerosolized *Bacillus anthracis* (anthrax) attacks in the U.S., which are of particular interest to public health experts owing to their relatively high probability of occurrence and potentially devastating impact [52]. Anthrax spores spread easily in the air to affect a large number of people through inhalation, with high fatality rates if left untreated for even just a few days. More precisely, if t is the time between

infection and treatment, then, according to the studies of [17, 18, 57, 58], the survival probability $\rho(t)$ of anthrax infected population can be approximated, for $t \leq 200$ hours, by

$$\rho(t) = e^{-(0.004t)^2}. \tag{5}$$

Considering the time it would take to detect an anthrax attack and the time to ship and dispense MCMs from warehouses to individuals, which is likely to be on the order of days [52], CDC is exploring different options in prepositioning MCMs and coordinating the SNS network to achieve better antibiotic efficacy. We now discuss how to calibrate our model to tackle this problems.

6.1. Model calibration

For the purposes of this study, we consider the (LLGC) formulation, that is, policymakers specify coverage targets and the model optimizes the required inventory, capacity, and dispensing costs.

Network. There are 12 federally-managed SNS warehouses for MCM storage, with their exact locations being classified for security reasons. It is known that the locations have been chosen such that MCMs can be transported from them to any state within 24-36 hours of a deployment call [46]. In our model, it is appropriate then to model these warehouses as a single virtual federal stockpile node. In case of an attack, inventories can be shipped from the federal inventory node to state-managed warehouses, where additional inventory is usually stored. We therefore model the SNS stockpile network with 1 federal stockpile node, 52 state and special district nodes. Inventory from the state warehouses can then be forwarded to city, county, and local authorities for dispensing to the general public. Using census data to obtain precise information about major cities or so-called Metropolitan Statistical Areas (MSAs), counties, we include 377 MSA nodes, 3221 county nodes, and 77072 neighborhood nodes. In addition, we also include 77072 household nodes, each denoting the aggregate demand in a neighborhood. The federal stockpile node serves all downstream state nodes. State nodes serve not only their downstream MSAs, but also all MSA nodes in neighboring states. MSAs serve predominantly their downstream counties, but also neighboring MSAs and counties, and so forth. In total, the network comprises 157795 nodes and 335010 edges.

Transportation times. As we remarked above, the time to ship MCMs from the federal node to any state level node varies between 24 to 36 hours. From state level to MSA/county level, we assume the transportation time to be within 2 to 6 hours, *e.g.*, if shipped via ground transportation, depending on the geography of the state. According to field experiments, including, for example, a study in Philadelphia in 2005 [1] and another one in Minneapolis in 2008 [8], policymakers have estimated the shipment of MCMs from local authorities to dispensing points to take 10 to 12 hours.

Holding costs. We consider the SNS design for one MCM type, *e.g.*, *ciprofloxacin* or *doxycycline*, to treat individuals affected by anthrax attacks in the U.S. Inventory costs include the cost of

purchasing, storage, management, replenishment, and shipment from manufacturer to warehouse. According to a commissioned paper by the Institute of Medicine [28], the costs for anthrax antibiotics stored at regional warehouses amount to $2.10 per person, while home kits are more expensive at $10 per person, due to higher packaging and delivery costs. We estimate the federal, state, MSA, county, and neighborhood holding costs to be between $1 and $2 per person. In general, upstream holding costs are cheaper than downstream ones due to anticipated scale economies. Annual management and replenishment costs are 85% of purchase cost according to [28].

Demand. To estimate the demand volume in a neighborhood at each MSA in case of an attack, we assume that an airborne attack has the same spread radius in every geographic region, and we use a value of 3000 square miles as in the field study on the Minneapolis MSA [8]. The affected population at node i can then be calculated as (population density at i) $\times$ $\min\{3000 \text{ sq miles}, \text{area of neighborhood}\}$. We obtained population density and area data of neighborhood from statistics of the Census Bureau [54].

Dispensing modes and costs. Two dispensing modes are currently available: (1) Point-of-Dispenses (PODs), and (2) door-to-door delivery by the United States Postal Service (USPS) [56]. PODs are what the supply chain literature would consider a "pull" strategy, in which people commute to pre-specified locations in their neighorhoods to pick up MCMs for their families. The USPS option is a "push" strategy, in which the staff members of USPS deliver MCMs into the mailboxes of each household. Dispensing capacity is shared among neighborhoods, so that there are approximately 2 PODs per county.

Primary costs associated with the PODs setup are: medical staff training costs and wages, supervisor training costs and wages, security wages, and public facility rental costs. Using the recommendation in [59], we estimate a labor wage of $18.64 per hour, $2785 security cost per POD per day, and $5000 administrative overhead cost per POD per day; for each POD, the total number of staff is 300 and can support a dispensing service rate of 1000 patients per hour. The resulting cost of operating a POD during dispensing phase is $3.39 per hour for a one-person/hour increase in dispensing capacity. For capacity installation costs, we assume that a two-day drill is performed annually, resulting in $162.72 per unit capacity per year.

The main costs associated with the USPS option are the wages for the delivery staff and wages for the accompanying security officers. Based on the Minneapolis-St. Paul study [28] with USPS delivery involving 179 staff, wage of $23.72 per person per hour, covering 205,000 households in 8.5 hours (estimate of 6.9 hours spent on transportation and 1.6 hours on material handling) for a population density of 1000 households per square mile, we derive a population-density-dependent USPS dispensing cost function: $\$ (356/\sqrt{\text{density}} + 19)$ per hour for every one-person/hour increase in delivery capacity. For capacity installation costs, we also assume a two-day drill every year. In

addition to these costs, the USPS option also has an upper-bound on how many workers can take up the delivery job, as each delivery worker has to be accompanied by one security officer, and the total number of security officers is likely a bottleneck during such an emergency situation. We assume that each county can support at most 10 delivery workers at any given time.

Survival parameters. We choose a time step of $\delta = 8$ hours, in accordance with the standard work shift length of 8 hours per group of staff for dispensing and delivery [52]. Then based on the survival probability calculation aforementioned in (5), for a detection time of 48 hours, we calculate $\{\rho_1, \rho_2, \rho_3, \rho_4, \dots\}$ to be $\{0.957, 0.944, 0.929, 0.912, \dots\}$. If a person inhales anthrax spores and is left untreated, his/her survival probability is estimated to be 0.2 [55].

6.2. Results

Having calibrated the model, we compute the optimal costs for different policy parameters: survivability target, attack scale, detection time, and dispensing mode. Table 3 is a summary of the model inputs and outputs. Table 4 provides solution times and information about the size of the resulting optimization problems.

Input	Output
Survivability target: ϵ_i	Total cost (inventory+capacity+dispensing)
Attack scale: Γ_i	Optimal inventory prepositioning strategy: $\mathbf{x}^\star$
Detection time: τ_0	Optimal capacity installation: $\mathbf{u}^\star$

Table 3 Input/output for our case study.

Min	1st Quantile	Median	Mean	3rd Quantile	Max
192	719	798	671	839	954

Table 4 Solution time (seconds). The resulting optimization problems for the case study instances have 1100000 variables and 115000 constraints, approximately, after pre-processing using Gurobi.

The default input parameters we consider are attack scale parameters $\Gamma_i = 2$, $\forall i \in V$, and detection time $\tau_0 = 48$ hours.[3] Below, we perform several sensitivity analyses to illustrate the flexibility of our framework and how it can be used to guide policymaking.

Varying survivability target. Figure 2 depicts the minimum required annual cost as a function of the survivability target. For example, for an 80% survivability target the annual budget needs to be about \$194 million dollars; if the survivability target is set to 92%, the budget would need to be increased to \$553 million, approximately. The curve in Figure 2 also illustrates that increasing

[3] For a discussion of anticipated detection times in case of an anthrax attack, see [52].

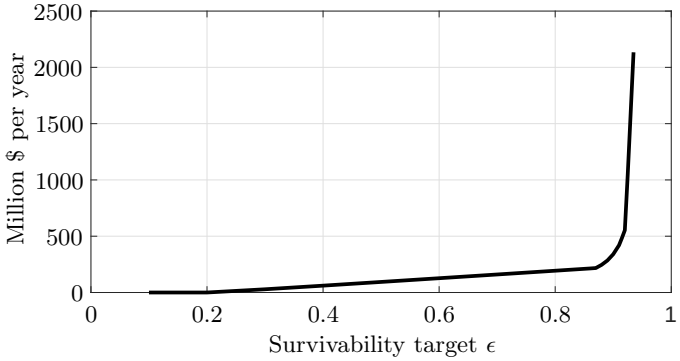

Figure 2 **Minimum required annual cost for different survivability targets.**

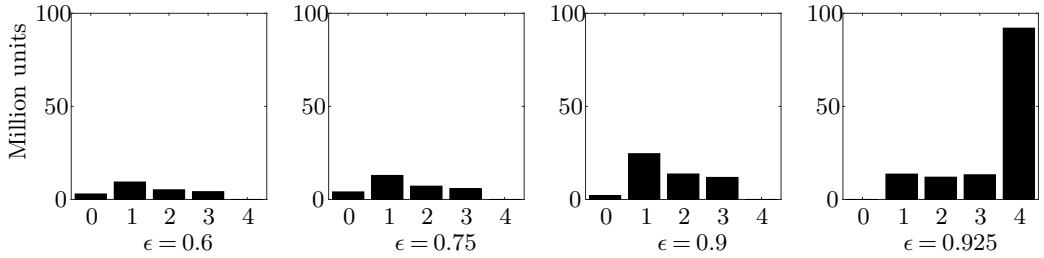

Figure 3 **Amount (y-axis) and location (x-axis) of prepositioned inventory for different survivability targets $\{0.6, 0.75, 0.9, 0.925\}$; at the x-axis, level 0 is federal, level 1 is state, level 2 is MSA, level 3 is county, and level 4 is predispensed medical kits in households.**

the survivability target beyond 90% requires a rather steep cost increase. We can also interpret the curve as a Pareto frontier associated with the cost-survivability tradeoff: the region left to the curve represents the achievable outcomes.

Figure 3 shows how much inventory and at which level it needs to be prepositioned for different survivability targets. It can be seen that as the target increases, more inventory is required. Notably, for a target less than 90%, predispensed medical kits are redundant, whereas for a higher target a rather large number of them is required.

Figure 4 shows county-level capacity decisions for different survivability targets. Although required capacity naturally increases with the desired target, interestingly, it drops at a steep rate beyond a certain point.

