Optimization Methods for Disease Prevention and Epidemic Control

Yan Deng
Department of Industrial and Operations Engineering, University of Michigan, USA

Siqian Shen
Department of Industrial and Operations Engineering
University of Michigan, Ann Arbor, MI, USA; e-mail: siqian@umich.edu

Yevgeniy Vorobeychik
Department of Electrical Engineering and Computer Science, Vanderbilt University

Abstract. This paper investigates problems of disease prevention and epidemic control (DPEC), in which we optimize two sets of decisions: (i) vaccinating individuals and (ii) closing locations, given respective budgets with the goal of minimizing the expected number of infected individuals after intervention. The spread of diseases is inherently stochastic due to the uncertainty about disease transmission and human interaction. We use a bipartite graph to represent individuals' propensities of visiting a set of location, and formulate two integer nonlinear programming models to optimize choices of individuals to vaccinate and locations to close. Our first model assumes that if a location is closed, its visitors stay in a safe location and will not visit other locations. Our second model incorporates compensatory behavior by assuming multiple behavioral groups, always visiting the most preferred locations that remain open. The paper develops algorithms based on a greedy strategy, dynamic programming, and integer programming, and compares the computational efficacy and solution quality. We test problem instances derived from daily behavior patterns of 100 randomly chosen individuals (corresponding to 195 locations) in Portland, Oregon, and provide policy insights regarding the use of the two DPEC models.

Key words: disease prevention and intervention; dynamic programming; 0-1 knapsack problem; compensatory behavior modeling; dynamic/static disease control

1 Introduction

In this paper, we study problems of infectious disease prevention and epidemic control (DPEC), which have became increasingly challenging in modern times given convenient means of travel. Meanwhile, disease control is also closely related to defending against bio-terrorism, where quick and high-quality responsive actions can significantly alleviate damages [9, 11]. For vaccinable diseases like influenza, vaccination is considered one of the most effective prevention strategies, and has been usually carried out as a primary response to new epidemics, including the H1N1 virus in 2009 and H3N2 virus in 2011 [10]. Closing locations (e.g., schools), on the other hand, is an effective intervention strategy to control population movements and to limit personal contact during epidemics. In practice, vaccinating individuals and closing locations can be simultaneously implemented to react to disease spread, while the two approaches cost and take effective at very different scales.

Given individual budgets for vaccinating individuals and closing locations, the goal of this paper is to formulate mathematical optimization models, for identifying which subset of individuals we should vaccinate and which locations we should close, to minimize the expected number of infected individuals.

We formulate DPEC problems by using a bipartite graph to model individuals' probabilities of visiting different locations. Suppose that disease may spread differently across subgroups of individuals in an at-risk population. We associate the uncertainty with disease carriers' infectious risk, location visiting behavior, and intervention effects. For each person who may visit a set of locations, we call the person has a compensatory behavior if he/she will choose an alternative place if the ones with higher preferences are closed, versus someone who does not have such a behavior will give up the visit if the destination location is closed. We consider two DPEC variants: one with and the other without individuals' compensatory behavior of visiting locations. We formulate integer nonlinear programming models for both cases, and develop algorithms based on greedy and dynamic programming (DP) strategies. We test instances describing behavior of 100 randomly chosen individuals visiting 195 locations in Portland, Oregon. For disease control over multiple time periods, we demonstrate the effectiveness of repeatedly implementing the results of the static DPEC problem, compared with an approach that dynamically updates people's infectious probabilities and resolves the problem at each time period. We also derive policy insights for disease control when considering compensatory behavior of visiting locations.

In the literature, related work has been performed in areas of modeling disease transmission dynamics [14], identifying critical individuals and predicting disease incidence by patient screening [12, 15, 16], preventing outbreaks by allocating medical resources [1, 2, 18], patient treatment [19], and dynamically closing locations [4, 7]. Our paper is of interest to policy makers who must decide how to allocate fixed budgets to both prevention and intervention phases as a whole [e.g., 17], to achieve the best control effect in terms of minimizing the expected number of infected individuals.

The remainder of the paper is organized as follows. Section 2 describes DPEC and introduces the notation. Section 3 formulates a basic model, denoted by DPEC-B, which excludes people's compensatory behavior of visiting locations. We show that the problem is NP-hard, and develop two alternative approximate approaches based on greedy and DP for solving DPEC-B. Section 4 considers an extension of DPEC-B (named as DPEC-E) by taking into account compensatory behavior using utility functions associated with subgroups' visiting preferences and binary variables for prioritizing visits. We use an exact DP algorithm for optimizing DPEC-E. In Section 5, we demonstrate the computational results, and derive policy insights via various types of testings. Section 6 concludes the paper and suggests future research directions.

2 Problem Statement and Notation

Let $\mathcal{F} = \{1, \ldots, n\}$ be a set of n different locations, and $\mathcal{P} = \{1, \ldots, m\}$ a set of m individuals. The preferences of visiting each location are characterized by probabilities p_{ij}, for all $i \in \mathcal{P}$ and $j \in \mathcal{F}$ with $\sum_{j \in \mathcal{F}} p_{ij} = 1$, $\forall i \in \mathcal{P}$. Person $i \in \mathcal{P}$ has a probability of being initially infectious, characterized by h_i ($0 \le h_i \le 1$), which accounts for Person i's diverse backgrounds, such as social connections, medical history, etc. We also consider locations that are sufficiently local, and therefore the data used in later experiments will not take into account individuals' traveling activities from/to locations outside set $\mathcal{F}$. In our model, chances of an individual becoming infected depend solely on whether there is an infectious individual at a location he/she visit, and is not sensitive to the number of infectious individuals in the location. This restriction is crucial to our ability to obtain a closed-form model of infectious disease spread, and is most reasonable when a location corresponds to a small area. Assume that a healthy individual $i \in \mathcal{P}$ gets infected with probability r_i^{BV} before vaccination, and r_i^{AV} after vaccination, given that there exists one or more infectious individuals in a location visited. By default, $r_i^{BV} > r_i^{AV}$ to reflect the effectiveness of vaccination.

This paper focuses on static DPEC models that take into account neither disease dynamics nor decision changes over time. We justify such an assumption by considering cases where policy makers need to quickly make initial prevention and/or intervention plans during a sporadic and explosive disease outbreak, using estimated parameters p_{ij} and h_i for $i \in \mathcal{P}$ and $j \in F$. In addition, decisions such as closing locations and vaccinating individuals are difficult and expensive to adjust dynamically. We later compare our static models with a benchmark dynamic model. Our results demonstrate solution stability in real-time control and the meaning of solving static DPEC models.

Define binary variables z_i for all $i \in \mathcal{P}$, and x_j for all $j \in \mathcal{F}$, such that

$$z_i = \begin{cases} 1 & \text{if Person } i \text{ is vaccinated,} \\ 0 & \text{otherwise;} \end{cases} \quad \text{and } x_j = \begin{cases} 1 & \text{if Location } j \text{ is closed,} \\ 0 & \text{otherwise.} \end{cases}$$

We denote a vaccination decision vector by $z = [z_i : \forall i \in \mathcal{P}]$, and a decision vector for closing locations by $x = [x_j : \forall j \in \mathcal{F}]$. Considering that closing locations and vaccinating people are usually at very different cost scales, we use two separate budgets for vaccination and for closing locations, denoted by B_z and B_x, respectively. Denote $c_i > 0$ as the cost of vaccinating Person i, $\forall i \in \mathcal{P}$, and $d_j > 0$ as the cost of closing Location j, $\forall j \in \mathcal{F}$. Let $\mathbb{E}[\#\text{ infected}]$ be the expected number of infected individuals at all locations. The goal of DPEC is to minimize $\mathbb{E}[\#\text{ infected}]$, while the costs of vaccination and closing locations are respectively limited by B_z and B_x. By the linearity of expectation,

$$\mathbb{E}[\#\text{ infected}] = \sum_{j \in \mathcal{F}} \mathbb{E}_j[\#\text{ infected}],$$

where $\mathbb{E}_j[\#\text{ infected}]$ represents the expected number of infected individuals at Location j, $\forall j \in \mathcal{F}$. We assume that a healthy individual has zero risk of getting infected at any location without the presence of infectious individuals. Following this,

$$\mathbb{E}_j[\#\text{ infected}] = \rho_j \cdot \mathbb{E}[\#\text{ infected at Location } j \mid \text{infection exists at Location } j].$$

where ρ_j represents the probability that infection exists at Location j. For each location $j \in \mathcal{F}$, we can compute ρ_j in advance as

$$\rho_j = 1 - \Pr\{\text{no infection at } j\} = 1 - \prod_{i \in P}(1 - p_{ij} h_i), \tag{1a}$$

The last equality in (1a) follows from the assumption that individuals' visiting probabilities are independent. Moreover, by the independence of vaccination,

$$\mathbb{E}[\#\text{ infected at Location } j \mid \text{infection exists at Location } j] = \sum_{i \in \mathcal{P}} p_{ij}(1 - h_i)\left(r_i^{AV} z_i + r_i^{BV}(1 - z_i)\right).$$

This yields the following objective function for our optimization problem:

$$\min \sum_{j \in \mathcal{F}} \rho_j \left(\sum_{i \in \mathcal{P}} p_{ij}(1 - h_i)\left(r_i^{AV} z_i + r_i^{BV}(1 - z_i)\right) \right)(1 - x_j)$$

Note that this objective is specified in closed form and, moreover, is linear in each of our decision variables. In contrast, most past epidemic models have been rather complex, requiring at times large-scale simulations to estimate infection spread [3]. The simplicity of the objective function we obtain is critical to our ability to formulate epidemic control as mathematical programs below. Next we formulate two integer nonlinear programming models for solving DPEC. The first model assumes no compensatory behavior of visiting locations, and designates individuals to stay home when the destination locations are closed. In contrast, the second model assumes that each individual visits a location with the highest preference among those that are open. In the next two sections, we formulate the two DPEC models, and discuss a few solution algorithms.

3 DPEC with No Compensatory Behavior

3.1 DPEC-B: Formulation and Complexity

The first DPEC model, denoted by DPEC-B, assumes that if a location is closed, individuals who would otherwise visit the location will choose to stay at home. This serves as the basic model, and is formulated as an integer nonlinear program as follows.

$$\textbf{DPEC-B:} \quad \min_{x,z} \quad \sum_{j\in\mathcal{F}} \rho_j \left(\sum_{i\in\mathcal{P}} p_{ij}(1-h_i)\left(r_i^{AV} z_i + r_i^{BV}(1-z_i)\right)\right)(1-x_j) \tag{2a}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{P}} c_i z_i \le B_z \tag{2b}$$

$$\sum_{j\in\mathcal{F}} d_j x_j \le B_x \tag{2c}$$

$$x_j \in \{0,1\} \quad \forall j \in \mathcal{F} \qquad z_i \in \{0,1\} \quad \forall i \in \mathcal{P}. \tag{2d}$$

Denoting $v_{ij} \equiv z_i x_j$ for all $i \in \mathcal{P}$ and $j \in \mathcal{F}$, we replace bilinear terms $z_i x_j$ in (2a), and linearize DPEC-B via a set of linear inequalities [13]:

$$v_{ij} \le z_i,\ v_{ij} \le x_j,\ v_{ij} \ge z_i + x_j - 1,\ v_{ij} \ge 0, \tag{3}$$

where v_{ij} is enforced to be the product of z_i and x_j when z_i and x_j are both binary. To see this, when either z_i or x_j is zero, the first two inequalities will enforce v_{ij} to be zero; otherwise, the third inequality together with the first two inequalities in (3) will ensure $v_{ij} = 1$.

Theorem 1. DPEC-B is NP-hard.

Proof. Let $\overline{\text{DPEC-B}}$ be a decision version of DPEC-B described as follows.

Given $V^+ > 0$, identify binary z and x, satisfying constraints (2b), (2c), and

$$\sum_{j\in\mathcal{F}} \rho_j \left(\sum_{i\in\mathcal{P}} p_{ij}(1-h_i)\left(r_i^{AV} z_i + r_i^{BV}(1-z_i)\right)\right)(1-x_j) \le V^+. \tag{4}$$

Verifying the feasibility of a given solution (z, x) to $\overline{\text{DPEC-B}}$ takes a polynomial number of steps, and thus the problem belongs to NP. Now consider a general 0-1 Knapsack problem [cf. 6] with a set I of items to be added to a knapsack with capacity B. For each item $i \in I$, the value and the weight of adding the item are respectively v_i and w_i. The goal is to find a subset of items in I, with the total weight being no more than B and the total value being at least V^-.