Put together, the analysis of varying survivability target illustrates a phase change that takes place around the target of 90%, approximately. To better understand this phenomenon, recall first that survivability targets translate into certain responsiveness requirements, through relationship (5), for example, with higher targets calling for higher responsiveness. For targets lower than 90% then, storing inventory upstream at federal/state levels provides acceptable responsiveness, and because it is so much more cost-efficient per unit and also enjoys pooling benefits (in the sense

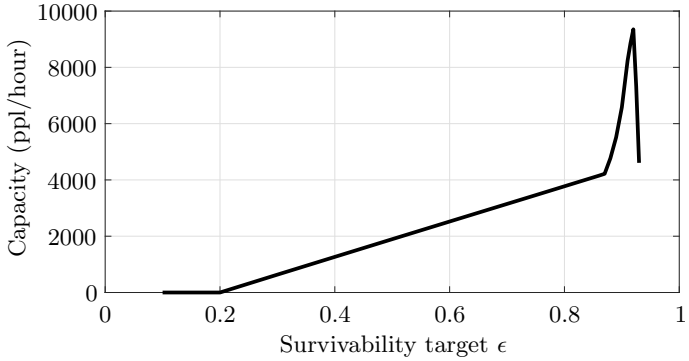

Figure 4 **Average capacity installation per county (number of people served per hour) for different survivability targets.**

that one kit could serve multiple downstream recipient households), it makes predispensed kits redundant. For targets higher than 90%, upstream storage no longer provides adequate responsiveness for the entire SNS network, and predispensed kits become necessary, which explains their introduction to the prepositioning strategy. The high rate at which these kits are introduced then, compared with kits stored upstream, can be explained by the lack of pooling benefits (one predispensed kit could only serve a single household). This in turn explains the observed steep increase in inventory costs and the decrease in required dispensing capacity.

Varying attack scale. We now explore different attack scale parameters. Figure 5 reports the minimum required annual costs and optimal inventory amounts at each level, under nine different cases: $\{1, 2, 3\}$ states being attacked and each state having $\{1, 2, 3\}$ MSAs or counties affected. Recall that the first parameter reflects the geographic complexity of the attack, and the second parameter the magnitude of an attack within a state.

Varying detection time. Figure 6 shows the cost-survivability trade-off for different detection times τ_0, highlighting the importance of timely response mechanisms. In particular, if the detection mechanism is able to identify an attack and start deployment within 60 hours, approximately \$788 million USD is needed for a 90% survivability target annually. For enhanced detection mechanism capable of detecting attacks within 24 hours, the corresponding annual costs reduce to just over \$210 million.

Choosing between different dispensing modes. Figure 7 illustrates the cost-survivability trade-offs for different dispensing modes. Using USPS dispensing mode alone is generally more costly as compared with using PODs, due to the upper bound on USPS capacity and reliance on home medical kits. Overall, we find that even though USPS is empirically efficient in high density neighborhoods, PODs should still remain as the primary dispensing channel due to their scalability.

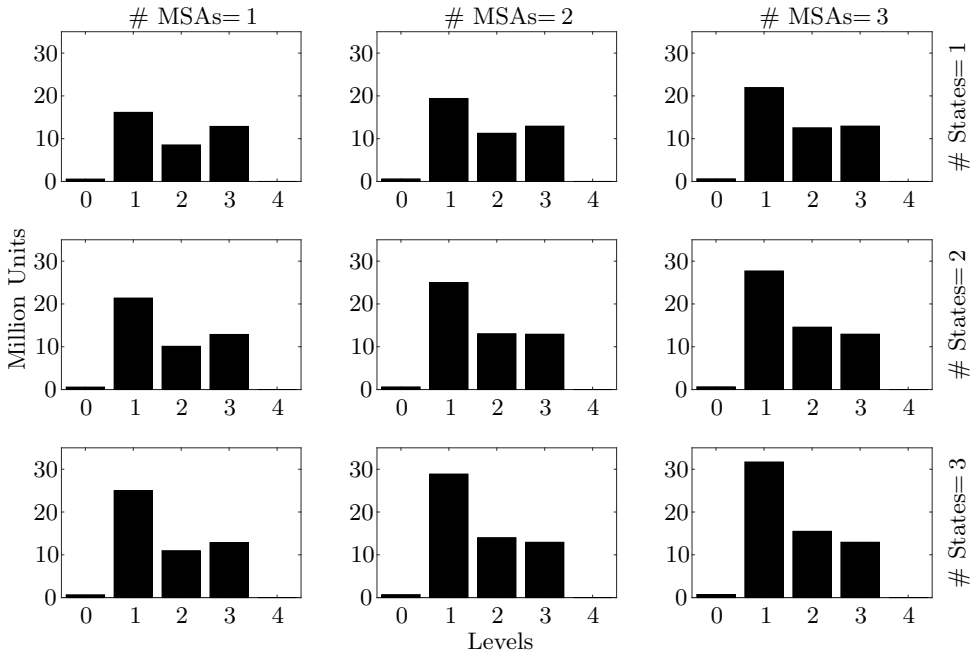

Figure 5 **Amount (y-axis) and location (x-axis) of prepositioned inventory for different attack scale parameters; at the x-axis, level 0 is federal, level 1 is state, level 2 is MSA, level 3 is county, and level 4 is predispensed medical kits in households.**

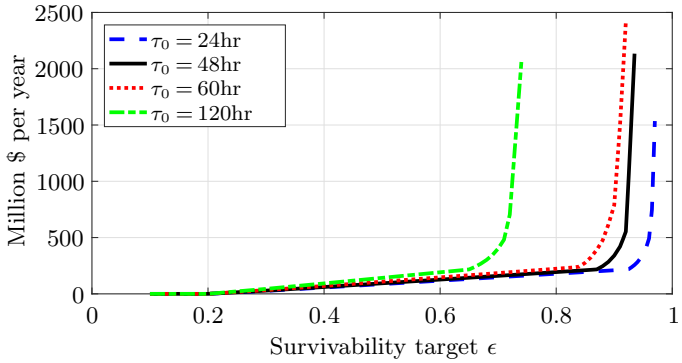

Figure 6 **Minimum required annual costs for different survivability targets and detection times.**

7. Extensions and Concluding Remarks

In this paper, we considered a framework to tackle the problem of designing a cost-effective and responsive public health stockpile supply chain for protection against bioterrorism attacks. Our framework captured many of the key drivers facing CDC in maintaining the Strategic National Stockpile, for example, holding costs, pooling, capacity, and responsiveness. Our methodology can be extended to capture additional SNS design considerations, or used in other supply chain applications. For example, an interesting direction for future research could be to allow for multiple

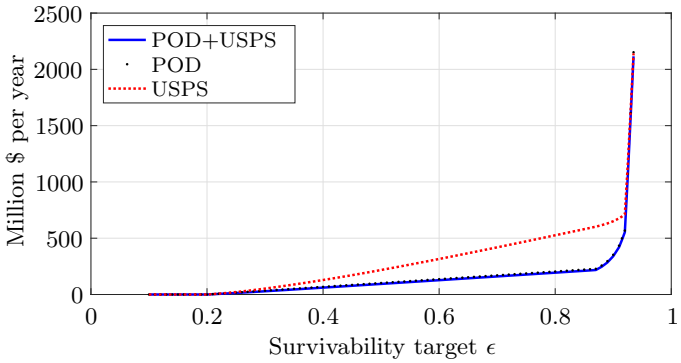

Figure 7 **Minimum required annual costs for different survivability targets and dispensing modes.**

types of MCMs, for the same or different biothreats. In that case, our model could be used to guide the design of CDC's overall portfolio and mix of MCMs in view of budgetary constraints.

Acknowledgment

This work is supported by Ford Motor Company, The Boeing Company, and United Technologies Corporation.

References

[1] Mary Agócs, Shannon Fitzgerald, Steven Alles, Graham John Sale, Victor Spain, Edward Jasper, Thomas Lee Grace, and Esther Chernak. Field Testing a Head-of-Household Method to Dispense Antibiotics. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, 5(3):255–267, sep 2007.

[2] Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. Network flows: theory, algorithms, and applications. *Prentice Hall*, 1993.

[3] David L. Alderson, Gerald G. Brown, Matthew W. Carlyle, and Louis Anthony Cox. Sometimes There Is No "Most-Vital" Arc: Assessing and Improving the Operational Resilience of Systems. *Military Operations Research*, 18(1):21–37, 2013.

[4] David L Alderson, Gerald G Brown, and W Matthew Carlyle. Assessing and Improving Operational Resilience of Critical Infrastructures and Other Systems. In *Bridging Data and Decisions*, number June 2017, pages 180–215. INFORMS, sep 2014.

[5] David L Alderson, Gerald G Brown, and W Matthew Carlyle. Operational Models of Infrastructure Resilience. *Risk Analysis*, 35(4):562–586, apr 2015.

[6] David L Alderson, Gerald G Brown, W Matthew Carlyle, and R Kevin Wood. Solving Defender-Attacker-Defender Models for Infrastructure Defense. In *12th INFORMS Computing Society Conference*, pages 28–49. INFORMS, apr 2011.

[7] Amir Ardestani-Jaafari and Erick Delage. Robust Optimization of Sums of Piecewise Linear Functions with Application to Inventory Problems. *Operations Research*, 64(2):474–494, apr 2016.

[8] Assistant Secretary for Preparedness and Response. National postal model for the delivery of medical countermeasures. Technical report, Department of Health and Human Services, Washington, D.C., 2010.

[9] Alper Atamtürk and Muhong Zhang. Two-Stage Robust Network Flow and Design Under Demand Uncertainty. *Operations Research*, 55(4):662–673, aug 2007.

[10] Michael Philip Atkinson. *Mathematical models of terror interdiction*. PhD thesis, Stanford University, 2009.

[11] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. Robust Optimization. *Princeton University Press*, 2009.

[12] Oded Berman and Arieh Gavious. Location of terror response facilities: A game between state and terrorist. *European Journal of Operational Research*, 177(2):1113–1133, mar 2007.

[13] Dimitris Bertsimas and Vineet Goyal. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming*, 134(2):491–531, sep 2012.

[14] Dimitris Bertsimas, Vineet Goyal, and Xu Andy Sun. A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Mathematics of Operations Research*, 36(1):24–54, 2011.

[15] Dimitris Bertsimas, Dan A. Iancu, and Pablo A. Parrilo. Optimality of Affine Policies in Multistage Robust Optimization. *Mathematics of Operations Research*, 35(2):363–394, may 2010.

[16] Dena M. Bravata, Gregory S. Zaric, Jon-Erik C. Holty, Margaret L. Brandeau, Emilee R. Wilhelm, Kathryn M. McDonald, and Douglas K. Owens. Reducing Mortality from Anthrax Bioterrorism: Strategies for Stockpiling and Dispensing Medical and Pharmaceutical Supplies. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, 4(3):244–262, sep 2006.

[17] Ron Brookmeyer. The statistical analysis of truncated data: application to the Sverdlovsk anthrax outbreak. *Biostatistics*, 2(2):233–247, jun 2001.

[18] Ron Brookmeyer, Elizabeth Johnson, and Sarah Barry. Modelling the incubation period of anthrax. *Statistics in Medicine*, 24(4):531–542, feb 2005.