To design a special instance of $\overline{\text{DPEC-B}}$, let set $\mathcal{F}$ contain all the items in set I. Designate $B_z = 0$, $B_x = B$, and $c_i = w_i$, $\forall i \in \mathcal{F}$. Note that $B_z = 0$ makes $z_i = 0$, $\forall i \in \mathcal{P}$, and further changes (4) to

$$\sum_{j \in \mathcal{F}} \left(\rho_j \sum_{i \in \mathcal{P}} p_{ij}(1 - h_i) r_i^{BV} \right) (1 - x_j) \leq V^+. \tag{5}$$

With (2c) ensuring the total weight of selected items in I (or equivalently, closed locations in $\mathcal{F}$) being bounded by B, the Knapsack problem is equivalent of solving a special $\overline{\text{DPEC-B}}$ with designed values of ρ_j, p_{ij}, h_i, and r_i^{BV}, $\forall j \in \mathcal{F}$, $i \in \mathcal{P}$, which satisfy $\rho_j \sum_{i \in \mathcal{P}} p_{ij}(1 - h_i) r_i^{BV} = v_j$ for all $j \in \mathcal{F}$, and $V^+ = \sum_{i \in I} v_i - V^-$. (As an example, one can let $r_i^{BV} = 1$, $h_i = 0$ for all $i \in \mathcal{P}$, $\sum_{i \in \mathcal{P}} p_{ij} = 1$, $\forall j \in \mathcal{F}$, and $\rho_j = v_j$, $\forall j \in \mathcal{F}$.) Due to the NP-completeness of general 0-1 Knapsack, we conclude that $\overline{\text{DPEC-B}}$ is also NP-complete. □

DPEC-B contains knapsack constraints (2b) and (2c), and products of the corresponding 0-1 binary variables in the objective. The complexity of solving DPEC-B is therefore pseudo-polynomial similar to general Knapsack problems. Later we develop algorithms based on similar approaches for solving the Knapsack, by using greedy and DP strategies. Moreover, often in practice, vaccinating each person costs almost the same, and cost for shutting down locations may also be quite similar. Thus, (2b) and (2c) can be further simplified to cardinality knapsack constraints respectively with identical costs for vaccinating an individual and closing one location. This will save significant computational effort especially for large-scale problems.

3.2 Algorithms for Solving DPEC-B

Assume that vaccination protects individuals from being infected (i.e., $r_i^{AV} = 0$ for all $i \in \mathcal{P}$), and people without vaccination will surely be infected (i.e., $r_i^{BV} = 1$ $\forall i \in \mathcal{P}$) when exposed to infection. As a result, we simplify the objective (2a) in DPEC-B as

$$\min \sum_{j \in \mathcal{F}} \sum_{i \in \mathcal{P}} \lambda_{ij} (1 - z_i)(1 - x_j). \tag{6}$$

where $\lambda_{ij} = \rho_j p_{ij}(1 - h_i)$. Using general values of r_i^{AV} and r_i^{BV} will not change the complexity of algorithms we develop in this section. In Remark 1, we demonstrate how to modify our algorithms to compute general DPEC-B, where both r_i^{AV} and r_i^{BV} can be fractional and $r_i^{BV} > r_i^{AV}$, $\forall i \in \mathcal{P}$.

Consider DPEC-B as Formulation (2) with the objective (2a) replaced by (6), subject to decision independent constraints (2b) and (2c). The exact computation requires solving nested DP recursions and visiting all non-dominated states defined by feasible 0-1 solution combinations of x and z. The number of states is exponential, depending on problem parameters B_x and B_z.

We first develop an approximate algorithm following a greedy strategy, which iteratively updates the solution until no improvement can be made to the objective. Denote $(\hat{x}, \hat{z})$ as a feasible solution corresponding to a subset $\hat{\mathcal{F}}^0 \subseteq \mathcal{F}$ and a subset $\hat{\mathcal{P}}^0 \subseteq \mathcal{P}$, such that

$$\hat{x}_j = \begin{cases} 0 & \text{if } j \in \hat{\mathcal{F}}^0 \\ 1 & \text{if } j \in \mathcal{F} \setminus \hat{\mathcal{F}}^0 \end{cases}, \quad \hat{z}_i = \begin{cases} 0 & \text{if } i \in \hat{\mathcal{P}}^0 \\ 1 & \text{if } i \in \mathcal{P} \setminus \hat{\mathcal{P}}^0 \end{cases}$$

That is, sets $\hat{\mathcal{F}}^0$ and $\hat{\mathcal{P}}^0$ respectively keep track of locations that have not been closed, and people who have not been vaccinated. Now consider "switching" the values of some entries in $(\hat{x}, \hat{z})$, e.g., increasing $\hat{x}_j$ for some $j \in \hat{\mathcal{F}}^0$ or $\hat{z}_i$ for some $i \in \hat{\mathcal{P}}^0$ from zero to one. Note that such changes will decrease the current objective value, and will also reduce respective remaining budgets in DPEC-B.

We start with $(\hat{x}, \hat{z}) = (\mathbf{0}, \mathbf{0})$, i.e., $\hat{\mathcal{F}}^0 = \mathcal{F}$ and $\hat{\mathcal{P}}^0 = \mathcal{P}$. At each iteration, we change the value of $\hat{x}_j$ for some Location j from 0 to 1, and update $\hat{\mathcal{F}}^0 = \hat{\mathcal{F}}^0 \setminus \{j\}$. (Similarly, we change $\hat{z}_i$ from 0 to 1 for some Person i, and update $\hat{\mathcal{P}}^0 = \hat{\mathcal{P}}^0 \setminus \{i\}$.) The algorithm follows a greedy strategy by selecting Location j (or Person i) according to the ratio of "objective reduction" to "cost of action" described as follows. Because the objective decreases by $\sum_{i \in \hat{\mathcal{P}}^0} \lambda_{ij}$ at cost d_j, for closing Location $j \in \mathcal{F}$, we define the *value* of changing $\hat{x}_j$ from 0 to 1, with given $\hat{\mathcal{P}}^0$ and $\hat{\mathcal{F}}^0 \setminus \{j\}$ by

$$X(j, \hat{\mathcal{P}}^0) = \left(\sum_{i \in \hat{\mathcal{P}}^0} \lambda_{ij} \right) / d_j \quad \forall j \in \hat{\mathcal{F}}^0. \tag{7}$$

Similarly, the *value* of switching $\hat{z}_i$ from 0 to 1, with given $\hat{\mathcal{F}}^0$ and $\hat{\mathcal{P}}^0 \setminus \{i\}$, is defined by

$$Z(i, \hat{\mathcal{F}}^0) = \left(\sum_{j \in \hat{\mathcal{F}}^0} \lambda_{ij} \right) / c_i \quad \forall i \in \hat{\mathcal{P}}^0. \tag{8}$$

We recursively close Location j (or vaccinate Person i) by following a descending order of $X(j, \hat{\mathcal{P}}^0)$ for all $j \in \hat{\mathcal{F}}^0$ (or a descending order of $Z(i, \hat{\mathcal{F}}^0)$ for all $i \in \hat{\mathcal{P}}^0$). At the end of each iteration, the new objective value is given by

$$V := \sum_{i \in \hat{\mathcal{P}}^0} \sum_{j \in \hat{\mathcal{F}}^0} \lambda_{ij},$$

corresponding to the current sets $\hat{\mathcal{F}}^0$ and $\hat{\mathcal{P}}^0$. We iterate the foregoing procedures until V becomes relatively stable. Algorithm 1 outlines critical steps of the greedy algorithm, in which V_{pre} represents an objective value from a previous step. We stop the algorithm when $V - V_{\text{pre}} \leq \epsilon$, where ϵ is positive and sufficiently small.

Example 1. Consider a DPEC-B example where $\mathcal{P} = \{1, 2\}$, $\mathcal{F} = \{1, 2\}$, $\lambda_{11} = 1$, $\lambda_{12} = 0.6$, $\lambda_{21} = 0.4$, $\lambda_{22} = 0.7$, $c_1 = c_2 = 1$, $d_1 = d_2 = 1$, $B_z = 1$, $B_x = 1$. We demonstrate Algorithm 1

Algorithm 1 A greedy algorithm for approximating solutions to DPEC-B.

Input: A DPEC-B instance and a sufficiently small $\epsilon > 0$.
Output: An objective value V^*.

Compute λ_{ij} for all $i \in \mathcal{P}$ and $j \in \mathcal{F}$.
Initialize $V := \sum_{j\in\mathcal{F}} \sum_{i\in\mathcal{P}} \lambda_{ij}$ and $\hat{\mathcal{P}}^0 := \mathcal{P}$.
repeat
  Let $V_{\text{pre}} := V$ as the best objective from last iteration.
  Set $K_x := B_x$, $\hat{\mathcal{F}}^0 := \mathcal{F}$.
  for $j \in \mathcal{F}$ **do**
    Compute $X(j, \hat{\mathcal{P}}^0) := \left(\sum_{i\in\hat{\mathcal{P}}^0} \lambda_{ij}\right)/d_j$.
  end for
  Sort all j in $\mathcal{F}$ in a descending order of $X(j, \hat{\mathcal{P}}^0)$, denoted by $\mathcal{L}_x = \{j_1, \ldots, j_n\}$ such that $X(j_1, \hat{\mathcal{P}}^0) \geq \cdots \geq X(j_n, \hat{\mathcal{P}}^0)$.
  for $k = 1, \ldots, n$ **do**
    if $d_{j_k} \leq K_x$ **then**
      Update $\hat{\mathcal{F}}^0 := \hat{\mathcal{F}}^0 \setminus \{j_k\}$, $K_x := K_x - d_{j_k}$.
    end if
  end for
  Set $K_z := B_z$, $\hat{\mathcal{P}}^0 := \mathcal{P}$.
  for $i \in \mathcal{P}$ **do**
    Compute $Z(i, \hat{\mathcal{F}}^0) := \left(\sum_{j\in\hat{\mathcal{F}}^0} \lambda_{ij}\right)/c_i$.
  end for
  Sort all i in $\mathcal{P}$ in a descending order of $Z(i, \hat{\mathcal{F}}^0)$, denoted by $\mathcal{L}_z = \{i_1, \ldots, i_m\}$, such that $Z(i_1, \hat{\mathcal{F}}^0) \geq \cdots \geq Z(i_m, \hat{\mathcal{F}}^0)$.
  for $k = 1, \ldots, m$ **do**
    if $c_{i_k} \leq K_z$ **then**
      Update $\hat{\mathcal{P}}^0 := \hat{\mathcal{P}}^0 \setminus \{i_k\}$, $K_z := K_z - c_{i_k}$.
    end if
  end for
  Compute $V := \sum_{i\in\hat{\mathcal{P}}^0} \sum_{j\in\hat{\mathcal{F}}^0} \lambda_{ij}$.
until $V_{\text{pre}} - V \leq \epsilon$
return $V^* = V$.

for $\epsilon = 10^{-6}$ as follows. Initially, both $\hat{\mathcal{P}}^0$ and $\hat{\mathcal{F}}^0$ are $\{1, 2\}$. In the first iteration, we compute $X(j, \hat{\mathcal{P}}^0)$ for all $j \in \mathcal{F}$, yielding $X(1, \{1, 2\}) = 1.4$ and $X(2, \{1, 2\}) = 1.3$. Exclude $j = 1$ from $\hat{\mathcal{F}}^0$, and update $\hat{\mathcal{F}}^0 = \{2\}$. Compute $Z(i, \hat{\mathcal{F}}^0)$ for all $i \in \mathcal{P}$, yielding $Z(1, \{2\}) = 0.6$ and $Z(2, \{2\}) = 0.7$. As a result, update set $\hat{\mathcal{P}}^0 = \{1\}$, and the objective value $V = 0.6$. We pass $\hat{\mathcal{P}}^0 = \{1\}$ to the next iteration, which sets $\hat{\mathcal{F}}^0$ back to $\{1, 2\}$. By computing and ordering $X(j, \{1\})$ for all $j \in \mathcal{F}$, we have $\hat{\mathcal{F}}^0 = \{2\}$ because $X(1, \{1\}) > X(2, \{1\})$. Therefore, we end with the same solutions and objective value. We terminate the algorithm and return $V^* = 0.6$.

Theorem 2. Algorithm 1 converges in a finite number of steps.

Proof. We show the result by contradiction. Let V^k denote the value of V obtained at the end of iteration k ($k \geq 1$ and integer). Suppose that Algorithm 1 does not converge. We then have an infinite sequence $\{V^1, \ldots, V^k, \ldots\}$ in which $V^k < -(k-1)\epsilon + V^1$ for any k. As k goes to infinity, $V^k \to -\infty$. Because any feasible objective value of DPEC-B must be bounded below by 0. This is a contradiction. □

Optimality gap. Algorithm 1 might terminate at a solution whose objective value is significantly higher than the optimal objective value. In Example 1, if we enumerate all feasible solutions x and z, the optimal objective value is 0.4. However, the greedy algorithm returns a suboptimal solution with an objective gap being

$$\frac{0.6 - 0.4}{0.4} \times 100\% = 50\%.$$

To show that such a gap can sometimes be arbitrarily large, we artificially design an example as follows. Consider $\mathcal{P} = \{1, 2\}$, $\mathcal{F} = \{1, 2\}$, $c_i = 1$ for all $i \in \mathcal{P}$, $d_j = 1$ for all $j \in \mathcal{F}$, and $B_x = 1$ and B_z=1. The parameters λ_{ij} satisfy

$$\lambda_{11} \geq \lambda_{12},\ \lambda_{22} \geq \lambda_{12},\ \lambda_{11} + \lambda_{21} \geq \lambda_{12} + \lambda_{22},\ \lambda_{12} > \lambda_{21}. \tag{9}$$

Given solutions $\tilde{x} = (\tilde{x}_1, \tilde{x}_2) = (1, 0)$, and $\tilde{z} = (\tilde{z}_1, \tilde{z}_2) = (0, 1)$, let $V(\tilde{z}, \tilde{x})$ be the corresponding objective value in (6), i.e., λ_{12}. According the first three conditions in (9),

$$V(\tilde{z}, \tilde{x}) \leq V(\tilde{z}, x) \text{ for any feasible } x \text{ to DPEC-B},$$
$$\text{and } V(\tilde{z}, \tilde{x}) \leq V(z, \tilde{x}) \text{ for any feasible } z \text{ to DPEC-B}.$$

Hence, Algorithm 1 will always arrive at solution $(\tilde{z}, \tilde{x})$ by following the greedy strategy, and return $V^* = V(\tilde{z}, \tilde{x}) = \lambda_{12}$. However, due to the last condition in (9), $V^* = \lambda_{21}$ is a better objective value, corresponding to solutions $z_1^* = 1$, $z_2^* = 0$, $x_1^* = 0$, $x_2^* = 1$. The gap between the two objectives is $100(\lambda_{12} - \lambda_{21})/\lambda_{21}\%$, which can be arbitrarily large if $\lambda_{12} \gg \lambda_{21}$.

Complexity of Algorithm 1. Values of V obtained from adjacent iterations k and $k+1$ in Algorithm 1 must follow $V^k - V^{k+1} > \epsilon$, while values of V^k in all iterations k are bounded from below by zero. Therefore, the number of iterations taken by the algorithm is bounded by $1 + \lfloor V^1/\epsilon \rfloor$. The number of steps at each iteration is $O(m+n)$, leading to an overall complexity $O((m+n)/\epsilon)$.

Alternatively, one can use a hybrid approach to approximately solve DPEC-B by integrating the greedy criteria with DP recursions with respect to the two different sets of decisions. We consider two possible schemes for implementing the hybrid approach:

- Scheme 1: Greedily close locations and subsequently vaccinate individuals via a DP algorithm.

- Scheme 2: Greedily vaccinate individuals and subsequently close locations via a DP algorithm.

In Appendix A, we demonstrate the algorithmic steps by following Scheme 1. For Scheme 2, Similar procedures can be performed, which we omit in this paper but only provide the complexity result.

Remark 1. Both Algorithm 1 and Algorithm 2 are for special DPEC-B where $r_i^{BV} = 1$ and $r_i^{AV} = 0$. (That is, at a location with the presence of infectious individuals, chances for people with and without vaccination to be infected are 0 and 100%, respectively.) We modify our algorithms as follows to make them applicable for solving general DPEC-B with $0 \leq r_i^{AV} < r_i^{BV} \leq 1$.

Now the objective of a general DPEC-B takes the form

$$\min \sum_{j\in\mathcal{F}} \sum_{i\in\mathcal{P}} \lambda_{ij} \left(r_i^{BV} - (r_i^{BV} - r_i^{AV})z_i\right)(1 - x_j),$$

with coefficients of z variables determined by general r_i^{BV} and r_i^{AV}, rather than 0 and 1 in (6). The values of changing entries in x and in z in Algorithm 1 respectively change to

$$X(j, \hat{\mathcal{P}}^0) = \left(\sum_{i\in\hat{\mathcal{P}}^0} \lambda_{ij} r_i^{BV} + \sum_{i\in\mathcal{P}\setminus\hat{\mathcal{P}}^0} \lambda_{ij} r_i^{AV}\right) /d_j, \text{ and } Z(i, \hat{\mathcal{F}}^0) = (r_i^{BV} - r_i^{AV}) \left(\sum_{j\in\hat{\mathcal{F}}^0} \lambda_{ij}\right) /c_i.$$

This also applies to the greedy steps in Algorithm 2 (see Appendix A), such that the value of closing Location j changes to $X(j, \mathcal{P}) = \left(\sum_{i\in\mathcal{P}} \lambda_{ij} r_i^{BV}\right)/d_j,\ \forall j \in \mathcal{F}$. As a result, we revise the DP recursions in Algorithm 2 by

$$f_i(k) = \min_{z_i\in\{0,1\}:\ c_i z_i \leq k} \left\{ \sum_{j\in\mathcal{F}\setminus\hat{\mathcal{F}}^1} \lambda_{ij} \left(r_i^{BV} - (r_i^{BV} - r_i^{AV})z_i\right) + f_{i+1}(k - c_i z_i) \right\},$$

for all $i \in \mathcal{P}$ and $k = 0, \ldots, B_z$. The complexities of both Algorithm 1 and Algorithm 2 are not affected by these changes.

4 DPEC with Compensatory Behavior

The previous DPEC-B model is based on the assumption that if a location is closed, individuals who otherwise visit the location will stay at home and do not visit any other locations. Such an assumption may not be effective when travel restrictions issued by the government are not severe, or when closed locations have other alternatives providing similar service.

In this section, we consider a DPEC problem variant, named DPEC-E, in which subgroups of individuals have visiting preferences and will visit alternative locations if the ones that are more preferred are closed. We formulate the problem as an integer program, which defines additional specially ordered set of type one (SOS1) variables for determining locations visited by each individual after intervention. This section also develops an exact DP-based algorithm for optimizing DPEC-E.

4.1 An Integer Programming Model for DPEC-E

We model individuals' probabilities of visiting each location via utility functions representing the visiting preferences and the status of the location (open or close). The population now contains multiple groups of rational individuals, who will visit locations with the highest utilities among those that are open. Denote the set of behavioral groups by $\Theta = \{1, \dots, w\}$ with $w \ll n$. For people in each group $\theta \in \Theta$, a distinct utility u_j^θ is assigned for visiting Location j, $j \in \mathcal{F}$. Every person i is associated with a probability vector $\mathbf{f}_i$ to characterize his/her likelihood of belonging to each group, where $\mathbf{f}_i = [f_{i,\theta} : \theta \in \Theta]$.

To calculate the probability of Person i visiting Location j, we use variables that are SOS1, and their values are determined by both solution x and utility u. Define binary variables a_j^θ such that $a_j^\theta = 1$ if all individuals in Group θ visit Location j, and $a_j^\theta = 0$ otherwise. Define scalar U^θ as the utility associated with the location that Group θ visits, for all $\theta \in \Theta$. For any Group θ, only one of the a_j^θ's can take value 1 depending on which open location has the highest utility value. Let M be an arbitrarily large number, set as $M = \max_{j \in \mathcal{F}, \theta \in \Theta} u_j^\theta$ in our later computation. The following constraints ensure that individuals in each group will visit an open location with the highest utility:

$$\sum_{j \in \mathcal{F}} a_j^\theta = 1 \quad \forall \theta \in \Theta,\ a_j^\theta \in \{0, 1\} \quad \forall j \in \mathcal{F},\ \theta \in \Theta \tag{10a}$$

$$U^\theta - u_j^\theta (1 - x_j) \geq 0 \quad \forall\ j \in \mathcal{F} \text{ and } \theta \in \Theta \tag{10b}$$

$$U^\theta - u_j^\theta (1 - x_j) \leq M(1 - a_j^\theta) \quad \forall\ j \in \mathcal{F} \text{ and } \theta \in \Theta. \tag{10c}$$

(10a) makes sure that each group of individuals only visits one location in $\mathcal{F}$; (10b) and (10c) together ensure that for all $\theta \in \Theta$, Group θ visits an open location that has the maximum utility. To see this, for every $\theta \in \Theta$, when $x_j = 1$ as Location j is closed, both (10b) and (10c) are relaxed, and a_j^θ is set to 0 in (10c) to keep it feasible. Otherwise, (10b) enforce U^θ to be the maximum u_j^θ among all locations with $x_j = 0$. Let $j^*(\theta)$ be such a location whose $u_{j^*(\theta)}^\theta = U^\theta$. Then, the $j^*(\theta)^{\text{th}}$ constraint in (10c) will set $a_{j^*(\theta)}^\theta = 1$, with all other a_j^θ kept to zero as their left-hand sides of the corresponding constraints in (10c) are positive. (Recall that we assume all utility values are distinct, and thus $j^*(\theta)$ is unique for every group θ.) The probability of Person i visiting Location

j (i.e., p_{ij}) is

$$p_{ij} = \sum_{\theta} f_{i,\theta} a_j^{\theta}, \tag{10d}$$

We give the integer programming model of DPEC-E as follows.

$$\textbf{DPEC-E:} \quad \min_{x\in\{0,1\}^n,\ z\in\{0,1\}^m,\ a\in\{0,1\}^{w\times n}} \{(2a) : (2b),\, (2c),\, (10a)\text{–}(10d)\}\,.$$

4.2 An Exact Algorithm for Optimizing DPEC-E

The idea of an exact algorithm for optimizing DPEC-E is to formulate DP recursions to traverse all possible *effective* solutions x, and then optimize a restricted DPEC-E with fixed x to obtain the corresponding best vaccination decisions z.

4.2.1 Preference lists and frontier locations

For every group $\theta \in \Theta$, we define a *preference list*, denoted by σ^θ, as a permutation of set $\mathcal{F} = \{1, \ldots, n\}$ with a descending order of the utilities viewed by Group θ. For Group $\theta \in \Theta$, its preference list is given by

$$\sigma^\theta = \{\sigma_1, \ldots, \sigma_n\} : \sigma_j \in \mathcal{F}\ \forall j = 1, \ldots, n \text{ with } u^\theta_{\sigma_1} > \ldots > u^\theta_{\sigma_n}.$$

Group 1	Group 2	Group 3	Group 4	Group 5
113	14	66	30	14
10	80	113	58	58
52	102	55	158	169
61	73	181	168	46
111	143	148	185	65

(a)

G1	G2	G3	G4	G5
113	14	**66**	**30**	14
10	**80**	113	58	**58**
52	102	55	158	169
61	73	181	168	46
111	143	148	185	65

(b)

Figure 1: Examples of preference lists and frontier locations

Figure 1(a) illustrates an example of preference lists, for five behavioral groups. For each group, the corresponding column contains the location IDs ordered by their utility values, with the highest utility showing at the top. For example, $\sigma^1 = \{113, 10, 52, 61, 111\}$ means that Location 113 is the most favorable and Location 10 is the second favorable considered by Group 1.

Figure 1(b) demonstrates how decisions of closing locations affect which location people in each group visit. The shaded entries correspond to a closed location (Location 14), and for each group, we underline locations that are visited, as their utilities are the highest among all open locations. The location chosen by a group is named "frontier location." For example, Locations 113, 80, 66, 30, and 58 are frontier locations for Groups 1–5, respectively.

4.2.2 Algorithmic details

Closing non-frontier locations is *ineffective*, as it will not reroute people to visit different locations and thus will not change the objective. Hence, we only need to sequentially close frontier locations to progressively improve the objective value. This becomes the key in the development of an exact algorithm for optimizing DPEC-E. We consider DPEC-E as a series of location closing followed by vaccination decisions at the end. More specifically, given a subset of locations to close at a given stage, our decisions are either (i) to stop closing locations and start vaccination, or (ii) to close one more location (if budget allows) and proceed to the next stage.

Denote $g(\mathcal{F}^1, k)$ as the minimum expected infection size that is achievable, given that all locations in set $\mathcal{F}^1$ have already been closed and a budget k remains for closing additional locations. (The budget for vaccination is a constant B_z for any $g(\mathcal{F}^1, k)$.) For all possible $\mathcal{F}^1 \subseteq \mathcal{F}$ and $0 \le k \le B_x$, the DP recursions are

$$g(\mathcal{F}^1, k) = \min \begin{cases} \min\limits_{\theta \in \Theta : d_{(\sigma^\theta \setminus \mathcal{F}^1)_1} \le k} & g\left(\mathcal{F}^1 + (\sigma^\theta \setminus \mathcal{F}^1)_1, k - d_{(\sigma^\theta \setminus \mathcal{F}^1)_1}\right), \\ \\ \kappa^*(\mathcal{F}^1), & \text{i.e., the optimal objective of DPEC-E given} \\ & x_j = 1 \ \forall j \in \mathcal{F}^1 \text{ and } x_j = 0 \ \forall j \in \mathcal{F} \setminus \mathcal{F}^1. \end{cases}$$

Here we use $(\bullet)_1$ to represent the first element of set $\bullet$. Set $\sigma^\theta \setminus \mathcal{F}^1$ is obtained from subtracting every location in $\mathcal{F}^1$ from list σ^θ while keeping the order of the original locations that are still open. Thus, $(\sigma^\theta \setminus \mathcal{F}^1)_1$ gives the frontier location for Group θ after closing all locations in $\mathcal{F}^1$. When set $\left\{\theta \in \Theta : d_{(\sigma^\theta \setminus \mathcal{F}^1)_1} \le k\right\} = \emptyset$ for some k and $\mathcal{F}^1$, we set $g(\mathcal{F}^1, k)$ positively infinite.

An initial state is given by $(\mathcal{F}^1, k) = (\emptyset, B_x)$. For state $(\mathcal{F}^1, k)$, if we stop closing locations and proceed to vaccinate individuals, the resulting infection spread is obtained from solving DPEC-E with x being fixed. If we decide to close more locations, the above recursion closes one more location from the current frontier locations $(\sigma^1 \setminus \mathcal{F}^1)_1, \ldots, (\sigma^w \setminus \mathcal{F}^1)_1$ and then embarks on the next stage. We update set $\mathcal{F}^1$ by including the just-closed location, and reduce budget k by deducting the corresponding cost for closing.