[19] Gerald Brown, Matthew Carlyle, Javier Salmer?n, and Kevin Wood. Defending Critical Infrastructure. *Interfaces*, 36(6):530–544, dec 2006.

[20] Gerald G Brown, W Matthew Carlyle, and R Kevin Wood. Optimizing Department of Homeland Security Defense Investments : Applying Defender-Attacker (-Defender) Optimization To Terror Risk Assessment and Mitigation. In *Department of Homeland Security Bioterrorist Risk Assessment: A Call for Change*. National Academies Press, Washington, D.C., 2008.

[21] Aakil M. Caunhye, Xiaofeng Nie, and Shaligram Pokharel. Optimization models in emergency logistics: A literature review. *Socio-Economic Planning Sciences*, 46(1):4–13, mar 2012.

[22] Melih Çelik, Özlem Ergun, Ben Johnson, Pinar Keskinocak, Álvaro Lorca, Pelin Pekgün, and Julie Swann. Humanitarian Logistics. In *2012 TutORials in Operations Research*, pages 18–49. INFORMS, oct 2012.

[23] Timothy C. Y. Chan, Z. J. M. Shen, and Auyon Siddiq. Robust Defibrillator Deployment Under Cardiac Arrest Location Uncertainty via Row-and-Column Generation. 2017.

[24] David L. Craft, Lawrence M. Wein, and Alexander H. Wilkins. Analyzing Bioterror Response Logistics: The Case of Anthrax. *Management Science*, 51(5):679–694, may 2005.

[25] Özlem Ergun, Gonca Karakus, Pinar Keskinocak, Julie Swann, and Monica Villarreal. Operations Research to Improve Disaster Supply Chain Management. In *Wiley Encyclopedia of Operations Research and Management Science*. John Wiley & Sons, Inc., Hoboken, NJ, USA, jan 2011.

[26] Boaz Golany, Edward H Kaplan, Abraham Marmur, and Uriel G Rothblum. Nature plays with dice ? terrorists do not: Allocating resources to counter strategic versus probabilistic risks. *European Journal of Operational Research*, 192(1):198–208, jan 2009.

[27] Boaz Golany, Moshe Kress, Michal Penn, and Uriel G Rothblum. Network Optimization Models for Resource Allocation in Developing Military Countermeasures. *Operations Research*, 60(1):48–63, feb 2012.

[28] James Guyton, Robert Kadlec, Chandresh Harjivan, Shabana Farooqi, Sheana Cavitt, and Joseph Buccina. A Cost and Speed Analysis of Strategies for Prepositioning Antibiotics for Anthrax. *Commissioned paper by the Institute of Medicine*, 2011.

[29] Mengran Hao, Shilan Jin, and Jun Zhuang. Robustness of Optimal Defensive Resource Allocations in the Face of Less Fully Rational Attacker. Technical report, Homeland Security Center, 2009.

[30] Nathaniel Hupert, Wei Xiong, Kathleen King, Michelle Castorena, Caitlin Hawkins, Cindie Wu, and John A. Muckstadt. Uncertainty and Operational Considerations in Mass Prophylaxis Workforce Planning. *Disaster Medicine and Public Health Preparedness*, 3(S2):S121–S131, dec 2009.

[31] Dan A Iancu, Maxim Sviridenko, Mayank Sharma, and Maxim Sviridenko. Supermodularity and Affine Policies in Dynamic Robust Optimization. *Operations Research*, 61(4):941–956, aug 2013.

[32] Thomas Inglesby and Barbara Ellis. Division of Strategic National Stockpile (DSNS) Program Review: A Report from the Board of Scientific Counselors (BSC). Technical report, Center for Disease Control and Prevention, Atlanta, 2012.

[33] Jocelyn Kaiser. Taking Stock of the Biodefense Boom. *Science*, 333(6047):1214–1214, sep 2011.

[34] Edward H. Kaplan. OR Forum - Intelligence Operations Research: The 2010 Philip McCord Morse Lecture. *Operations Research*, 60(6):1297–1309, dec 2012.

[35] Kathleen King. Logistical Models For Planning And Operating Medical Countermeasure Distribution Networks During Public Health Emergencies. *PhD dissertation, Cornell University*, 2012.

[36] Kathleen King and John A. Muckstadt. Evaluating Planned Capacities for Public Health Emergency Supply Chain Models. Technical report, Cornell University, 2009.

[37] Gary Lazzaro. *Tri-level optimization algorithms for solving defender-attacker-defender network models*. Phd dissertation, Naval Postgraduate School, 2016.

[38] Eva K. Lee, Chien-Hung Chen, Ferdinand Pietz, and Bernard Benecke. Modeling and Optimizing the Public-Health Infrastructure for Emergency Response. *Interfaces*, 39(5):476–490, oct 2009.

[39] Eva K. Lee, Siddhartha Maheshwary, Jacquelyn Mason, and William Glisson. Decision support system for mass dispensing of medications for infectious disease outbreaks and bioterrorist attacks. *Annals of Operations Research*, 148(1):25–53, nov 2006.

[40] Eva K. Lee, Siddhartha Maheshwary, Jacquelyn Mason, and William Glisson. Large-Scale Dispensing for Emergency Response to Bioterrorism and Infectious-Disease Outbreak. *Interfaces*, 36(6):591–607, dec 2006.

[41] Eva K. Lee, Hannah K. Smalley, Yang Zhang, Ferdinand Pietz, and Bernard Benecke. Facility location and multi-modality mass dispensing strategies and emergency response for biodefence and infectious disease outbreaks. *International Journal of Risk Assessment and Management*, 12(2/3/4):311, 2009.

[42] National Research Council. *Department of Homeland Security Bioterrorism Risk Assessment: A Call for Change*. The National Academies Press, Washington DC, 2008.

[43] Anna Nicholson, Scott Wollek, Benjamin Kahn, and Jack Hermann. The Nation's Medical Countermeasure Stockpile. *National Academies Press*, 2016.

[44] Office of Public Health Preparedness and Response. Board of Scientific Counselors, 2016.

[45] Office of Public Health Preparedness and Response. Public Health Preparedness Capabilities: National Standards for State and Local Planning, 2016.

[46] Office of Public Health Preparedness and Response. Strategic National Stockpile, 2016.

[47] Jesus Rios and David Rios Insua. Adversarial Risk Analysis for Counterterrorism Modeling. *Risk Analysis*, 32(5):894–915, may 2012.

[48] David Simchi-Levi, William Schmidt, and Yehua Wei. From superstorms to factory fires: managing unpredictable supply-chain disruptions. *Harvard Business Review*, 92(1-2):96, 2014.

[49] David Simchi-Levi, William Schmidt, Yehua Wei, Peter Yun Zhang, Keith Combs, Yao Ge, Oleg Gusikhin, Michael Sanders, and Don Zhang. Identifying Risks and Mitigating Disruptions in the Automotive Supply Chain. *Interfaces*, 45(5):375–390, oct 2015.

[50] David Simchi-Levi, He Wang, and Yehua Wei. Increasing Supply Chain Robustness through Process Flexibility and Inventory. *Working paper, MIT*, 2016.

[51] David Simchi-Levi, He Wang, and Yehua Wei. Constraint Generation for Two-Stage Robust Network Flow Problem. *INFORMS Journal on Optimization, forthcoming*, 2018.

[52] Clare Stroud, Kristin Viswanathan, Tia Powell, and Robert R Bass. Prepositioning Antibiotics for Anthrax. *Institute of Medicine*, (September):1–387, 2011.

[53] Joline Uichanco. A robust model for pre-positioning emergency relief items before a typhoon with an uncertain trajectory. *Working paper, University of Michigan*, 2015.

[54] United States Census Bureau. Population, Housing Units, Area, and Density: 2000, 2000.

[55] U.S. Food and Drug Administration. Anthrax, 2015.

[56] L. M. Wein, D. L. Craft, and E. H. Kaplan. Emergency response to an anthrax attack. *Proceedings of the National Academy of Sciences*, 100(7):4346–4351, apr 2003.

[57] D. A. Wilkening. Sverdlovsk revisited: Modeling human inhalation anthrax. *Proceedings of the National Academy of Sciences*, 103(20):7589–7594, may 2006.

[58] D. A. Wilkening. Modeling the Incubation Period of Inhalational Anthrax. *Medical Decision Making*, 28(4):593–605, jul 2008.

[59] Gregory S. Zaric, Dena M. Bravata, Jon-Erick Cleophas Holty, Kathryn M. McDonald, and Margaret L. Brandeau. Modeling the Logistics of Response to Anthrax Bioterrorism. *Medical Decision Making*, 28(3):332–350, 2008.

[60] Jun Zhuang and Vicki M. Bier. Balancing Terrorism and Natural Disasters?Defensive Strategy with Endogenous Attacker Effort. *Operations Research*, 55(5):976–991, oct 2007.

Appendix A: Proofs

Proof of Theorem 1. Under Assumption 2, our model can be reformulated to one where, instead of inventory shipments over the edges of (V,E), $\{f_e : e \in E\}$, we consider inventory shipments over the edges of the directed out-tree (V,T) with root node 0 and $T := \{(i,j) : i \in V, j \in \mathcal{C}(i)\}$. This is because, under Assumption 2, any inventory shipment to node j occurs from some "parent node" i such that $j \in \mathcal{D}(i)$, *i.e.*, $\mathcal{P}^{L-\ell}(j) = i$ for some ℓ, where $\mathcal{P}^n(k) := \mathcal{P}(\mathcal{P}^{n-1}(k))$ for n positive integer, $k \in V$ and $\mathcal{P}^0(k) := k$. Thus, any f_{ij} can be thought of as a flow along the unique path $i \to \mathcal{P}^{L-\ell-1}(j) \to \ldots \mathcal{P}(j) \to j$. It can then be readily seen that our model entails a *path-based* network flow formulation for inventory shipments in the graph (V,T). In the rest of the proof, we consider the associated *edge-based* formulation, and denote the shipments over the edges with $\mathbf{y} = \{y_t : t \in T\}$. Clearly, this is without loss, since for every feasible $(\mathbf{f},\mathbf{s})$ in the path-based formulation, there exists $\mathbf{y}$ such that $(\mathbf{y},\mathbf{s})$ is feasible for the edge-based formulation, and vice versa (for more details see Chapter 3.5 in [2]). Furthemore, we assume that demand occurs at every node; specifically, demand at node i is $d_i = \hat{d}_i \xi_i$ for all $i \in V$ and some $\boldsymbol{\xi} \in \Xi$. Unmet demand at each node in $V \setminus V_D$ is treated in the same way as at nodes in V_D: either penalized at a cost rate b under the (LLC), or subject to a guarantee under the (LLG).