Given a fixed solution x to DPEC-E characterized by set $\mathcal{F}^1 = \{j \in \mathcal{F} : x_j = 1\}$, we compute

$$a_j^\theta = \mathbf{1}\left(\text{argmax}_{l \in \mathcal{F} \setminus \mathcal{F}^1} \{u_l^\theta\} = j\right) \quad \text{for all } j \in \mathcal{F},\ \theta \in \Theta \tag{12a}$$

$$p_{ij} = \textstyle\sum_{\theta \in \Theta} f_{i,\theta} \mathbf{1}\left(\text{argmax}_{l \in \mathcal{F} \setminus \mathcal{F}^1} \{u_l^\theta\} = j\right) \quad \text{for all } i \in \mathcal{P},\ j \in \mathcal{F}, \tag{12b}$$

where $\mathbf{1}(\bullet)$ returns 1 if $\bullet$ is true and 0 otherwise. For a given set of closed locations in $\mathcal{F}^1$, selecting which individuals to vaccinate is given by

$$\min_{z\in\{0,1\}^m} \left\{ \sum_{j\in\mathcal{F}\setminus\mathcal{F}^1} \sum_{i\in\mathcal{P}} \rho_j p_{ij}(1-h_i)(1-z_i) \ : \ \sum_{i\in\mathcal{P}} c_i z_i \le B_z \right\}, \tag{13}$$

as a 0-1 Knapsack problem that can be optimized by off-the-shelf MIP solvers.

Remark 2. This algorithm searches for locations to close only among the current frontier locations at each step. Consider a special case with cardinality budget constraint for closing locations (i.e., $d_1 = \cdots = d_n = 1$), such that B_x represents the maximum number of locations that one can close. Frontier locations throughout the recursion are subsets of $(\sigma^1)_{B_x} \cup \cdots \cup (\sigma^w)_{B_x}$, where $(\sigma^\theta)_{B_x}$ returns a set of first B_x elements of list σ^θ. A search over such a subset, with its cardinality $\le wB_x$, is much more efficient than a search over the entire $\mathcal{F}$.

5 Computational Results

We test both DPEC-B and DPEC-E on problem instances derived from real-world datasets. The computation emphasizes on (i) demonstrating computational efficacy of the solution algorithms, (ii) varying parameters and deriving policy insights from the DPEC results. All computations are performed on a HP Workstation Z210 Windows 7 machine with Intel(R) Xeon(R) CPU 3.20 GHz, and 8GB memory. All involved integer programs are solved by default CPLEX 12.4 [8] via ILOG Concert Technology with C++.

5.1 Experimental Design and Computational Setup

We use EpiSims simulator [3] to produce instances used for evaluating our models and approaches. EpiSims simulator models epidemic spread and probabilities that represent individual movement patterns among locations. The data is for Portland, Oregon, and describes typical movement of the city's population over a 24-hour period. The dynamic movement of individuals simulated by EpiSims is formed by actual census, land-use, and population mobility data. We therefore view it as representative of actual travel patterns.

Each data point in our data set corresponds to an individual engaged in an activity at a specific location, and contains the following characteristics: An ID of the individual, an ID of the household to which this individual belongs, an ID of the activity that this data point represents, the purpose of this activity (for example, home or shopping), start time and duration of the activity, and, finally, an ID of the location where the activity takes place. From this, we extract only the person ID, location ID, and activity duration. We then process this information to obtain p_{ij} for all $i \in \mathcal{P}$ and

$j \in \mathcal{F}$ as follows: After a pass through the entire data set, we collect for Person $i \in \mathcal{P}$ the set of locations that he/she has visited. We set $p_{ij} = 0$ for any location j that is never visited by Person i. If j is visited by i, we compute the total time that i spends at this location over all visits to the location, and divide this value by the total time that Person i spends at *all* locations to obtain p_{ij}. While our data contains approximately 1.6 million people moving among approximately 250,000 different locations, we randomly choose 100 individuals from it for evaluating our algorithms, which subsequently restricts our attention to 195 corresponding locations that the 100 people visit.

To test DPEC-E, we categorize the entire population into five behavioral groups $\Theta = \{1, \ldots, 5\}$. Each pair of group and location is assigned a random integer $u_j^\theta,\ j \in \mathcal{F},\ \theta \in \Theta$, ranging from 1 to 100 representing the group's preference for visiting Location j. For each person we generate five random numbers and normalize them to represent the likelihood that a person belongs to five behavioral groups.

We adopt three different sets of initial infectious probabilities $h = [h_i : i \in \mathcal{P}]$. The first set is randomly generated from an exponential distribution with mean = 1.5% to represent the most general pattern of disease-carrying initial state in a population. This set is used in the majority of our computation, with the following two exceptions. Tests in Section 5.2.3 rely on a less variable h, randomly generated as initial infectious probabilities from a uniform distribution over [0.05,0.15], to compute dynamic infection status through time. Tests in Section 5.2.5 aim to compare different approaches for solving DPEC-E. We randomly generate entry values in 100 different vectors h from an exponential distribution with mean = 1%.

Two different sets of cost values are tested and analyzed. In Section 5.2, we randomly generate the vaccination cost $c_i,\ \forall i \in \mathcal{P}$ from normal distribution $N(\mu, \sigma^2) = N(10, 1^2)$ and facility closing cost $d_j,\ \forall j \in \mathcal{F}$ from normal distribution $N(\mu, \sigma^2) = N(100, 50^2)$. Tests in Section 5.3 use cardinality cost values, and set $c_1 = \cdots = c_m = 1$ and $d_1 = \cdots = d_n = 1$.

5.2 Results of DPEC-B

We replace bilinear terms $z_i x_j$ in the DPEC-B formulation (2) by inequalities in (3) and solve the resulting binary integer programming model in the optimization solver CPLEX to compute the optimal objective. Our analysis of DPEC-B covers the following aspects. In Section 5.2.1, we analyze the effect of varying budgets B_x and B_z on disease control, and articulate solution patterns such as the types of closed locations and vaccinated individuals. In Section 5.2.2, we justify the superiority of integrating decisions of vaccination and location closing, by examining results of separating the two strategies in disease control. In Section 5.2.3, we demonstrate the stability of static DPEC-B, and compare with a benchmark dynamic strategy, where the infectious probabilities evolve over time. In Section 5.2.4, we compare the results of DPEC-B with a ubiquitous "high-degree" strategy in practice. In Section 5.2.5, we compare Algorithm 1 and Algorithm 2 for approximating solutions to DPEC-B.

5.2.1 Sensitivity analysis and solution patterns

We denote $B_x\%$ as the ratio of location closing budget to the cost of closing all location, i.e., $B_x / \sum_{j\in\mathcal{F}} d_j$, and $B_z\%$ denotes the ratio of vaccination budget to the cost of vaccinating the entire population, i.e., $B_z / \sum_{i\in\mathcal{P}} c_i$. Figure 2 depicts the optimal objectives of DPEC-B, corresponding to $B_x\%$ ranging between [0%, 6%] and $B_z\%$ ranging between [0%, 20%]. When budgets increase, the

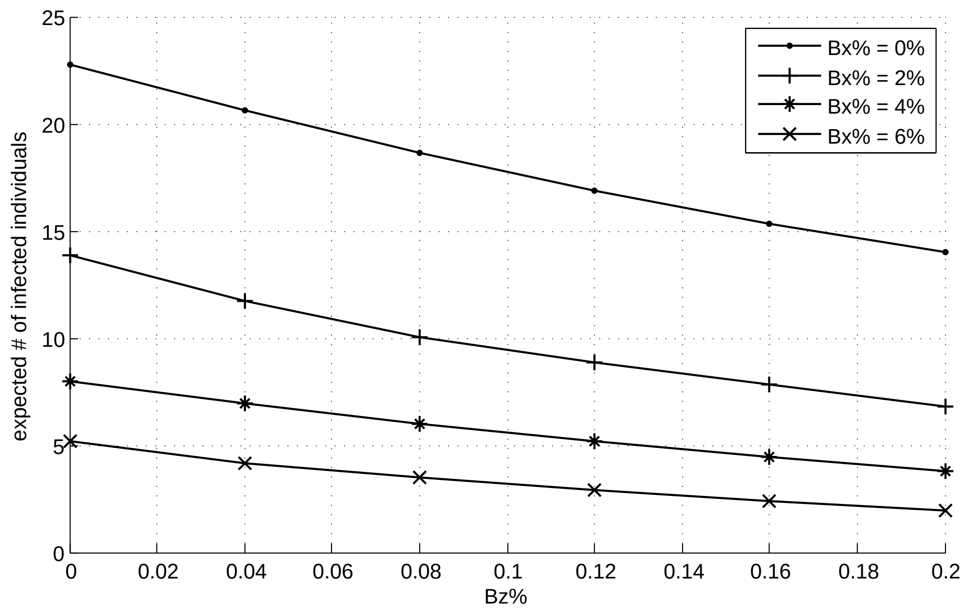
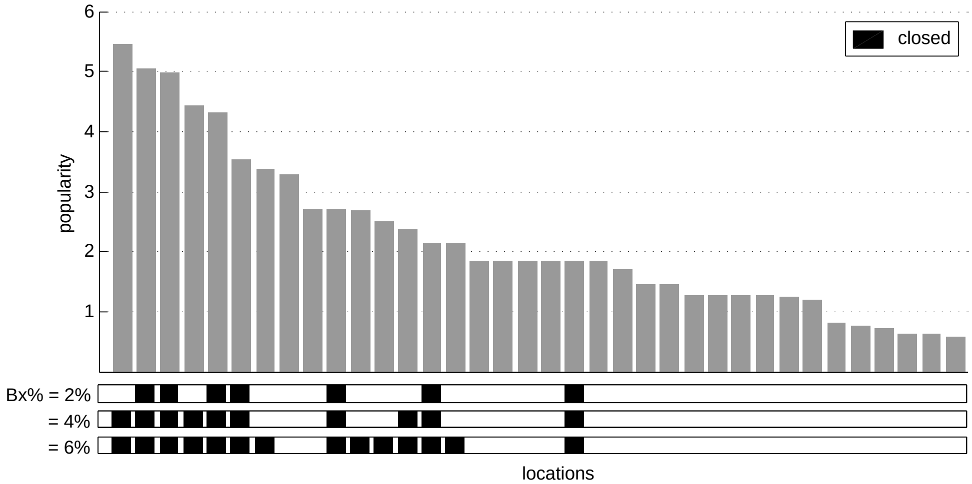


Figure 2: Optimal objectives of DPEC-B under different B_x and B_z

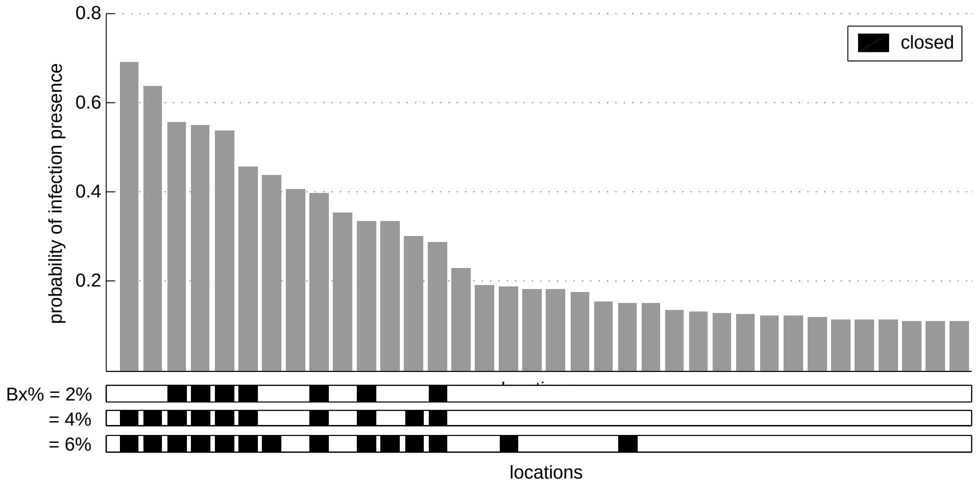
expected number of infections decreases at a decreasing rate, reflecting a submodular disease control effect for each unit of additional budget input in both prevention and intervention procedures.

Characterizing closed locations. The DPEC models we study involve two types of heterogeneity: the initial infectious probabilities (characterized by h_i) and propensities of visiting locations (characterized by p_{ij} in DPEC-B, and by $f_{i,\theta}$, u_j^θ in DPEC-E). Intuitively, one prefers to close locations that are the most popular among all individuals, and locations that have the highest chances of having infected individuals in presence.