To prove optimality of affine policies, it suffices to show that for any static inventory allocation decision $\mathbf{x} \in X$, there exist policies that are affine in the uncertain demand for the adjustable decisions and achieve the same worst-case cost under fully-adjustable policies. Thus, we henceforth consider the static inventory allocation decision $\mathbf{x}$ as fixed.

We first deal with (LLC). At the end of the proof, we argue how (LLG) can be cast as a special case of (LLC). Using the edge-based formulation, the more general demand model, and a fixed inventory allocation $\mathbf{x} \in X$, it can be readily seen that when $\rho = 1$, the (LLC) problem is equivalent to (in the sense that they have the same optimal set)[4]

$$\begin{aligned}
\min_{\mathbf{y}(\cdot),\mathbf{s}(\cdot)} \max_{\boldsymbol{\xi}\in\Xi} \quad & \mathbf{1}^T \mathbf{s}(\boldsymbol{\xi}) \\
\text{subject to} \quad & s_i(\boldsymbol{\xi}) + y_{\mathcal{P}(i)i}(\boldsymbol{\xi}) + x_i \geq \sum_{j\in\mathcal{C}(i)} y_{ij}(\boldsymbol{\xi}) + \hat{d}_i \xi_i, \ \forall i \in V, \forall \boldsymbol{\xi} \in \Xi \\
& y_{\mathcal{P}(i)i}(\boldsymbol{\xi}) + x_i \geq \sum_{j\in\mathcal{C}(i)} y_{ij}(\boldsymbol{\xi}), \ \forall i \in V, \forall \boldsymbol{\xi} \in \Xi \\
& \mathbf{y}(\cdot), \mathbf{s}(\cdot) \geq 0,
\end{aligned}$$

[4] Technically, this means that if $(\mathbf{f},\mathbf{s})$ is optimal for (LLC), then there exists $\mathbf{y}$ such that $(\mathbf{y},\mathbf{s})$ is optimal for the corresponding edge-based formulation, and vice versa.

where $\mathbf{1}$ is the vector of all ones. Let z_{F} be the optimal value of this fully-adjustable formulation. Correspondingly, let z_{A} be the optimal value of its affinely adjustable counterpart, *i.e.*, when we restrict $\mathbf{y}(\cdot)$ and $\mathbf{s}(\cdot)$ to be affine in $\boldsymbol{\xi}$. It suffices then to show that $z_{\mathrm{F}} = z_{\mathrm{A}}$.

We now introduce some useful notation:

- For some index set I, let $\boldsymbol{\xi}_I := \{\xi_i : i \in I\}$ and $\mathrm{proj}_I \Xi := \{\boldsymbol{\xi}_I : \boldsymbol{\xi} \in \Xi\}$.
- Let V^l be the set of nodes at the lth level, *i.e.*, $V^l := \{j \in V : \mathcal{P}^l(j) = 0\}$.
- Let A_i be the set of ancestor nodes of i, *i.e.*, $A_i := \{i, \mathcal{P}(i), \mathcal{P}^2(i), \ldots, 0\}$.
- We will frequently look into subgraphs of (V,T), specifically out-trees rooted at some node $i \in V^l$, denoted by (V_i, T_i), where $V_i := \{j \in V : \mathcal{P}^k(j) = i \text{ for some } k = 0,1,2,\ldots,L-l\}$, and $T_i := \{(k,l) : k \in V_i, l \in \mathcal{C}(k)\}$.
- Let $\mathcal{Q}_i(\xi_i', x_i, \mathbf{x}_{-i})$ by the set of feasible policies for (V_i, T_i), with inventory at node i being x_i, the inventory on $V_i \setminus \{i\}$ being $\mathbf{x}_{-i}$, and the uncertain parameter being ξ_i' at node i, *i.e.*,

$$\begin{aligned}
\mathcal{Q}_i(\xi_i', x_i, \mathbf{x}_{-i}) := \{&\{\mathbf{y}(\cdot), \mathbf{s}(\cdot)\} : \\
& x_i \geq \sum_{j \in \mathcal{C}(i)} y_{ij}(\boldsymbol{\xi}),\ \forall \boldsymbol{\xi} \in \Xi_i \\
& s_i(\boldsymbol{\xi}) + x_i \geq \sum_{j \in \mathcal{C}(i)} y_{ij}(\boldsymbol{\xi}) + \hat{d}_i \xi_i',\ \forall \boldsymbol{\xi} \in \Xi_i \\
& y_{\mathcal{P}(k)k}(\boldsymbol{\xi}) + x_k \geq \sum_{j \in \mathcal{C}(k)} y_{kj}(\boldsymbol{\xi}),\ \forall k \in V_i \setminus \{i\}, \forall \boldsymbol{\xi} \in \Xi_i \\
& s_k(\boldsymbol{\xi}) + y_{\mathcal{P}(k)k}(\boldsymbol{\xi}) + x_k \geq \sum_{j \in \mathcal{C}(k)} y_{kj}(\boldsymbol{\xi}) + \hat{d}_k \xi_k,\ \forall k \in V_i \setminus \{i\}, \forall \boldsymbol{\xi} \in \Xi_i \\
& \mathbf{y}(\boldsymbol{\xi}), \mathbf{s}(\boldsymbol{\xi}) \geq 0,\ \forall \boldsymbol{\xi} \in \Xi_i\}.
\end{aligned}$$

where $\Xi_i(\xi_i') := \{\boldsymbol{\xi} \mid \boldsymbol{\xi} \in \Xi, \xi_i = \xi_i'\}$. We use shorthand notation Ξ_i when there is no ambiguity in ξ_i'.

- Let $\Theta_i(\xi_i', \mathbf{s})$ be the worst-case demand loss in (V_i, T_i) under $\xi_i = \xi_i'$ and (feasible) resource allocation policy $\{\mathbf{y}(\cdot), \mathbf{s}(\cdot)\}$, $\Omega_i^F(\xi_i', x_i, \mathbf{x}_{-i})$ be the worst-case demand loss in V_i under $\xi_i = \xi_i'$ with an optimal fully-adjustable policy, and $\mathcal{Q}_i^\star(\xi_i', x_i, \mathbf{x}_{-i})$ be the set of all optimal policies:

$$\begin{aligned}
\Theta_i(\xi_i', \mathbf{s}) &:= \max_{\boldsymbol{\xi} \in \Xi_i(\xi_i')} \sum_{k \in V_i} s_k(\boldsymbol{\xi}), \\
\Omega_i^F(\xi_i', x_i, \mathbf{x}_{-i}) &:= \min_{\mathbf{y}(\cdot), \mathbf{s}(\cdot)} \Theta_i(\xi_i', \mathbf{s}) \text{ subject to } \{\mathbf{y}(\cdot), \mathbf{s}(\cdot)\} \in \mathcal{Q}_i(\xi_i', x_i, \mathbf{x}_{-i}), \\
\mathcal{Q}_i^\star(\xi_i', x_i, \mathbf{x}_{-i}) &:= \arg\min_{\mathbf{y}(\cdot), \mathbf{s}(\cdot)} \Theta_i(\xi_i', \mathbf{s}) \text{ subject to } \{\mathbf{y}(\cdot), \mathbf{s}(\cdot)\} \in \mathcal{Q}_i(\xi_i', x_i, \mathbf{x}_{-i}).
\end{aligned}$$

Similarly, we define the affinely-adjustable counterpart $\Omega_i^A(\xi_i', x_i, \mathbf{x}_{-i})$ as the worst-case demand loss in V_i, restricted to $\{\mathbf{y}(\cdot), \mathbf{s}(\cdot)\} \in \mathcal{Q}_i(\xi_i', x_i, \mathbf{x}_{-i})$ and affine.

Using our notation, $z_{\mathrm{F}} = \Omega_0^F(1, x_0, \mathbf{x}_{-0})$ and $z_{\mathrm{A}} = \Omega_0^A(1, x_0, \mathbf{x}_{-0})$. It thus suffices to show that

$$\Omega_0^F(1, x_0, \mathbf{x}_{-0}) = \Omega_0^A(1, x_0, \mathbf{x}_{-0}). \tag{9}$$

We now present two useful intermediate results. The first provides a recursive expression for Ω_i^F, and the second shows that Ξ has binary vertices.

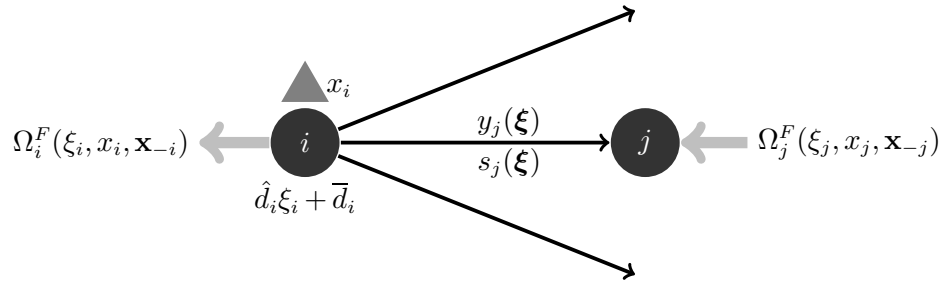

Figure 8 **Recursive representation of Ω_i^F.**

PROPOSITION 1. *For all $i \in V$, $\xi_i' \in [0,1]$, $\mathbf{x} \geq 0$,*

$$\Omega_i^F(\xi_i', x_i, \mathbf{x}_{-i}) = \left(\hat{d}_i \xi_i' - x_i + \max_{\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)} \Xi_i(\xi_i')} \sum_{j \in \mathcal{C}(i)} \Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j}) \right)^+ .$$

Moreover, $\Omega_i^F(\xi_i, x_i, \mathbf{x}_{-i})$ is convex in ξ_i over $\xi_i \in [0,1]$.

PROPOSITION 2. *Ξ is a polytope with binary vertices,* i.e., *$\boldsymbol{\xi} \in ext(\Xi) \implies \boldsymbol{\xi} \in \{0,1\}^n$.*

Proposition 1 shows the intuitive fact that given a fixed inventory decision and under the optimal policy, the demand shortage in V_i should be the sum of shortages in $\{V_j\}_{j \in \mathcal{C}(i)}$ plus the demand on node i, subtracting away inventory x_i. Based on such recursive definition, we can view the tree as a collection of hub-and-spoke clusters, each consisting of one node and several edges pointing away from this node (Figure 8).

Now we prove (9) by induction. For a given tree (V_i, T_i) and uncertainty set Ξ_i, we limit our attention to $\{\mathbf{y}(\cdot), \mathbf{s}(\cdot)\}$ that satisfies the following properties:

(I) $\forall m \in V_i$, $j \in \mathcal{C}(m)$, $\boldsymbol{\xi} \in \Xi_i : s_m(\boldsymbol{\xi}) = s_m(\xi_m)$, $y_{mj}(\boldsymbol{\xi}) = y_{mj}(\xi_j)$, *i.e.*, the policies are univariate in ξ_m, ξ_j respectively.