Our computational results in Figure 3 show evidences of both intuitions. Define the popularity



(a) Relation between location closing and location popularity ($B_z\% = 8\%$)

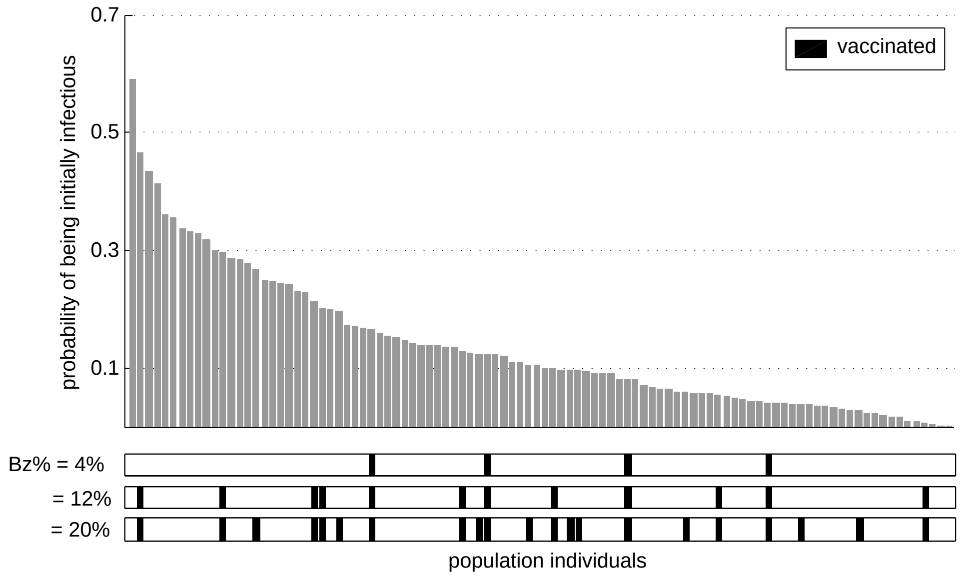


(b) Relation between location closing and the presence of infection ($B_z\% = 8\%$)

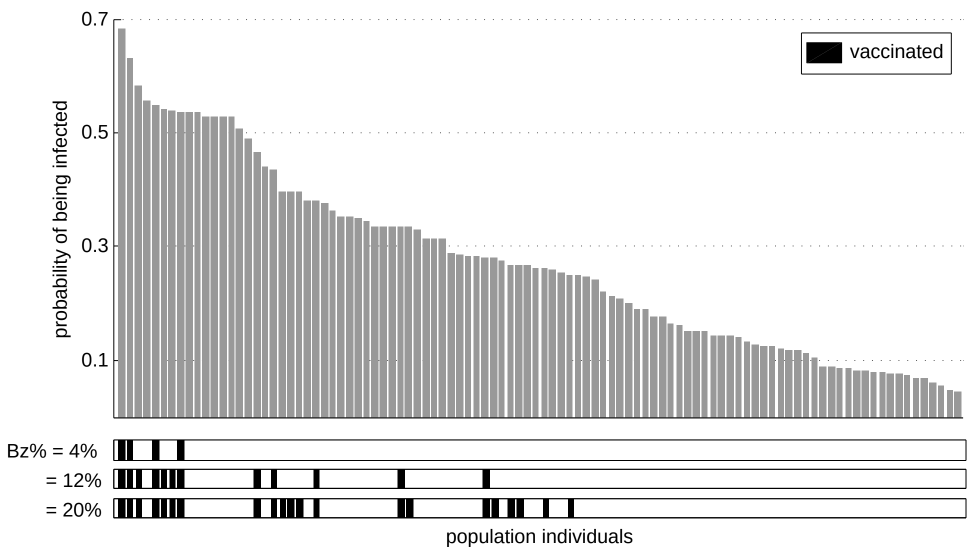
Figure 3: Closed locations in DPEC-B and their characteristics

of Location j by the summation of its visiting propensities from all individuals, i.e., $\sum_{i\in\mathcal{P}} p_{ij}$. Figure 3(a) highlights closed locations in DPEC-B for $B_x\% = 2\%$, 4%, and 6% (with a constant $B_z\% = 8\%$), and depicts their popularity in a decreasing order. The figure only contains the first 36 locations (out of 195) which include all locations that are closed. When $B_x\% = 2\%$ and $B_z\% = 8\%$, we observe that closed faculties are the most popular ones. This becomes even more obvious as we increase B_x to 4% and 6%. Figure 3(b) highlights closed locations according to the probabilities of having infection presence at different locations (i.e., $\rho_j,\ j \in \mathcal{F}$), where it is evident that locations with higher chances of having infectious individuals are more likely to be closed.

Characterizing vaccination solutions. Figure 4 depicts relations of vaccination solutions with the initial infectious probabilities (i.e., h_i for all $i \in \mathcal{P}$) and with individuals' probabilities of being infected after vaccination, which we refer to as "post-vaccination infection probabilities" (i.e., $(1 - h_i) \sum_{j \in \mathcal{F}^0} p_{ij} \rho_j$ for all $i \in \mathcal{P}$).



(a) Relation between vaccination and initial infectious probabilities ($B_x\% = 2\%$)



(b) Relation between vaccination and post-vaccination infection probabilities ($B_x\% = 2\%$)

Figure 4: Optimal vaccination in DPEC-B and their characteristics

Figure 4 has each vertical set of bars/blocks represent an individual in $\mathcal{P}$: the black blocks, if exist, signify the vaccinated individuals; the gray bars in Figure 4(a) stretch proportional to the initial infectious probabilities, and in Figure 4(b) stretch proportional to post-vaccination infection probabilities. We order each individual following a descending order of initial infectious probabilities in Figure 4(a) and a descending order of the post-vaccination infection probabilities in Figure 4(b). The results are given for $B_z\% = 4\%$, 12% and 20% with $B_x\% = 2\%$ being constant.

The distribution of vaccination subject in Figure 4(a) displays no special pattern. In Figure 4(b) it tends to concentrate towards the left end, corresponding to relatively high post-vaccination infection probabilities. This observation is in line with an intuition that vaccination does not depend on initial infectious probabilities but is affected by individuals' chances of being infected.

5.2.2 Effect of integrating vaccination and location closing

Strategies of vaccination and location closing are sometimes separately analyzed in the epidemic control literature. In this paper, we simultaneously optimize both variables, where closing locations controls infection spread by either forcing people to cancel their visits so as to reduce their exposure to possible infection (i.e., DPEC-B), or reshaping individuals' travel patterns (i.e., DPEC-E).

We compare DPEC-B with a decision model in which vaccination and location closing decisions are optimized respectively in Formulation (13) and in Formulation

$$\min_{x\in\{0,1\}^n} \left\{ \sum_{j\in\mathcal{F}} \sum_{i\in\mathcal{P}} \rho_j p_{ij}(1-h_i)(1-x_j) \; : \; \sum_{j\in\mathcal{F}} d_j x_j \le B_x \right\}. \tag{14}$$

We then fix the two sets of solutions in (2a) to calculate the corresponding expected number of infected individuals under separate prevention and intervention control. We set budgets B_z and B_x to 1%, 3%, 5%, 7% and 9% of $\sum_{i\in\mathcal{P}} c_i$ and $\sum_{j\in\mathcal{F}} d_j$, respectively. Our results show that vaccination and location closing are supplement to each other, and a good integration of the two could effectively control infection spread.

Figure 5 demonstrates the comparison, in which each ($B_x\%$,$B_z\%$) contains two bars to represent the expected number of infected individuals in DPEC-B and in the separate control model, respectively. The solid line in Figure 5 depicts the expected infected number before intervention. We observe that integrating prevention and intervention decisions leads to smaller infection spread in all instances. The winning margins get bigger when budgets B_x and B_z increase.

5.2.3 Static versus dynamic DPEC-B

The paper focuses on static models for closing locations and vaccinating population. This is reasonable if we consider the sporadic nature of disease outbreaks and fast responding actions that are needed. Here we justify the approach from another perspective and show that the static solutions, if applied repeatedly, yield results nearly as good as dynamically optimizing DPEC-B.

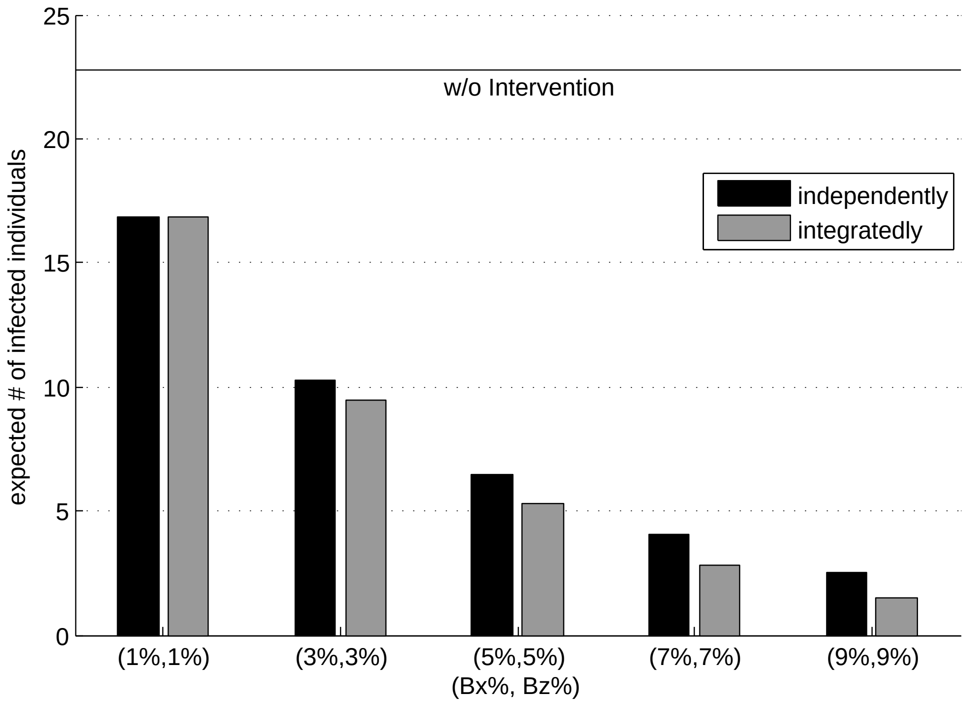


Figure 5: Comparing integrated and independent disease control decisions

Consider the disease control as a dynamic process over a finite time horizon $1, \ldots, T$ in which decisions are made at the beginning of each time period, while individuals' infectious probabilities evolve as we implement the solutions through time. We parameterize all decisions and the infectious probabilities with time t, as (x^t, z^t) and $h^t,\ t = 1, \ldots, T$, to represent decisions made at period t, and infectious probabilities at the end of period t, respectively. (Vector h^0 contains all initial infectious probabilities.) Values of h^t are recursively updated based on h^{t-1} and decisions (x^t, z^t), using

$$h_i^t = h_i^{t-1} + (1 - h_i^{t-1})(r_i^{AV} z_i^t + r_i^{BV}(1 - z_i^t)) \sum_{j \in \mathcal{F}} p'_j p_{ij} (1 - x_j^t) \quad \forall t = 1, \ldots, T,\ i \in \mathcal{P}. \tag{15}$$

That is, the probability of Person i being infectious at the end of time t equals to his/her infectious probability at the end of time $t - 1$ plus the complementary multiplied by the chance of Person i getting infected in time t. We justify this dynamic model by applications of annual disease control, where individuals may be vaccinated in multiple years for the same disease, and locations can be closed but reopen in different years.

Two strategies are considered for optimizing DPEC-B over multiple time periods.

- Strategy 1: At the beginning of period t, solve DPEC-B with $h = h^{t-1}$ and compute solutions (x^t, z^t). Update h^t according to (15) and repeat the process for $t = 1, \ldots, T$.

- Strategy 2: At the beginning of time 1, solve DPEC-B with $h = h^0$ and apply its solution (x^1, z^1) repeatedly to all time periods, i.e., $(x^t, z^t) = (x^1, z^1)$ for $t = 2, \ldots, T$.

Figure 6 illustrates the cumulative objectives of DPEC-B by using Strategy 1 ("Dynamic Optimization") and Strategy 2 ("Static Approximation"), over 35 time periods, for three $(B_x\%, B_z\%)$ combinations. In general, Static Approximation yields good approximations to results of Dynamic Optimization, with relevant deviation consistently smaller than 4%. This observation provides a strong support for this paper to focus on the static DPEC-B model. After solving one DPEC-B, a policy maker can apply the solution to multiple time periods, without updating the infectious probabilities of individuals. More importantly, this approach is much more efficient computationally compared with Dynamic Optimization.

5.2.4 DPEC-B versus a high-degree heuristic strategy

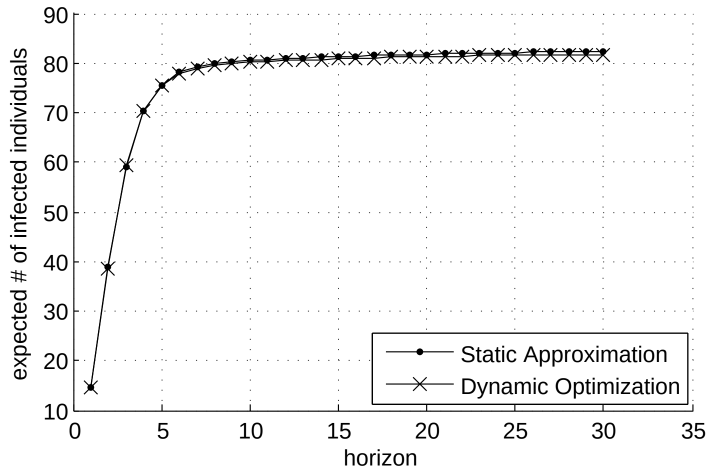
Gallos et al. [5] describes an approach commonly used in the practice of disease control, in which a policy maker simply vaccinates individuals who visit the most number of locations, and closes locations that are visited by the most people. We refer to this policy as the high-degree strategy, for which we define the degree of Location $j \in \mathcal{F}$ by $\mathcal{D}^{\mathcal{F}}(j)$ and the degree of Person $i \in \mathcal{P}$ by $\mathcal{D}^{\mathcal{P}}(i)$, where $\mathcal{D}^{\mathcal{F}}(j) = \sum_{i \in \mathcal{P}} p_{ij}, \ \forall j \in \mathcal{F}$, and $\mathcal{D}^{\mathcal{P}}(i) = \sum_{j \in \mathcal{F}} p_{ij} \mathcal{D}^{\mathcal{F}}(j), \ \forall i \in \mathcal{P}$.

We test $B_x\% = 0\%, \ 2\%, \ 4\%$ and 6%, and $B_z\%$ ranging between $[0\%, 16\%]$. The results are compared between DPEC-B and the high-degree strategy, where the latter repeatedly closes locations and vaccinates individuals with the highest degrees until exhausting budgets B_x and B_z. Figure 7 presents the results of DPEC-B and the high-degree strategy. The former is significantly better for all $(B_x\%, B_z\%)$ combinations.

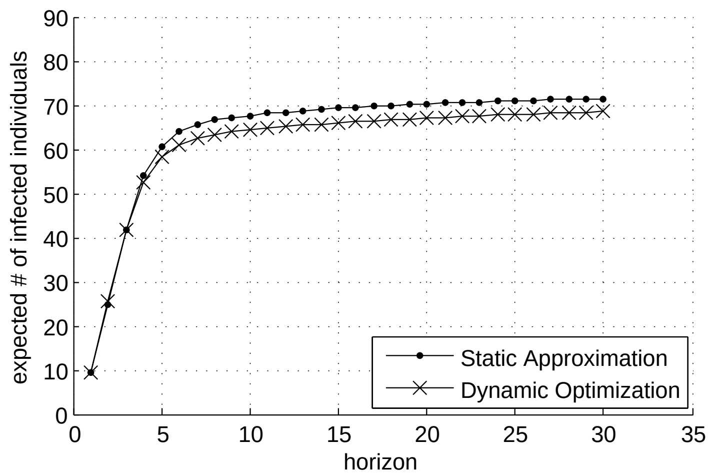
5.2.5 Algorithm 1 versus Algorithm 2

We generate 100 sets of initial infectious probabilities h_i for all $i \in \mathcal{P}$ for this test. Using each set of data, we implement Algorithm 1 and Algorithm 2 for solving DPEC-B with various budgets (e.g., $B_x\%$ set to 0.5%, 1%, 1.5% and 2%, and $B_z\%$ varied between $[0\%, 25\%]$).

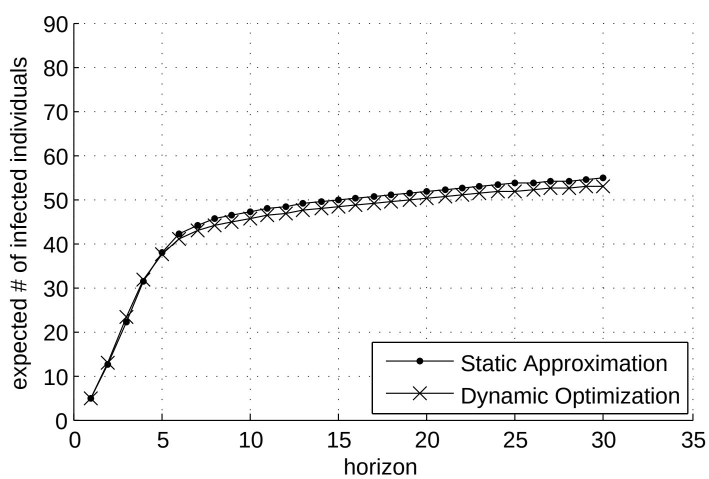
Table 1 presents the objective averages and CPU time given by the two algorithms and by the CPLEX solver (optimal). In general, both the approximate algorithms provide results that are close to the optimal objective values. For $B_z\% \leq 5\%$ and $B_x\% \leq 1\%$, they attain near-optimal solutions, of which the objective values deviate by no more than 3% from the optimal objectives. Relative deviation increases at an increasing rate as we increase budgets, which, though, is mainly attributed to reductions of objective values (i.e., the scales of infection spread shrink effectively



(a) $B_x = 0.5\%, B_z = 5\%$



(b) $B_x = 1\%, B_z = 15\%$



(c) $B_x = 2\%, B_z = 25\%$

Figure 6: Comparing dynamic and static DPEC-B solutions

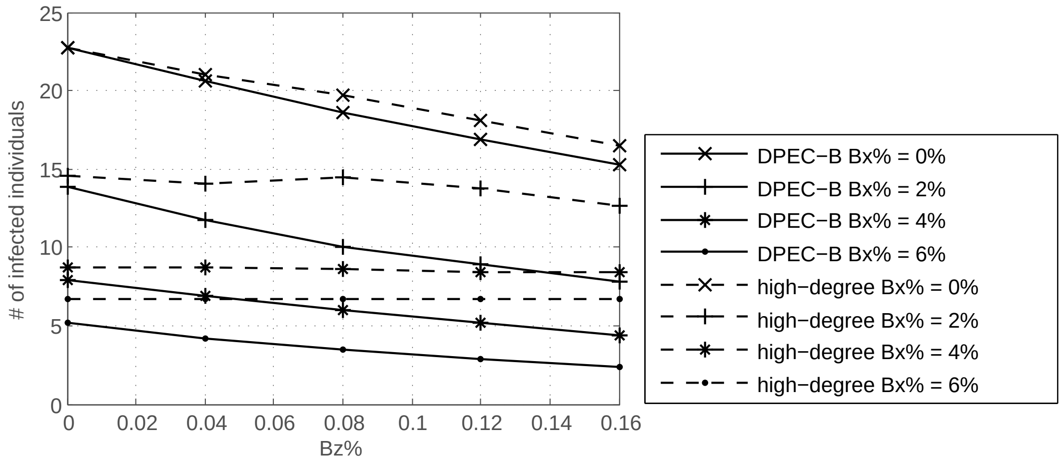


Figure 7: Comparing DPEC-B and the high-degree strategy

given more budgets for intervening). On the other hand, Algorithm 1 and Algorithm 2 use much less CPU time than solving DPEC-B as an integer program in CPLEX.

5.3 Results of DPEC-E

We solve DPEC-E by using the approach developed in Section 4.2. Throughout this section we assume $c_i = 1\ \forall i \in \mathcal{P}$, and $d_j = 1\ \forall j \in \mathcal{F}$. Therefore, B_z and B_x are simply referred to as the maximum number of individuals to vaccinate and the maximum number of locations to close. To visualize the impact of compensatory behavior, we design a DPEC-B comparison instance in line with the settings of DPEC-E as follows. Assume that people in Group θ visit the location with the highest utility and stay at home if it is closed. Accordingly, p_{ij} for all $i \in \mathcal{P}$ and $j \in \mathcal{F}$ and ρ_j for all $j \in \mathcal{F}$ can be calculated based on (12b) and (1a). We demonstrate how compensatory behavior affects disease control and the resulting infection spread.

5.3.1 Sensitivity analysis and solution patterns

Figure 8 depicts the objective values of DPEC-E and the comparison DPEC-B. Each curve, corresponding to a specific $B_x\%$ value, comprises a series of points horizontally located by an increasing sequence of $B_z\%$ values, and vertically by the optimization objectives. We set $B_x\%$ respectively to 0%, 1% and 1.5%, and vary $B_z\%$ between $[0\%, 20\%]$.

Without closing any location, both DPEC-E and DPEC-B reduce to the same formulation that minimizes the infection spread through budgeted vaccination, parameterized by equivalent visiting probabilities. Therefore, the curves in DPEC-E and DPEC-B for $B_x\% = 0\%$ coincide exactly. Meanwhile, the curves' vertical distribution reveals that increasing B_x results in notable reduction

Table 1: Objectives and CPU seconds of the greedy, hybrid, and exact approaches

Budgets		**Algorithm 1**		**Algorithm 2**		**CPLEX**	
B_x%	B_z%	$\mathbb{E}$[# infected][a]	CPU sec	$\mathbb{E}$[# infected]	CPU sec	$\mathbb{E}$[# infected]	CPU sec
0.5%	0%	11.4035	0.02	11.4035	0.01	11.1675	0.12
0.5%	5%	7.7925	0.02	7.8565	0.00	7.5678	1.23
0.5%	15%	3.1874	0.02	3.3859	0.02	3.0515	1.35
0.5%	25%	0.9395	0.01	1.1011	0.02	0.8764	1.31
0.5%	35%	0.1599	0.03	0.2282	0.02	0.1358	1.19
1%	0%	8.6115	0.02	8.6115	0.02	8.5186	0.10
1%	5%	5.5182	0.03	5.5822	0.01	5.3716	2.10
1%	15%	1.9306	0.01	2.0851	0.01	1.7855	2.17
1%	25%	0.4672	0.01	0.5629	0.02	0.3982	1.85
1.5%	0%	6.7063	0.02	6.7063	0.01	6.5814	0.10
1.5%	5%	4.0281	0.02	4.0859	0.01	3.8761	1.10
1.5%	15%	1.1968	0.01	1.2981	0.02	1.0286	1.22
1.5%	25%	0.2286	0.02	0.2883	0.02	0.1576	1.03
2%	0%	5.1832	0.02	5.1832	0.01	5.0389	0.11
2%	5%	2.9436	0.02	3.0060	0.02	2.6824	0.94
2%	15%	0.7865	0.02	0.8530	0.01	0.5343	1.04
2%	25%	0.1190	0.02	0.1553	0.01	0.0445	0.82

[a] Instances with objective values lower than 0.1000 have been omitted.

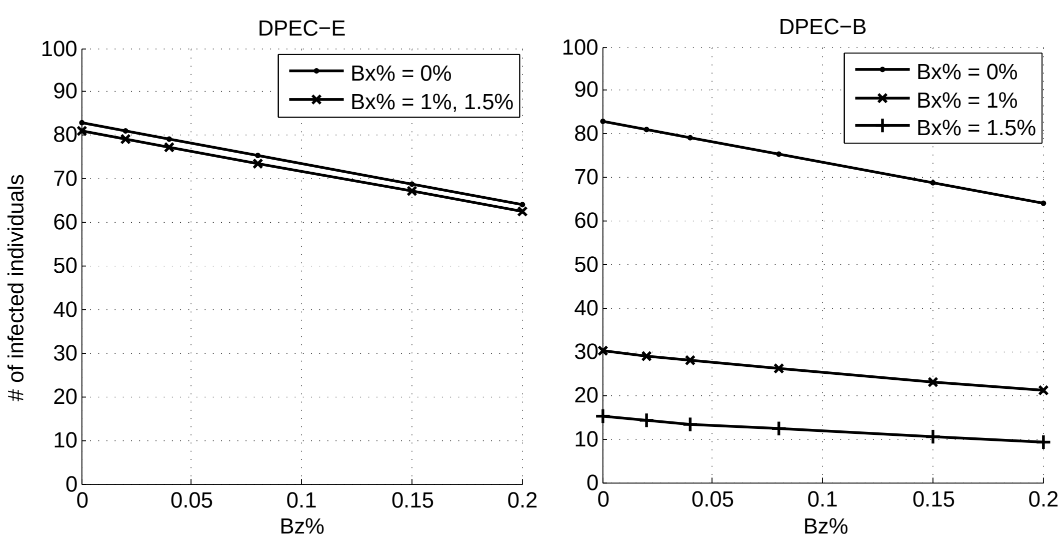


Figure 8: Optimal objectives of DPEC-E under different B_x and B_z

of infection spread in DPEC-B, while it has little impact in DPEC-E.

In DPEC-E, closing locations does not prohibit people from making visits (since they visit alternative locations if the top choice is closed). As a result, the intervention does not change the total number of visits, but only induces different patterns of people's presence in the locations. A higher budget B_x entails more feasible decisions x which potentially constitute more patterns of infectious presence and exposure. Therefore, the scale of infection decreases as we increase B_x, but not strictly.

Our observation here has important policy implications: when the primary effect of closing down locations is to let people simply congregate elsewhere, the action may have negligible benefit. In such a case, it may be better to utilize limited resources for vaccinating more individuals or applying alternative intervention strategies, e.g., travel restriction.

Characterizing closed locations. First, we show that closing additional locations does not always reduce the scale of infection in DPEC-E. For a specified set $\mathcal{F}^1$ of closed locations, we append DPEC-E with equality constraints $x_j = 1$ for $j \in \mathcal{F}^1$ and $x_j = 0$ otherwise. Table 2 presents the optimal objective values of DPEC-E for different selections of $\mathcal{F}^1$, where $B_z = 8$ in all tested instances.