(II) $\forall m \in V_i$: (a) $y_{\mathcal{P}(m)m}(0) = 0$, (b) $y_{jk}(0) = \Omega_k^A(0, x_k, \mathbf{x}_{-k}) = 0 \, \forall (j,k) \in T_m$, and (c) $s_j(0) = 0 \, \forall j \in V_m$, *i.e.*, these policies are linear.

Let

$$\overline{\mathcal{Q}}_i(\xi_i, x_i, \mathbf{x}_{-i}) := \{\{\mathbf{y}, \mathbf{s}\} \in \mathcal{Q}_i(\xi_i, x_i, \mathbf{x}_{-i}) \, : \, \{\mathbf{y}, \mathbf{s}\} \text{ satisfies (I)-(II)}\}$$
$$\overline{\mathcal{Q}}_i^\star(\xi_i, x_i, \mathbf{x}_{-i}) := \overline{\mathcal{Q}}_i(\xi_i, x_i, \mathbf{x}_{-i}) \cap \mathcal{Q}_i^\star(\xi_i, x_i, \mathbf{x}_{-i}).$$

We are now ready to formalize the induction process.

Induction Hypothesis. If for some $l \in \{0, 1, \ldots, L-1\}$, $\forall j \in V^{l+1}, \mathbf{x} \geq 0$, $\exists \{\bar{\mathbf{y}}^j(\cdot), \bar{\mathbf{s}}^j(\cdot)\}$ such that

$$\text{(Feasibility) } \forall \xi_j \in [0,1], \{\bar{\mathbf{y}}^j(\cdot), \bar{\mathbf{s}}^j(\cdot)\} \in \overline{\mathcal{Q}}_j(\xi_j, x_j, \mathbf{x}_{-j}), \tag{10}$$

$$\text{(Optimality) } \forall \xi_j \in \{0,1\}, \Theta_j(\xi_j, \bar{\mathbf{s}}^j) = \Omega_j^A(\xi_j, x_j, \mathbf{x}_{-j}) = \Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j}), \tag{11}$$

then, $\forall i \in V^l, \mathbf{x} \geq 0$, $\exists \{\bar{\mathbf{y}}^i(\cdot), \bar{\mathbf{s}}^i(\cdot)\}$ such that

$$\text{(Feasibility) } \forall \xi_i \in [0,1], \{\bar{\mathbf{y}}^i(\cdot), \bar{\mathbf{s}}^i(\cdot)\} \in \overline{\mathcal{Q}}_i(\xi_i, x_i, \mathbf{x}_{-i}), \tag{12}$$

$$\text{(Optimality) } \forall \xi_i \in \{0,1\}, \Theta_i(\xi_i, \bar{\mathbf{s}}^i) = \Omega_i^A(\xi_i, x_i, \mathbf{x}_{-i}) = \Omega_i^F(\xi_i, x_i, \mathbf{x}_{-i}). \tag{13}$$

Base Case ($l = L-1$). Given some $\mathbf{x} \geq 0$, for each $j \in V^L$ we define $\bar{s}^j_j(\boldsymbol{\xi}) = \bar{s}^j_j(\xi_j) = \xi_j(\hat{d}_j - x_j)^+$. This policy satisfies (I) and (II) by construction, and it is straightforward to check feasibility and optimality.

General Step. Now we prove the general induction step: suppose (10)-(11) hold for some $l \in \{0, 1, \ldots, L-1\}$ and $\forall j \in V^{l+1}, \mathbf{x} \geq 0$. To construct $\{\bar{\mathbf{y}}^i(\cdot), \bar{\mathbf{s}}^i(\cdot)\}$, we define the following. Since $\Omega^A_j(\xi_j, x_j, \mathbf{x}_{-j})$ is nondecreasing in ξ_j and nonnegative, there necessarily exists $g_j \geq 0$ such that $\Omega^A_j(\xi_j, x_j, \mathbf{x}_{-j}) = g_j\xi_j$ for $\xi_j \in \{0,1\}$. Let $\beta_i := \min\left(1, \frac{x_i}{\sum_{j=1}^{\Gamma_i} g_{(j)} + \hat{d}_i}\right)$ if $\sum_{j=1}^{\Gamma_i} g_{(j)} + \hat{d}_i > 0$, and $\beta_i := 1$ otherwise, where $g_{(j)}$ is the jth largest element of $\{g_j\}_{j \in \mathcal{C}(i)}$. ($\beta_i$ is the surge demand coverage ratio)

We are now ready to construct affine $\{\bar{\mathbf{y}}^i, \bar{\mathbf{s}}^i\}$:

$$\{\bar{\mathbf{y}}^i, \bar{\mathbf{s}}^i\} := \begin{cases} \bar{y}^i_{ij}(\xi_j) = \beta_i g_j \xi_j & \forall j \in \mathcal{C}(i) \\ \bar{s}^i_i(\xi_i) = (1-\beta_i)\hat{d}_i\xi_i & \\ \bar{y}^i_{km}(\xi_m) = \bar{y}^{j1}_{km}(\xi_m) & \forall (k,m) \in T_j, \forall j \in \mathcal{C}(i) \\ \bar{s}^i_k(\xi_k) = \bar{s}^{j1}_k(1)\xi_k & \forall k \in V_i \setminus \{i\}. \end{cases} \tag{14}$$

Let $\overline{\mathcal{Q}}^\star_{j1} := \overline{\mathcal{Q}}^\star_j(1, x_j + \bar{y}^i_{ij}(1), \mathbf{x}_{-j})$. By the induction hypothesis, $\overline{\mathcal{Q}}^\star_{j1} \neq \emptyset$, since $x_j + \bar{y}^i_{ij}(\xi_j) \geq 0$ for $\xi_j \in [0,1]$. We now check the feasibility and optimality of $\{\bar{\mathbf{y}}^i, \bar{\mathbf{s}}^i\}$.

For (Feasibility), note that Property (I-II) hold by construction. To show $\{\bar{\mathbf{y}}^i, \bar{\mathbf{s}}^i\}$ satisfies the constraints in $\mathcal{Q}_i(\xi'_i, x_i, \mathbf{x}_{-i})$, note that the constraints involving node i are satisfied by construction of $\{\bar{\mathbf{y}}^i, \bar{\mathbf{s}}^i\}$. For nodes downstream from i, note that $\bar{y}^i_{km}(\xi_m) = \bar{y}^{j1}_{km}(1)\xi_m$ and $\bar{s}^i_k(\xi_k) = \bar{s}^{j1}_k(1)\xi_k$, for all $(k,m) \in T_j, j \in \mathcal{C}(i)$ and $k \in V_i \setminus \{i\}$. Since $\{\bar{\mathbf{y}}^{j0}(\cdot), \bar{\mathbf{s}}^{j0}(\cdot)\}$ and $\{\bar{\mathbf{y}}^{j1}(\cdot), \bar{\mathbf{s}}^{j1}(\cdot)\}$ satisfy the constraints of $\mathcal{Q}_j(0, x_j + \bar{y}^i_{ij}(0), \mathbf{x}_{-j})$ and $\mathcal{Q}_j(1, x_j + \bar{y}^i_{ij}(1), \mathbf{x}_{-j})$ respectively, it is straightforward to check the constraint satisfaction of $\{\mathbf{y}^i, \mathbf{s}^i\}$ for $\mathcal{Q}_i(\xi'_i, x_i, \mathbf{x}_{-i})$.

For (Optimality), note that for $\xi'_i \in \{0,1\}$ we have

$$\begin{aligned}
\Theta_i(\xi'_i, \bar{\mathbf{s}}^i) =& \bar{s}^i_i(\xi'_i) + \max_{\boldsymbol{\xi} \in \Xi_i} \sum_{j \in \mathcal{C}(i)} \sum_{k \in V_j} \bar{s}^i_k(\xi_k) && \text{(15a)} \\
=& \bar{s}^i_i(\xi'_i) + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi'_i \\ \xi_j \in [0, \xi'_i], j \in \mathcal{C}(i)}} \sum_{j \in \mathcal{C}(i)} \max_{\boldsymbol{\xi} \in \Xi_j(\xi_j)} \sum_{k \in V_j} \bar{s}^{j1}_k(1)\xi_k && \text{(15b)} \\
=& \bar{s}^i_i(\xi'_i) + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi'_i \\ \xi_j \in \{0, \xi'_i\}, j \in \mathcal{C}(i)}} \sum_{j \in \mathcal{C}(i)} \max_{\boldsymbol{\xi} \in \Xi_j(\xi_j)} \sum_{k \in V_j} \bar{s}^{j1}_k(1)\xi_k && \text{(15c)} \\
=& \bar{s}^i_i(\xi'_i) + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi'_i \\ \xi_j \in \{0, \xi'_i\}, j \in \mathcal{C}(i)}} \sum_{j \in \mathcal{C}(i)} \Omega^A_j(\xi_j, x_j + \bar{y}^i_{ij}(\xi_j), \mathbf{x}_{-j}) && \text{(15d)} \\
=& \bar{s}^i_i(\xi'_i) + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi'_i \\ \xi_j \in \{0, \xi'_i\}, j \in \mathcal{C}(i)}} \sum_{j \in \mathcal{C}(i)} \Omega^F_j(\xi_j, x_j + \bar{y}^i_{ij}(\xi_j), \mathbf{x}_{-j}) && \text{(15e)} \\
=& \bar{s}^i_i(\xi'_i) + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi'_i \\ \xi_j \in \{0, \xi'_i\}, j \in \mathcal{C}(i)}} \sum_{j \in \mathcal{C}(i)} \left(\Omega^F_j(\xi_j, x_j, \mathbf{x}_{-j}) - \bar{y}^i_{ij}(\xi_j)\right)^+ && \text{(15f)} \\
=& \bar{s}^i_i(\xi'_i) + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi'_i \\ \xi_j \in \{0, \xi'_i\}, j \in \mathcal{C}(i)}} \left(\sum_{j \in \mathcal{C}(i)} \Omega^F_j(\xi_j, x_j, \mathbf{x}_{-j}) - \sum_{j \in \mathcal{C}(i)} \bar{y}^i_{ij}(\xi_j)\right) && \text{(15g)} \\
=& \left(\hat{d}_i\xi'_i - x_i + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi'_i \\ \xi_j \in \{0, \xi'_i\}, j \in \mathcal{C}(i)}} \sum_{j \in \mathcal{C}(i)} \Omega^F_j(\xi_j, x_j, \mathbf{x}_{-j})\right)^+ && \text{(15h)}
\end{aligned}$$

$$= \left(\hat{d}_i \xi_i' - x_i + \max_{\substack{\boldsymbol{\xi}_{\mathcal{C}(i)} : \mathbf{1}^T \boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i \xi_i' \\ \xi_j \in [0, \xi_i'], j \in \mathcal{C}(i)}} \sum_{j \in \mathcal{C}(i)} \Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j}) \right)^+ \tag{15i}$$

$$= \Omega_i^F(\xi_i', x_i, \mathbf{x}_{-i}) = \Omega_i^A(\xi_i', x_i, \mathbf{x}_{-i}). \tag{15j}$$

Equalities (15a-15b) are straightforward applications of definitions of Θ_i and $\bar{\mathbf{s}}^i$.