Table 2: DPEC-E results for fixed sets $\mathcal{F}^1$ of closed locations

# of locations to close ($\lvert\mathcal{F}^1\rvert$)	locations to close ($\mathcal{F}^1$)	expected # of infected individuals
0	None	75.3
1	14	73.6
1	66	77.1
2	14,66	75.4
2	14,113	73.6
3	14,66,113	73.6

In Table 2, by having all locations open, the minimum expected number of infected individuals is 75.3. If we close Location 14, the objective decreases to 73.6. However, by closing Location 66, we increase the objective to 77.1, leading to an observation that closing additional locations does not always improve the objective value of DPEC-E. To see why, people who planned to visit the closed locations will instead visit other alternatives, which may cause worse cases that inevitably have high infectious individuals gathering in the same location, leading to the objective value increase.

Our results also demonstrate a correlation between the number of frontier locations and the infection spread in DPEC-E. We test DPEC-E and DPEC-B on a specific instance with $B_z = 8$ and $B_x = 2$. (As a reference, the optimal objective value with $B_x = 0$ and the same B_z is 75.3.)

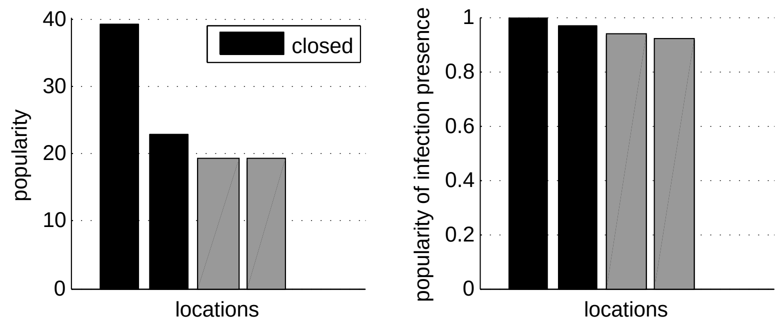
Table 3: Results of DPEC-B and DPEC-E for $B_x = 2$ and $B_z = 8$

Model	Objective	Closed locations at optimum
DPEC-B	26.2	{14,66}
DPEC-E	73.6	{14}, {14,58}, {14,80}, {14,113}

In Table 3, the objective of DPEC-B decreases by $(75.3 - 26.2)/75.3 = 65.2\%$, through closing Locations 14 and 66. The DPEC-E objective values are much higher than DPEC-B, indicating the importance of individuals' cooperation, as well as the government's enforcement on restricting travels and raising public awareness during epidemic periods.

G0	G1	G2	G3	G4
113	14 (shaded)	66 (shaded)	**30**	14 (shaded)
10	80	113	58	58
52	102	55	158	169
61	73	181	168	46
111	143	148	185	65

(a) Closed locations (shaded) and fronter locations (underscored)



(b) Relation of closed locations, location popularity, and probability of the presence of infection

Figure 9: Locations closed and relations to parameters in DPEC-B

Results of closing Locations 14 and 66 given by DPEC-B are given in Figure 9(a), where individuals in Groups 1, 2 and 4 are enforced to stay home, considerably decreasing the number of visits performed in all locations. Solution patterns of location closing are in line with previous observations about DPEC-B that popular locations or locations with high presence rates of infection are usually closed (see Figure 9(b)).

Meanwhile, closing up to two locations in DPEC-E reduces the expected number of infected individuals by $(75.3 - 73.6)/75.3 = 2.3\%$. For the four alternative closing solutions, Figure 10 shades closed locations and underlines the corresponding frontier locations. Individuals from behavioral groups visit locations that are different from one another, and there are five distinct frontier locations. This observation indicates that solutions of closing locations intends to maximize the number of distinct locations visited by different behavior groups.

To provide further evidence, we compute the objective values of DPEC-E corresponding to a set of feasible solutions for $B_z = 8$ and $B_x = 3$. For each feasible solution, we identify the number

G1	G2	G3	G4	G5
113	14	**66**	**30**	14
10	**80**	113	58	**58**
52	102	55	158	169
61	73	181	168	46
111	143	148	185	65

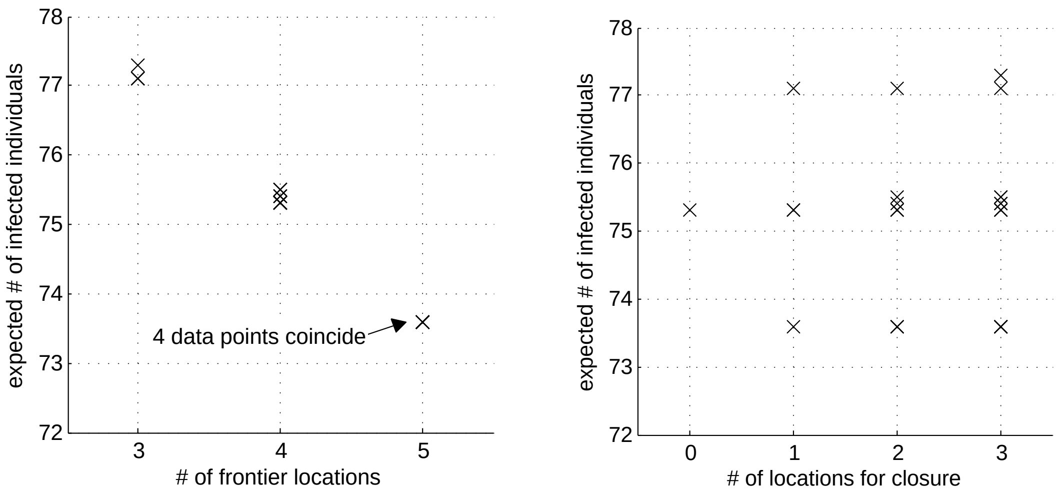
G1	G2	G3	G4	G5
113	14	**66**	**30**	14
10	80	113	58	**58**
52	**102**	55	158	169
61	73	181	168	46
111	143	148	185	65

G1	G2	G3	G4	G5
113	14	**66**	**30**	14
10	**80**	113	58	58
52	102	55	158	**169**
61	73	181	168	46
111	143	148	185	65

G1	G2	G3	G4	G5
113	14	**66**	**30**	14
10	**80**	113	58	**58**
52	102	55	158	169
61	73	181	168	46
111	143	148	185	65

Figure 10: Frontier locations and closed locations in DPEC-E

of distinct frontier locations selected by all five groups. Figure 11(a) presents each solution as a point, with the horizontal axis denotes the number of distinct frontier locations and the vertical axis denotes the objective value. As a comparison, we revise the horizontal axis in Figure 11(b) to denote the number of locations that are closed correspondingly.



(a) (b)

Figure 11: Relation of infection spread and distinct frontier locations in DPEC-E

Figure 11 demonstrates a strong negative correlation between infection spread and the number of frontier locations (not the number of closed locations). A good location closing decision features more distinct frontier locations, such that crowd disperses across more locations in smaller groups. Following this pattern, an infectious person will potentially infect fewer people, and thus leads to a smaller infection spread.

The four coinciding data points in Figure 11(a) correspond to the four closed locations in Figure 10, where frontier locations of all groups are mutually different. If we consider locations as unordered, the four cases represent the same location visiting pattern. This explains why they yield the same infection spread. On the other hand, we observe small fluctuation of the objective when there exists at least one location visited by more than one behavioral groups (i.e., the number of frontier locations equal to 3 or 4 (but < 5) in Figure 11(a)). The fluctuation is due to different behavioral group sizes. This type of disparities in infection spread caused by group size heterogeneity, is much smaller than disparities caused by different patterns of frontier locations. The latter is a consequence of merging (or splitting) the current population in different behavioral groups to locations with diverse infection risk.

5.3.2 Results of DPEC-B implemented in DPEC-E

We compute the objective values of DPEC-E by using optimal solutions of DPEC-B, for $B_x\% = 1\%$ and 1.5% (respectively corresponding to closing one and two locations), and $B_z\%$ ranging between $[0\%, 20\%]$. Table 4 reports the objective values of DPEC-B and DPEC-E, while Column **B-in-E** presents the values of $\mathbb{E}[\#\text{ infected}]$ when DPEC-B solutions are implemented in DPEC-E.

Table 4: Objectives of DPEC-E given by optimal DPEC-B solutions

Budgets		$\mathbb{E}[\#\text{ infected}]$[a]			Budgets		$\mathbb{E}[\#\text{ infected}]$		
B_x	$B_z\%$	DPEC-B	B-in-E	DPEC-E	B_x	$B_z\%$	DPEC-B	B-in-E	DPEC-E
1	0%	30.36	83.03	81.05	2	0%	15.10	85.16	81.05
1	2%	29.15	81.31	79.16	2	2%	14.15	79.26	79.16
1	6%	27.16	77.74	75.44	2	6%	12.83	75.61	75.44
1	10%	25.31	74.22	71.76	2	10%	11.71	72.17	71.76
1	15%	23.16	69.72	67.20	2	15%	10.39	67.86	67.20
1	20%	21.14	66.36	62.66	2	20%	9.27	63.62	62.66

[a] As a reference, the expected number of infected individuals without vaccination or closing locations is 82.97.

The objective values in both **B-in-E** and **DPEC-E** are much larger than in **DPEC-B**, while the **B-in-E** results are sometimes even worse than the results given by solutions of no vaccination or location closing. Compared with the optimal objective values of DPEC-E, solutions of DPEC-B do not yield good results when the compensatory behavior is actually in presence, but gaps between the two objectives become smaller when more locations can be closed.

5.3.3 DPEC-E versus the high-degree strategy

We consider the high-degree strategy for the DPEC-E case as follows. Define the degrees of Location $j,\ j \in \mathcal{F}$, Person $i,\ i \in \mathcal{P}$, and Group $\theta,\ \theta \in \Theta$ by $\mathcal{D}_E^{\mathcal{F}}(j)$, $\mathcal{D}_E^{\mathcal{P}}(j)$, and $\mathcal{D}_E^{\Theta}(\theta)$, respectively. Let

$$\begin{aligned}
\mathcal{D}_E^{\mathcal{F}}(j) \quad &= \Big(\sum_{\theta\in\Theta} u_j^\theta\Big)\Big/\Big(\sum_{j\in\mathcal{F}}\sum_{\theta\in\Theta} u_j^\theta\Big) \quad \forall j \in \mathcal{F} \\
\mathcal{D}_E^{\Theta}(\theta) \quad &= \Big(\sum_{j\in\mathcal{F}} u_j^\theta\Big)\Big/\Big(\sum_{j\in\mathcal{F}}\sum_{\theta\in\Theta} u_j^\theta\Big) \quad \forall \theta \in \Theta \\
\mathcal{D}_E^{\mathcal{P}}(i) \quad &= \sum_{\theta\in\Theta} f_{i,\theta}\mathcal{D}^{\Theta}(\theta) \quad \forall i \in \mathcal{P}
\end{aligned}$$

Without closing any locations, the frontier locations of all five groups are: 113 for Group 1, 14 for Group 2, 66 for Group 3, 30 for Group 4, and 14 for Group 5, of which the degrees are ranked the 27$^{\text{th}}$, 21$^{\text{st}}$, 4$^{\text{th}}$, 8$^{\text{th}}$, and 21$^{\text{st}}$ highest among all locations in $\mathcal{F}$, respectively.

Subject to budgets B_z and B_z, the high-degree strategy vaccinates individuals and closes locations according to decreasingly ordered degrees $\{\mathcal{D}_E^{\mathcal{P}}(i),\ i \in \mathcal{P}\}$ and $\{\mathcal{D}_E^{\mathcal{F}}(j),\ j \in \mathcal{F}\}$, respectively. For $B_x\% = 0\%,\ 1\%$ and 1.5% (corresponding to closing up to zero, one, and two locations), the high-degree strategy will not close any frontier locations given above, and thus leads to the same results for all tested $B_z\%$ ranging between $[0\%, 20\%]$.

Figure 12 presents the objective values of DPEC-E computed by using the optimization model and the high-degree strategy, where the latter solutions are suboptimal in all cases.

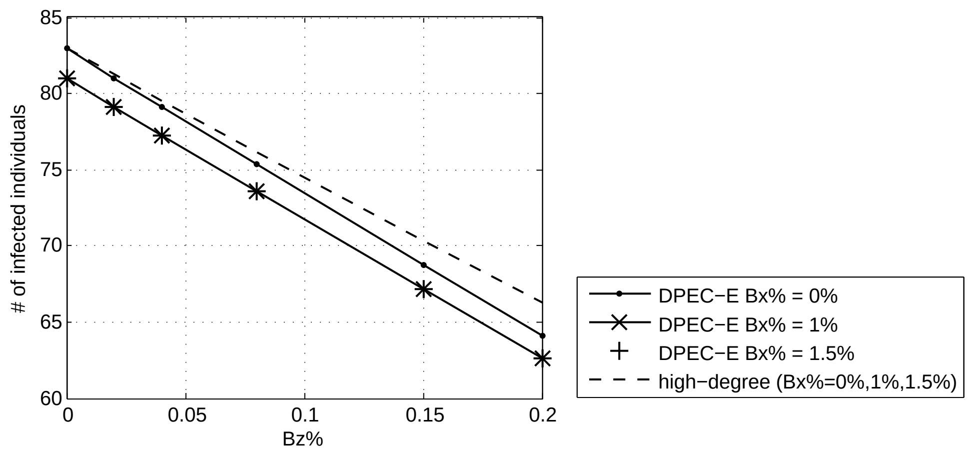


Figure 12: Comparing DPEC-E and the high-degree strategy

6 Conclusions

In this paper, we study problems of disease prevention and epidemic control (DPEC) where people who may carry infectious virus visit a set of locations with certain probabilities. Unlike most DPEC literature, we integrate decisions of disease prevention (e.g., vaccination) and intervention (i.e., closing locations). The paper mainly focuses on two DPEC problem variants, with and without compensatory behavior of visiting locations, respectively. We formulate both problems as integer programming models that can be optimized by off-the-shelf solvers. Moreover, we develop approximate algorithms for efficiently computing feasible solutions to the first DPEC variant without compensatory behavior, and demonstrate their computational efficacy.