To show equality (15c), we argue that is it sufficient to check $\xi_j = 1$ and $\xi_j = 0$ for $j \in \mathcal{C}(i)$. Suppose that $\Xi_j(\xi_j)$ has n extreme points $\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^n$, and $W_m = \{k \in V_j : \xi_k^m = \xi_j\}$, $\overline{W}_m = V_j \setminus W_m$. Since for each extreme point, ξ_k takes value of either ξ_j or 0 (*cf.* Proposition 2), we can write

$$\max_{\boldsymbol{\xi} \in \Xi_j(\xi_j)} \sum_{k \in V_j} \bar{s}_k^{j1}(1) \xi_k = \max_{m \in \{1, \ldots, n\}} \sum_{k \in V_j} \left\{ \bar{s}_k^{j1}(1) \xi_k \mathbb{1}_{k \in W_m} \right\},$$

which is a convex function of ξ_j ($\mathbb{1}_{i \in I} = 1$ if $i \in I$ and 0 otherwise). By this token, we can replace $\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)} \Xi_i$ with $\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)} \Xi_i \cap \{0,1\}^{|\mathcal{C}(i)|}$ for the argument of the outer maximization, since $\Xi_i(0)$ and $\Xi_i(1)$ have binary vertices.

To show equality (15d), note that for $\xi_j = 1$, $\Omega_j^A(1, x_j + \bar{y}_{ij}^i(1), \mathbf{x}_{-j}) = \Theta_j(1, \bar{\mathbf{s}}^{j1}) = \bar{s}_j^{j1}(1) + \max_{\boldsymbol{\xi} \in \Xi_j(1)} \sum_{k \in V_j \setminus \{j\}} \bar{s}_k^{j1}(\xi_k)$, which is equal to $\max_{\boldsymbol{\xi} \in \Xi_j(1)} \sum_{k \in V_j} \bar{s}_k^{j1}(1) \xi_k$. For $\xi_j = 0$, $\Omega_j^A(0, x_j + \bar{y}_{ij}^i(0), \mathbf{x}_{-j}) = 0$. Equality (15e) holds by the induction hypothesis. Equality (15f) follows from the definition of Ω_j^F and (19). For equality (15g), by the definitions of β_i, we have $(\Omega_j^A(\xi_j, x_j, \mathbf{x}_{-j}) - \bar{y}_{ij}^i(\xi_j)) \geq 0$ for $j \in \mathcal{C}(i)$, therefore we can rewrite (15f) into (15g). For the transition into (15h), it is straightforward to consider $\xi_i' = 0$ and $\xi_i' = 1$ separately. Lastly, (15i) holds by the convexity property of $\sum_{j \in \mathcal{C}(i)} \Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j})$ and (15j) holds by the recursive property, both shown in Proposition 1. We have now completed the proof for the inductive step, thus showing $z_{\text{F}} = z_{\text{A}}$.

To complete the proof, we now turn our attention to the (LLG) formulation. By solving for $s_i(\mathbf{d})$ from (1c) and substituting in (2b), we can eliminate (1c) and re-write (2b) as

$$\sum_{j:(j,i) \in E} f_{ji}(\mathbf{d}) \geq \frac{1 - \epsilon_i - \overline{\rho}}{\rho - \overline{\rho}} d_i, \quad \forall i \in V, \forall \mathbf{d} \in U.$$

At optimality, the above constraint can be taken to be active without loss (if it is not, we can scale down the associated flows into i so that it becomes active). Thus, the (LLG) is equivalent to

$$\begin{aligned}
\min_{\mathbf{x}, \mathbf{f}(\cdot), \mathbf{s}(\cdot)} \quad & \mathbf{h}^T \mathbf{x} \\
\text{subject to} \quad & x_i \geq \sum_{j:(i,j) \in E} f_{ij}(\mathbf{d}), \quad \forall i \in V, \forall \mathbf{d} \in U' \\
& s_i(\mathbf{d}) + \sum_{j:(j,i) \in E} f_{ji}(\mathbf{d}) = d_i, \quad \forall i \in V, \forall \mathbf{d} \in U' \\
& s_i(\mathbf{d}) = 0, \quad \forall i \in V, \forall \mathbf{d} \in U' \\
& f_{ij}(\mathbf{d}) \geq 0, \quad \forall (i,j) \in E, \forall \mathbf{d} \in U' \\
& \mathbf{x} \in X,
\end{aligned}$$

where $U' := \left\{\mathbf{d}' : d'_i = \frac{1-\epsilon_i-\overline{\rho}}{\rho-\overline{\rho}} d_i \, \forall i \in V, \mathbf{d} \in U\right\}$. Then, this is in turn equivalent to

$$\begin{aligned}
\min_{\mathbf{x}, \mathbf{f}(\cdot), \mathbf{s}(\cdot)} \quad \mathbf{h}^T\mathbf{x} \quad + \quad \max_{\mathbf{d}\in U'} \quad & b' \sum_{i\in V} (1-\overline{\rho}) s_i(\mathbf{d}) \\
\text{subject to} \quad & x_i \geq \sum_{j:(i,j)\in E} f_{ij}(\mathbf{d}), \quad \forall i \in V, \forall \mathbf{d} \in U' \\
& s_i(\mathbf{d}) + \sum_{j:(j,i)\in E} f_{ji}(\mathbf{d}) = d_i, \quad \forall i \in V, \forall \mathbf{d} \in U' \\
& \sum_{j:(j,i)\in E} f_{ji}(\mathbf{d}) \leq d_i, \quad \forall i \in V, \forall \mathbf{d} \in U' \\
& f_{ij}(\mathbf{d}) \geq 0, \quad \forall (i,j) \in E, \forall \mathbf{d} \in U' \\
& \mathbf{x} \in X
\end{aligned}$$

for b' large enough. However, this corresponds to an (LLC) instance that satisfies our assumptions and therefore admits an optimal affine adjustable policy. $\square$

Proof of Proposition 1. Note that we can express $\Omega_i^F(\xi'_i, x_i, \mathbf{x}_{-i})$ as $\max_{\boldsymbol{\xi}\in\Xi_i(\xi'_i)} r_i(\boldsymbol{\xi}, x_i, \mathbf{x}_{-i})$, where

$$\begin{aligned}
r_i(\boldsymbol{\xi}, x_i, \mathbf{x}_{-i}) := \min_{\mathbf{y},\mathbf{s}} \quad & \sum_{k\in V_i} s_k \\
\text{subject to} \quad & x_i \geq \sum_{j\in\mathcal{C}(i)} y_{ij} \\
& s_i + x_i \geq \sum_{j\in\mathcal{C}(i)} y_{ij} + \hat{d}_i \xi_i \\
& y_{\mathcal{P}(k)k} + x_k \geq \sum_{j\in\mathcal{C}(k)} y_{kj}, \; \forall k \in V_i \setminus \{i\} \\
& y_{\mathcal{P}(k)k} + s_k + x_k \geq \sum_{j\in\mathcal{C}(k)} y_{kj} + \hat{d}_k \xi_k, \; \forall k \in V_i \setminus \{i\} \\
& \mathbf{y}, \mathbf{s} \geq 0.
\end{aligned}$$

It is straightforward to check that the above optimization problem admits an optimal solution such that $s_i = (\hat{d}_i\xi'_i - x_i)^+$ and $\sum_{j\in\mathcal{C}(i)} y_{ij} = (x_i - \hat{d}_i\xi'_i)^+$. Also,

$$r_i(\boldsymbol{\xi}, x_i + \Delta, \mathbf{x}_{-i}) = (r_i(\boldsymbol{\xi}, x_i, \mathbf{x}_{-i}) - \Delta)^+, \quad \forall \Delta \geq 0. \tag{19}$$

Another fact we will use is that

$$\min_{\substack{\mathbf{x}\geq 0 \\ \mathbf{1}^T\mathbf{x} = K}} \sum_{i=1}^n (q_i - x_i)^+ = (\mathbf{1}^T\mathbf{q} - K)^+, \quad \forall \mathbf{q} \in \mathbb{R}^n, K \geq 0. \tag{20}$$

Using all these properties we have

$$\begin{aligned}
\Omega_i^F(\xi'_i, x_i, \mathbf{x}_{-i}) =& (\hat{d}_i\xi'_i - x_i)^+ + \max_{\boldsymbol{\xi}\in\Xi_i} \min_{\substack{\{y_{ij}\}_{j\in\mathcal{C}(i)} : y_{ij}\geq 0, \\ \sum_{j\in\mathcal{C}(i)} y_{ij} = (x_i - \hat{d}_i\xi'_i)^+}} \sum_{j\in\mathcal{C}(i)} r_j(\boldsymbol{\xi}, x_j + y_{ij}, \mathbf{x}_{-j}) && \text{(21a)} \\
=& (\hat{d}_i\xi'_i - x_i)^+ + \max_{\boldsymbol{\xi}\in\Xi_i} \min_{\substack{\{y_{ij}\}_{j\in\mathcal{C}(i)} : y_{ij}\geq 0, \\ \sum_{j\in\mathcal{C}(i)} y_{ij} = (x_i - \hat{d}_i\xi'_i)^+}} \sum_{j\in\mathcal{C}(i)} (r_j(\boldsymbol{\xi}, x_j, \mathbf{x}_{-j}) - y_{ij})^+ && \text{(21b)} \\
=& (\hat{d}_i\xi'_i - x_i)^+ + \left(\max_{\boldsymbol{\xi}\in\Xi_i} \sum_{j\in\mathcal{C}(i)} r_j(\boldsymbol{\xi}, x_j, \mathbf{x}_{-j}) - (x_i - \hat{d}_i\xi'_i)^+ \right)^+ && \text{(21c)}
\end{aligned}$$

$$=(\hat{d}_i\xi_i' - x_i)^+ + \left(\max_{\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)}\Xi_i} \max_{\substack{\boldsymbol{\xi}\in\Xi_j(\xi_j),\\ j\in\mathcal{C}(i)}} \sum_{j\in\mathcal{C}(i)} r_j(\boldsymbol{\xi}, x_j, \mathbf{x}_{-j}) - (x_i - \hat{d}_i\xi_i')^+ \right)^+ \tag{21d}$$

$$=(\hat{d}_i\xi_i' - x_i)^+ + \left(\max_{\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)}\Xi_i} \sum_{j\in\mathcal{C}(i)} \max_{\boldsymbol{\xi}\in\Xi_j(\xi_j)} r_j(\boldsymbol{\xi}, x_j, \mathbf{x}_{-j}) - (x_i - \hat{d}_i\xi_i')^+ \right)^+ \tag{21e}$$

$$=(\hat{d}_i\xi_i' - x_i)^+ + \left(\max_{\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)}\Xi_i} \sum_{j\in\mathcal{C}(i)} \Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j}) - (x_i - \hat{d}_i\xi_i')^+ \right)^+ \tag{21f}$$

$$= \left(\hat{d}_i\xi_i' - x_i + \max_{\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)}\Xi_i} \sum_{j\in\mathcal{C}(i)} \Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j}) \right)^+ . \tag{21g}$$

Equality (21a) follows from the definition of r_i; (21b) follows from (19); (21c) follows from (20); (21d) is an equivalent way of writing the maximization operator; exchanging the inner maximization and summation operators is allowed since r_j only depends on $\{\xi_k, k \in V_j\}$, leading to (21e); (21f) follows from the definition of r_j. Equality (21g) holds by considering the fact that $\Omega_j^F \geq 0$ for any $j \in \mathcal{C}(i)$, and checking two cases: one with $\hat{d}_i\xi_i' - x_i < 0$, and the other with $\hat{d}_i\xi_i' - x_i \geq 0$.

To show convexity of $\Omega_i^F(\xi_i', x_i, \mathbf{x}_{-i})$, we use an induction argument on the level of node i. First note that for $i \in V^L$, we have $\Omega_i^F(\xi_i', x_i, \mathbf{x}_{-i}) = (\hat{d}_i\xi_i' - x_i)^+$, which is convex and non-decreasing in ξ_i'. For i in any other level of the network, given the result we just proved (21g), it suffices to show $\max_{\boldsymbol{\xi}_{\mathcal{C}(i)} \in \text{proj}_{\mathcal{C}(i)}\Xi_i(\xi_i')} \sum_{j\in\mathcal{C}(i)} \Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j})$ is convex in ξ_i'. Since $\Omega_j^F(\xi_j, x_j, \mathbf{x}_{-j})$ is convex in ξ_j (by the induction hypothesis), maximum is obtained at the extreme points. Given that $\text{proj}_{\mathcal{C}(i)}\Xi_i(\xi_i') = \{\boldsymbol{\xi}_{\mathcal{C}(i)} : \xi_j \in [0, \xi_i'], \mathbf{1}^T\boldsymbol{\xi}_{\mathcal{C}(i)} \leq \Gamma_i\xi_i'\}$ and Ω_j^F is non-decreasing in ξ_j, we can equivalently write: $\max_{|S|=\Gamma_i, S\subseteq\mathcal{C}(i)} \left\{ \sum_{j\in S} \Omega_j^F(\xi_i', x_j, \mathbf{x}_{-j}) + \sum_{j\in\mathcal{C}(i)\setminus S} \Omega_j^F(0, x_j, \mathbf{x}_{-j}) \right\}$, which is convex, non-decreasing in ξ_i'. □

Proof of Proposition 2. For the sake of reaching a contradiction, suppose there exists some extreme point $\boldsymbol{\xi} \in \text{ext}(\Xi)$, such that its kth entry is fractional, $\xi_k \in (0,1)$. Note that by definition of Ξ, $\xi_0 = 1$, and thus $k \neq 0$. Let $p = \mathcal{P}(k)$ be its parent, *i.e.*, $k \in \mathcal{C}(p)$. Then it must hold that $0 < \sum_{i\in\mathcal{C}(p)} \xi_i \leq \Gamma_p\xi_p$, and thus $\Gamma_p > 0$ and $\xi_p > 0$.

Suppose $\xi_p = 1$. Then the above condition becomes $0 < \sum_{i\in\mathcal{C}(p)} \xi_i \leq \Gamma_p$. Furthermore, $\xi_i \in [0,1]\ \forall i \in \mathcal{C}(p)$. Now consider the $|\mathcal{C}(p)|$-dimensional polytope $\{\boldsymbol{\gamma} : \sum_{i\in\mathcal{C}(p)} \gamma_i \leq \Gamma_p, \gamma_i \in [0,1]\, \forall i \in \mathcal{C}(p)\}$, which can be readily seen to have binary extreme points. Let these extreme points be $\{\boldsymbol{\gamma}^j, j = 1, \ldots, N\}$. Since $\boldsymbol{\xi}_{\mathcal{C}(p)}$ belongs to this polytope, we can express it as

$$\boldsymbol{\xi}_{\mathcal{C}(p)} = \sum_{j=1}^N \alpha_j\boldsymbol{\gamma}^j, \text{ where } \sum_{j=1}^N \alpha_j = 1, \alpha_j \geq 0, \forall j \in \{1, \ldots, N\}.$$

In particular, for those $i \in \mathcal{C}(p)$ that $\xi_i > 0$, we have:

$$\sum_{j=1}^N \alpha_j\gamma_i^j = \xi_i \iff \sum_{j=1}^N \alpha_j \frac{1}{\xi_i}\gamma_i^j = 1. \tag{22}$$

Now we use each $\boldsymbol{\gamma}^j$ to construct a vector $\boldsymbol{\xi}^j \in \Xi$ and show that we can write $\boldsymbol{\xi}$ as a convex combination of $\{\boldsymbol{\xi}^j, j = 1, \ldots, N\}$, and $\boldsymbol{\xi} \neq \boldsymbol{\xi}^j, \forall j \in \{1, \ldots, N\}$—which will contradict our assumptions.

We construct $\boldsymbol{\xi}^j$, $j=1,\ldots,N$, as follows:

$$\begin{aligned}
\boldsymbol{\xi}^j_{V\setminus V_p} &= \boldsymbol{\xi}_{V\setminus V_p},\\
\xi^j_p &= \xi_p,\\
\boldsymbol{\xi}^j_{\mathcal{C}(p)} &= \boldsymbol{\gamma}^j,\\
\boldsymbol{\xi}^j_{V_i\setminus\{i\}} &= \begin{cases} \frac{1}{\xi_i}\boldsymbol{\xi}_{V_i\setminus\{i\}} & \text{if } \xi_i>0 \text{ and } \xi^j_i=1\\ 0 & \text{otherwise}\end{cases}, \quad \forall i\in\mathcal{C}(p).
\end{aligned}$$

It is straightforward to verify that for all $j=1,\ldots,N$, $\boldsymbol{\xi}^j\in\Xi$ and $\boldsymbol{\xi}^j\neq\boldsymbol{\xi}$. We next argue that $\sum_{j=1,\ldots,N}\alpha_j\boldsymbol{\xi}^j=\boldsymbol{\xi}$ by checking component-wise. This is straightforward for components $V\setminus V_p$, p, and $\mathcal{C}(p)$. For any $l\in V_i\setminus\{i\}$ for $i\in\mathcal{C}(p)$ such that $\xi_i=0$, we have that $\xi^j_l=\xi_l=0$ for all $j=1,\ldots,N$. Finally, for any $l\in V_i\setminus\{i\}$ for $i\in\mathcal{C}(p)$ such that $\xi_i>0$, we have that

$$\sum_{j=1}^N\alpha_j\xi^j_l=\sum_{j:\xi^j_i=1}\alpha_j\xi_l/\xi_i+\sum_{j:\xi^j_i=0}0=\sum_{j=1}^N\alpha_j\xi_l/\xi_i\xi^j_i=\sum_{j=1}^N\left(\alpha_j\xi^j_i/\xi_i\right)\xi_l\overset{(22)}{=}\sum_{j=1}^N\left(\alpha_j\gamma^j_i/\xi_i\right)\xi_l\overset{(22)}{=}\xi_l.$$

Consequently, $\boldsymbol{\xi}$ cannot be an extreme point if $\xi_p=1$. Since $\xi_p>0$, we must have $\xi_p\in(0,1)$. We can propagate this argument upstream to eventually show that $\xi_0\in(0,1)$, which contradicts that $\xi_0=1$ for all $\boldsymbol{\xi}\in\Xi$, and the proof is complete. □

Appendix B: Numerical Studies on AP Heuristic

B.1. Relaxing Assumptions 1 and 2

We provide details on how problem instances are generated for Studies 1 and 2 in §4.2, alongside a quantitative analysis of our results. For each problem instance, we generate a tree-style graph as follows:

1. Uniformly sample the target total number of nodes, $n\in\{100,200,500,1000,2000,5000,10000\}$.
2. Uniformly sample the number of levels, $l+1\in\{2,3,4,5\}$.
3. For each non-leaf level $k=0,1,\ldots,l-1$, sample a random integer f_k between 5 and 15.
4. Find F (*e.g.*, by line search) such that $\texttt{round}(1+f_0/F+f_0f_1/F^2+\cdots+f_0\cdots f_{l-1}/F^{l-1})=n$, where $\texttt{round}(\cdot)$ is the usual rounding operation.
5. Grow the tree from root in a breadth-first way: for each node in level $k=0,1,\ldots,l-1$, assign to it c_i children, where c_i is randomly picked from $\{(\prod_{j=0}^k f_j/F^k-1)^+,\prod_{j=0}^k f_j/F^k,\prod_{j=0}^k f_j/F^k+1\}$.

With this procedure, we generate 5000 trees. For Study 2, for each of the 5000 trees, we create a non-tree counterpart by introducing additional edges into the graph. Specifically, for each pair of nodes that are in adjacent levels in the previously constructed tree, we assign them an edge with probability p_{arc}, which is randomly sampled from $[0,0.01]$ for each instance. The total number of edges in the non-tree counterpart comes out to have 0% to 200% more edges than its base tree.

To generate cost, efficacy, demand, and attack severity parameter values, we first normalize the per unit demand loss cost, $b=100$. Then for each instance, we do the following:

1. Generate (float-valued) inventory cost upper and lower bounds uniformly: $\underline{h}\in[0,10]$, and $\bar{h}\in[10,20]$, and $\delta h=\bar{h}-\underline{h}$.

2. For each node i, assign an inventory cost $h_i \in [\underline{h}, \bar{h}]$ to it. More specifically, h_i is uniformly sampled from $[\underline{h} + \frac{l(i)+0.5}{l+1} \times \delta h, \underline{h} + \frac{l(i)+1}{l+1} \times \delta h]$, where $l(i) \in \{0, 1, \ldots, l\}$ is the level in which node i resides.
3. Uniformly sample the average efficacy $\bar{\rho} \in [0, 1]$. For each path p, generate $\rho_p \in [0.5, 1] \times \bar{\rho}$. Assign a flow cost efficacy $100\rho_p$ to the path.
4. For each demand node i, the maximum demand is $\hat{d}_i$, which is uniformly sampled from $[0, 1000]$. (This value is then fed into the optmization formulation $d_i(\xi_i) = \hat{d}_i \xi_i$.)
5. To determine the attack scale for each node, Γ_i, first generate $\bar{\Gamma} \in [0, 0.2]$. Then for each node i in this network instance, uniformly sample an integer Γ_i between $\texttt{round}(0.8\bar{\Gamma}c_i)$ and $\texttt{round}(1.2\bar{\Gamma}c_i)$.