Our computation is performed on a simulated instance, involving 100 people, 195 locations, and parameters extracted from real-world datasets describing typical movement of population in

Portland, Oregon, over a 24-hour period. We compare two implementation strategies of the DPEC decisions: One repeatedly solves an updated DPEC by using new infectious probabilities at each time period; the other solves the DPEC once and applies the same static solutions to all time periods. It indicates that solutions solved by static models can yield relatively good results in a setting with dynamic evolving infectious rates. Moreover, we study the relationship between optimal solutions and parameter settings. Finally, by considering compensatory behavior, our models yield the following policy insights. First, closing locations when compensatory behavior exists does not necessarily reduce the number of visits in all locations, but induce different visiting patterns, resulting in minor disease control effects or even infection scale increasing. Therefore, it is important to introduce government interference procedures such as travel restriction, during vaccination and location closing, and to carefully consider the ramifications of these choices. Second, location closing decisions that yield good infection control results intend to increase the number of distinct frontier locations, by triaging a population to as many different locations as possible.

Future research directions include investigating DPEC models in a multi-objective programming context, and comparing the results for different types of diseases. We are also interested in incorporating more types of prevention/intervention decisions into the current models, and study efficient algorithms for optimizing the solutions under extreme events.

Acknowledgment

The authors gratefully acknowledge the editor and two anonymous reviewers. Dr. Vorobeychik is thankful to Sandia National Laboratories as a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

[1] N. G. Becker and D. N. Starczak. Optimal vaccination strategies for a community of households. *Mathematical Biosciences*, 139(2):117–132, 1997.

[2] M. L. Brandeau. Resource allocation for epidemic control. In *Encyclopedia of Optimization*, pages 3291–3295. Springer, 2009.

[3] S. Eubank, H. Guclu, V. S. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429:180–184, 2004.

[4] N. M. Ferguson, D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452, 2006.

[5] L. K. Gallos, F. Liljeros, P. Argyrakis, A. Bunde, and S. Havlin. Improving immunization strategies. *Physical Review E*, 75(4):045104, 2007.

[6] M. Garey and D. Johnson. *Computers and intractability*, volume 174. Freeman San Francisco, CA, 1979.

[7] T. C. Germann, K. Kadau, I. M. Longini Jr, and C. A. Macken. Mitigation strategies for pandemic influenza in the united states. In *Proceedings of the National Academy of Sciences of the United States of America*, volume 103, pages 5935–5940. JSTOR, 2006.

[8] ILOG. *CPLEX 12.1 User's Manual*. ILOG, Armonk, NY.

[9] E. H. Kaplan, D. L. Craft, and L. M. Wein. Analyzing bioterror response logistics: The case of smallpox. *Mathematical biosciences*, 185(1):33–72, 2003.

[10] J. H. Kreijtz, A. D. Osterhaus, and G. F. Rimmelzwaan. Vaccination strategies and vaccine formulations for epidemic and pandemic influenza control. *Human Vaccines*, 5(3):126–135, 2009.

[11] M. Kress. Policies for biodefense revisited: The prioritized vaccination process for smallpox. *Annals of Operations Research*, 148(1):5–23, 2006.

[12] H. L. Lee and W. P. Pierskalla. Mass screening models for contagious diseases with no latent period. *Operations research*, 36(6):917–928, 1988.

[13] G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I – convex underestimating problems. *Mathematical Programming*, 10(1):147–175, 1976.

[14] A. S. R. S. Rao and M. Kakehashi. A combination of differential equations and convolution in understanding the spread of an epidemic. *Sadhana*, 29(3):305–313, 2004.

[15] J. S. Schwartz, B. P. Kinosian, W. P. Pierskalla, and H. Lee. Strategies for screening blood for human immunodeficiency virus antibody. *JAMA: The Journal of the American Medical Association*, 264(13): 1704–1710, 1990.

[16] L. M. Wein and S. A. Zenios. Pooled testing for hiv screening: Capturing the dilution effect. *Operations Research*, 44(4):543–569, 1996.

[17] World Health Organization. WHO and UNICEF joint report. `http://www.who.int/immunization_monitoring/routine/joint_reporting/en/index.html`, January 2011.

[18] J. T. Wu, L. M. Wein, and A. S. Perelson. Optimization of influenza vaccine selection. *Operations Research*, 53(3):456–476, 2005.

[19] J. Zhang, J. E. Mason, B. T. Denton, and W. P. Pierskalla. Applications of operations research to the prevention, detection, and treatment of disease. In *Wiley Encyclopedia of Operations Research and Management*. Wiley, 2011.

APPENDIX

A An Alternative Algorithm for Solving DPEC-B

We describe an algorithm alternative to Algorithm 1 for solving DPEC-B. It follows a hybrid scheme (i.e., Scheme 1 mentioned in Section 3.2) that greedily close locations and subsequently vaccinate individuals based on DP recursions. The details are given as follows.

A greedy strategy to close locations. Steps of greedily closing locations follow the same criteria in Algorithm 1. We start with $\hat{\mathcal{P}}^0 = \mathcal{P}$, and evaluate value $X(j, \mathcal{P}) = \left(\sum_{i \in \mathcal{P}} \lambda_{ij}\right) / d_j$ of closing Location j, $\forall j \in \mathcal{F}$. We reorder locations in $\mathcal{F}$, and consider a list $\mathcal{L} = \{j_1, \ldots, j_n\}$ with $X(j_1, \mathcal{P}) \geq \cdots \geq X(j_n, \mathcal{P})$. We first close locations according to the order in $\mathcal{L}$, until the budget

B_x is exhausted. We denote a feasible solution by $\hat{x}$. For any subset $\hat{\mathcal{F}}^1$ with $\hat{x}_j = 1, \forall j \in \hat{\mathcal{F}}^1$ and $\hat{x}_j = 0, \forall j \in \mathcal{F} \setminus \hat{\mathcal{F}}^1$, the DPEC-B reduces to a 0-1 Knapsack problem:

$$\min \left\{ \sum_{i \in \mathcal{P}} \left(\sum_{j \in \mathcal{F} \setminus \hat{\mathcal{F}}^1} \lambda_{ij} \right) (1 - z_i) : \sum_{i \in \mathcal{P}} c_i z_i \leq B_z,\ z_i \in \{0, 1\}\ \forall i \in \mathcal{P} \right\}. \tag{A-1}$$

DP for solving Formulation (A-1)**.** Apply DP to solve Formulation (A-1) for a given $\hat{\mathcal{F}}^1$. The algorithm involves m stages that sequentially decide whether or not to vaccinate Persons $1, \ldots, m$ subject to a budget limit B_z. At stage i, if we do not vaccinate People i, the objective will increase by $\sum_{j \in \mathcal{F} \setminus \hat{\mathcal{F}}^1} \lambda_{ij}$ with the same remaining budget. Otherwise, the budget B_z decreases by c_i with the objective being unchanged. For all $i \in \mathcal{P}$ and $k = 0, \ldots, B_z$, the DP recursions are

$$\begin{aligned} f_i(k) \quad &= \min_{z_i \in \{0,1\}:\ c_i z_i \leq k} \left\{ \left(\sum_{j \in \mathcal{F} \setminus \hat{\mathcal{F}}^1} \lambda_{ij} \right) (1 - z_i) + f_{i+1}(k - c_i z_i) \right\} \\ &= \begin{cases} \sum_{j \in \mathcal{F} \setminus \hat{\mathcal{F}}^1} \lambda_{ij} + f_{i+1}(k) & \text{if } c_i > k \\ \min \left\{ \sum_{j \in \mathcal{F} \setminus \hat{\mathcal{F}}^1} \lambda_{ij} + f_{i+1}(k),\ f_{i+1}(k - c_i) \right\} & \text{otherwise,} \end{cases} \end{aligned}$$

with boundary conditions being

$$f_{m+1}(k) = 0 \text{ for all } k = 0, \ldots, B_z.$$

We track backwards, and compute $f_i(k),\ \forall k = 0, \ldots, B_z$ for $i = m, \ldots, 1$. Value $f_1(B_z)$ provides the best objective value of Formulation (A-1) for some fixed $\hat{\mathcal{F}}^1$.

Algorithm 2 summarizes critical steps in the hybrid algorithm, where Steps 2-11 corresponds to the greedy part for closing locations, and Steps 12-18 corresponds to the DP iterations for vaccinating individuals.

Optimality gap. Consider Example 1 given in Section 3.2. After the greedy steps in Algorithm 2, we have $X(1, \mathcal{P}) = 1.4$ and $X(1, \mathcal{P}) = 1.3$, and therefore $\mathcal{F}^1 = \{1\}$. We start the DP with $\Lambda_1 = 0.6$ and $\Lambda_2 = 0.7$. Given incumbent $f_3(0) = 0$ and $f_3(1) = 0$, we obtain $f_2(0) = 0.7$ with $x_2 = 0$, and $f_2(1) = 0$ with $z_2 = 1$. In the subsequent stage, $f_1(1)$ attains the minimum 0.6 with $z_1 = 0$, i.e., the solutions are to vaccinate Person 2 and to close Location 1, with an objective value being 0.6. However, $x_1 = 0$, $x_2 = 1$, $z_1 = 1$, $z_2 = 0$ are the optimal solutions with the minimum objective 0.4.

Indeed, similar to the pure greedy algorithm (i.e., Algorithm 1), Algorithm 2 might also yield solutions with arbitrarily large optimality gaps. Consider the same example we design in Section 3.2, the third condition in (9) will set $x_1 = 1$ according to the greedy criteria, and the first and

Algorithm 2 A hybrid approximate algorithm for solving DPEC-B

Input: A DPEC-B instance.

Output: An objective value V^*.

Compute λ_{ij} for all $i \in \mathcal{P}$ and $j \in \mathcal{F}$.
Initialize $K_x := B_x$, $\hat{\mathcal{F}}^1 := \emptyset$.
for $j \in \mathcal{F}$ **do**
  Compute $X(j, \mathcal{P}) := \left(\sum_{i \in \mathcal{P}} \lambda_{ij}\right) / d_j$.
end for
Sort all j in $\mathcal{F}$ in a descending order of $X(j, \mathcal{P})$, denoted by $\mathcal{L} = \{j_1, \ldots, j_n\}$, such that $X(j_1, \mathcal{P}) \geq \cdots \geq X(j_n, \mathcal{P})$.
for $k = 1, \ldots, n$ **do**
  if $d_{j_k} \leq K_x$ **then**
    Update $\hat{\mathcal{F}}^1 := \hat{\mathcal{F}}^1 + j_k$, $K_x := K_x - d_{j_k}$.
  end if
end for
Compute $\Lambda_i := \sum_{j \in \mathcal{F} \setminus \hat{\mathcal{F}}^1} \lambda_{ij}$ for all $i \in \mathcal{P}$.
Define boundary conditions: $f_{m+1}(k) = 0$ for all k.
for $i = m$ to 1 **do**
  for $K_z = 0$ to B_z **do**
    Compute $f_i(k) = \min_{z_i \in \{0,1\}:\ c_i z_i \leq k} \{\Lambda_i(1 - z_i) + f_{i+1}(k - c_i z_i)\}$.
  end for
end for
return $V^* = f_1(B_z)$.

second conditions in (9) will further guarantee the DP recursions vaccinating Person 2 (with $z_2 = 1$ and thus $z_1 = 0$). With the same solution $(\tilde{x}, \tilde{z})$ and suboptimal objective value λ_{12} yielded by Algorithm 2, we can have the optimality gap arbitrarily large by letting $\lambda_{12} \gg \lambda_{21}$.

Complexity Analysis of Algorithm 2. The greedy steps to obtain a subset $\hat{\mathcal{F}}^1$ are $O(n)$. To compute $f_i(k)$ for each state (i, k), we perform (i) one summation step as $\lambda_i + f_{i+1}(k - c_i z_i)\big|_{z_i=1}$ and (ii) one value comparison step as $\lambda_i + f_{i+1}(k - c_i z_i)\big|_{z_i=1}$ versus $f_{i+1}(k - c_i z_i)\big|_{z_i=0}$. Thus, implementing the DP algorithm for each fixed $\hat{\mathcal{F}}^1$ takes $O(mB_z)$ steps. In the worse case, for every person i we calculate $f_i(k)$ for all $k = 0, \ldots, B_z$. The complexity is then proportional to the size of all states, i.e., $O(mB_z)$. The overall complexity of the hybrid approach is $O(n + mB_z)$. Alternatively, Algorithm 2 has a complexity of $O(m + nB_x)$ if we follow Scheme 2, i.e., using the greedy subroutine to decide solution $\hat{\mathcal{P}}^0$ and then using DP to decide solution x.