Figures 9-14 depict, either through scatter plots or standard box plots, the upper bounds on the AP heuristic suboptimality gaps we obtained for Studies 1 & 2 for varying number of nodes, tree depth, inventory cost parameters, attack severity parameters, antibiotic efficacy, and tree-violating edges. It can be seen that suboptimality gaps are small and do not appear to have strong correlation with any of the parameter values tested, including, the number of nodes and the attack severity.

B.2. Relaxing Infinite Capacity Assumption

To explore the AP heuristic's performance for the joint inventory and dispensing capacity optimization problem, (LLGC) in §5, we conducted numerical studies similar to Studies 1 & 2 described above.

Recall that (LLGC) differed from (LLG) by introducing capacity constraints (4e) for subsets of edges in the graph, and the associated capacity variables and costs. We conducted two studies, in both of which we considered the same setup as in Study 1. In the first of these two new studies, we generated these subsets of edges in a random fashion. In the second, we considered subsets of edges that had similar structure as in (4e)—specifically, sets of edges which all shared the same demand node. Both studies yielded similar results; we focus only on the latter in the remainder.

To introduce dispensing capacity, for each demand node, we partitioned its incoming edges to k equi-sized sets; k was sampled from $\{2, 4, 6, 8\}$. For each resulting set of edges, we introduced a capacity variable and a constraint, as in (4e). The associated capacity cost coefficient was sampled as a fraction of the life loss cost.

Table 5 provides statistics of the upper bounds on the suboptimality gaps we obtained. Figure 15 depicts the

Min	1st Quantile	Median	Mean	3rd Quantile	Max
0.00%	0.00%	0.00%	0.15%	0.00%	0.22%

Table 5 Statistics of upper bounds on AP heuristic's suboptimality gaps for the numerical studies on the joint inventory and dispensing capacity optimization problem.

suboptimality gap bounds for varying number of nodes. As before, our analysis reveals that the suboptimality gap is small and does not grow with the size of the network.

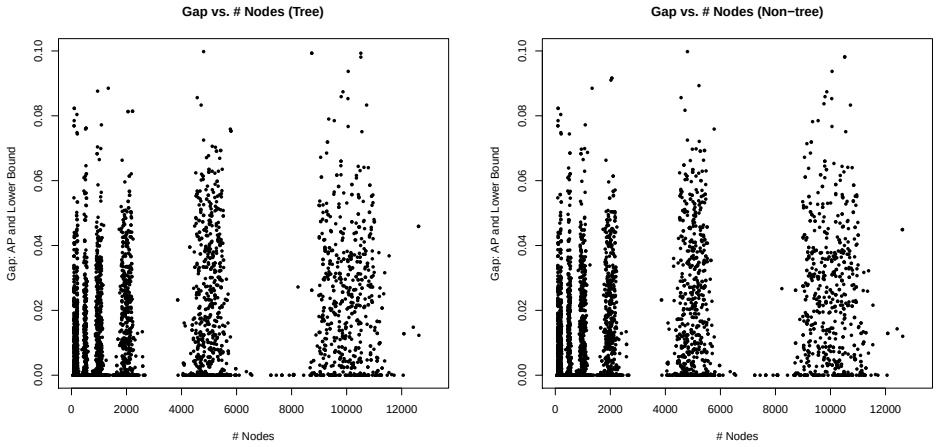

Figure 9 **Upper bounds on AP heuristic's suboptimality gap for varying number of nodes.**

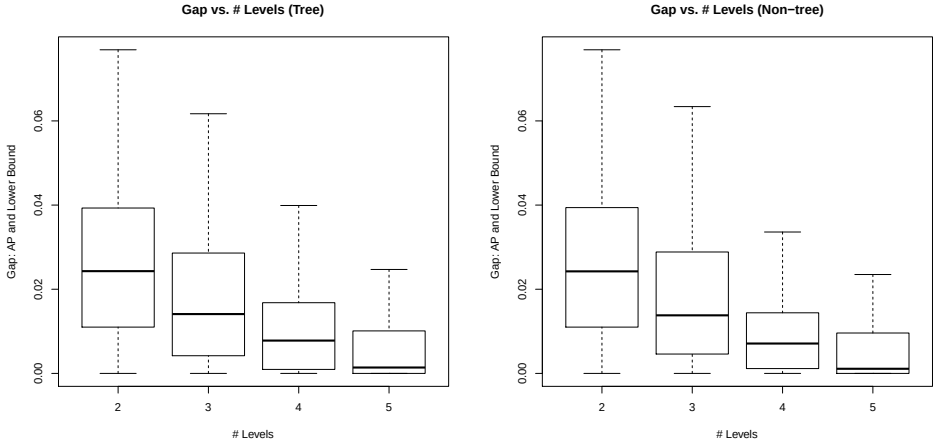

Figure 10 **Upper bounds on AP heuristic's suboptimality gap for varying number of levels in the graph.**

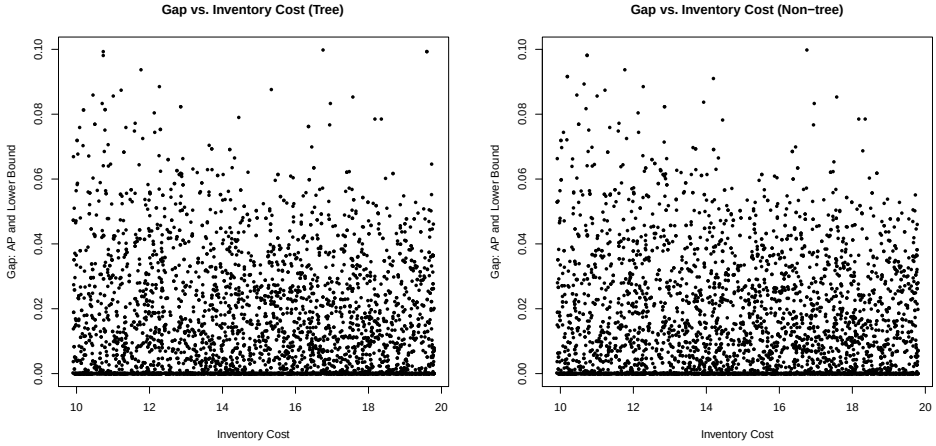

Figure 11 **Upper bounds on AP heuristic's suboptimality gap for varying inventory cost parameter, $\bar{h}$.**

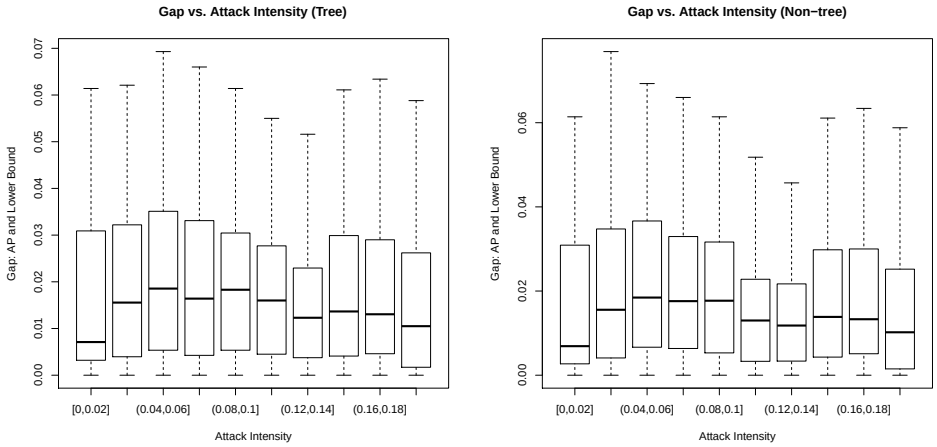

Figure 12 **Upper bounds on AP heuristic's suboptimality gap for varying attack severity, $\bar{\Gamma}$.**

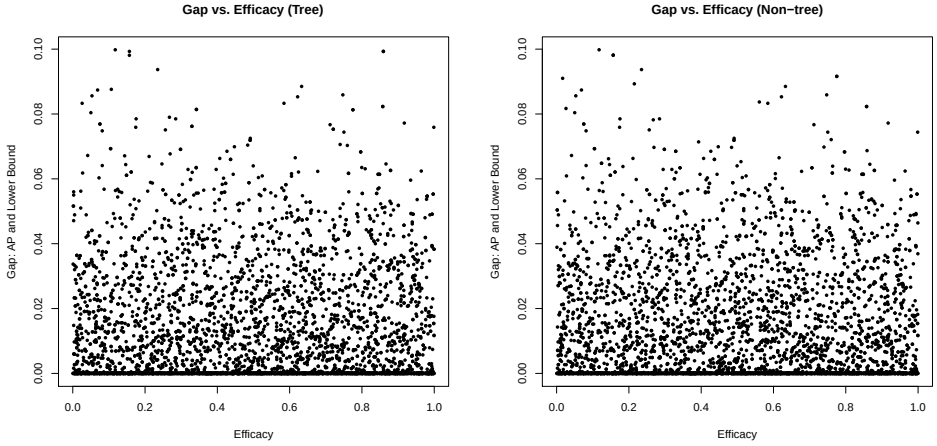

Figure 13 **Upper bounds on AP heuristic's suboptimality gap for varying antibiotic efficacy, $\bar{\rho}$.**

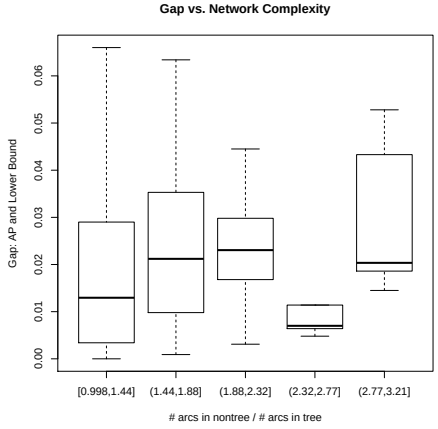

Figure 14 **Upper bounds on AP heuristic's suboptimality gap for varying network complexity.**

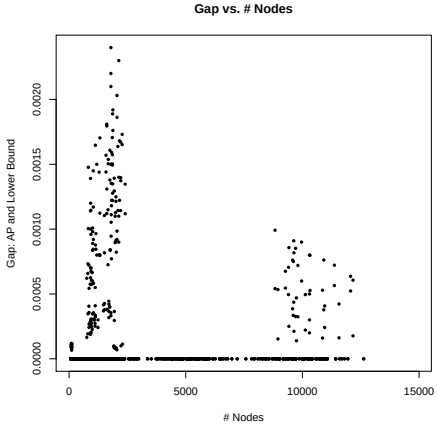

Figure 15 **Upper bounds on AP heuristic's suboptimality gap for varying number of nodes, capacitated networks.**