Optimizing Medical Resource Allocation for Mortality Reduction during Pandemic Outbreaks Using Non-Homogenous Markov Chain

Xiangheng (Annie) Chen

Senior Thesis
Presented to the Department of Economics and Program in Mathematical Methods in the Social Sciences Northwestern University
In Candidacy for Interdisciplinary Honors

Spring 2023

Advisor: Professor Piotr Dworczak
Abstract

This study proposes a theoretical framework using non-homogeneous Markov chain to simulate the dynamics of a pandemic outbreak and to predict mortality rates under different resource allocation strategies. By categorizing individuals based on infection status and underlying health conditions, the model captures transitions between categories with time-dependent probabilities. We calibrate the model using data from Mexico's pandemic outbreak in 2020 and conduct simulations to derive insights on optimal medical resource allocation. The findings indicate that reserving resources for non-pandemic patients not only balances access to treatment but also reduces overall mortality. Additionally, a flexible reserve system that adapts to the increasing severity of a pandemic yields lower mortality rates than fixed reserve systems, making it a more favorable allocation strategy for policymakers.

Key Words: Pandemic outbreak, resource allocation, Markov chain, mortality rate, underlying health conditions, reserve mechanism
Acknowledgments

I would like to offer my sincerest gratitude to my thesis advisor, Professor Piotr Dworczak, for mentoring me through this challenging process with his extensive knowledge on this topic, his unwavering patience toward my inexperience with theory and simulations, and his kindness in providing me with support whenever needed. This paper would not have been possible without his invaluable guidance in shaping and upholding the rigor of my research process. I also want to thank Professor Robert Porter and Professor Molly Schnell for providing me with valuable insights on my thesis topic at the early stage of my research.

I am deeply grateful for all my mentors and peers in the MMSS program. As I reflect upon the four years spent here, I am delighted to realize that my learning surpassed even my highest anticipations, encompassing both academic and non-academic spheres. I want to give a special thank-you to Dorothy Tarasul, with whom I spent countless hours in the library on coursework and thesis-writing throughout the four years. I am grateful for our friendship and mutual support, and for us making it through this rewarding academic experience together.

I am appreciative of the SONIC Research Group, where I broadened my horizon on academic research and had the opportunity to collaborate with exceptional peers on various projects. I am particularly thankful to Professor Noshir Contractor and Dr. Kyosuke Tanaka, with whom I had one of my earliest research experiences and learned so much from their dedication to academic pursuits as well as impeccable work ethics.

I am privileged to have many supportive friends throughout my college years. Thank you to Nancy, Joyce, Charlotte, Angel, Coco, Sam, Vicky, Shannon, Kailyn, Kaitlin, Austin, Marcel, and countless others … for being there for me during significant and insignificant moments, providing me with mental and emotional support. They are the ones that kept me cheerful and made my journey truly meaningful.

Above all, my deepest gratitude goes to my parents and grandparents. Their unconditional support enabled me to embark on this uncertain path of attending college at Northwestern. Despite the physical distance imposed by the pandemic, I have always felt their unwavering presence in my growth. They played a crucial role in influencing my choice of the thesis topic, and it is to them and to all those who recognized the profound care for and from their loved ones amidst the pandemic outbreak that this paper is dedicated.
# Table of Contents

1 Introduction.................................................................................................................... 4

2 Literature Review ......................................................................................................... 6

3 The Non-Homogenous Markov Chain Model............................................................... 8
   3.1 Setting up an Absorbing Markov Chain Framework................................................ 8
   3.2 Determining Health Outcomes through Resource Eligibility................................. 10
   3.3 Modeling Non-Homogeneity of the Transition Matrix .............................................. 12
       3.3.1 The Changing Susceptibility of the Population............................................... 13
       3.3.2 The Changing Share of Resource-Eligible Population..................................... 14

4 MATLAB Simulation .................................................................................................... 17
   4.1 Methodology ............................................................................................................ 17
       4.1.1 Empirical Setting ............................................................................................. 17
       4.1.2 Data Calibration Approach ........................................................................... 18
       4.1.3 Optimization Approach .................................................................................. 19
   4.2 Data Calibration ....................................................................................................... 19
       4.2.1 Parameters and Assumptions ......................................................................... 19
       4.2.2 The Baseline Model and Interpretations ......................................................... 23
   4.3 Simulation-Based Optimization ................................................................................ 26
       4.3.1 Optimization subject to the Baseline Model.................................................... 26
       4.3.2 Optimization subject to Varying Effective Reproduction Rates....................... 27
       4.3.3 Optimization subject to Flexible Reserve System .......................................... 30

5 Discussion ..................................................................................................................... 33

Appendix A ..................................................................................................................... 35

Appendix B ..................................................................................................................... 37

Bibliography ................................................................................................................... 38
1 Introduction

The COVID-19 pandemic has brought into picture a new wave of academic insights on allocating scarce medical resources during public health crises. One of the major challenges of pandemic outbreaks is that medical resources are relatively inelastic in the short term, which cannot fulfill the sudden increase in demand for medical care. A guideline recommending resource prioritization criteria during COVID-19 points out that there were only 85,000 adult ICU beds, 62,000 full-featured ventilators, and 76,000 full-time respiratory therapists across the US during the pandemic, far less than what was needed to address all pandemic patients [1]. The situation demands that medical resources are allocated optimally to treat those who are most in need.

Studies have suggested that individuals with underlying health conditions have been disproportionately affected by the pandemic outbreak. On one hand, mortality rate from COVID-19 is significantly higher for patients with underlying health conditions than those without. A study evaluating the mortality rates of 566,602 patients with COVID-19 suggests that the mortality rate of patients with underlying health conditions was 12%, 4 times higher than that of patients without underlying health conditions [2]. Another survey-based study points out that more than 70% of physician consider comorbidities in diseases such as chronic respiratory failure, diabetes, obesity, etc., as important influencing factors on the COVID-19 prognosis [3].

On the other hand, as more resources tilt toward pandemic patients, non-pandemic patients with underlying health conditions lack sufficient medical attention when in need. While to the best of our knowledge, limited empirical study or data collection has been conducted to quantify the insufficiency of medical resources received by non-pandemic patients, the survey-based study reports that 72.7% of physicians believe that COVID-19 free patients are medically neglected during the pandemic outbreak [3]. Additionally, a study investigating the leading causes of excess mortality in Mexico during the COVID-19 pandemic outbreak suggests that 34.4% of excess mortality in Mexico were not directly associated with COVID-19 infection [4]. Cause-specific death tolls reveal that mortality resulting from major health conditions other than COVID-19 infections, including ischemic heart diseases, diabetes, and respiratory infections, all increased by more than 30% [4]. The study points out that these excess mortalities could be due to lack of timely access to needed, ongoing health care treatment [4].

Hence, this paper conjectures that during a pandemic outbreak, recognizing the medical needs from individuals with underlying health conditions (with or without pandemic disease
infection) could facilitate overall mortality reduction. We choose overall mortality reduction as the main objective of our model and simulations because it was ranked as one of the most preferred criteria for scarce medical resource by medical experts and the general public [1] [5]. However, it was observed that in certain medical systems, reducing pandemic-related mortality was prioritized whereas mortality from other causes was neglected, reflected through unreasonable allocation of resources between these two groups. We thus aim to explore what kind of medical resource allocation strategy could best fulfill the aforementioned objective of overall mortality reduction, taking into account mortalities from both pandemic and non-pandemic related causes.

To address the challenge on optimizing medical resource allocation for all patient categories, this paper adopts the absorbing Markov chain as a framework and builds upon it a dynamic model that simulates the different transition probabilities and death rates among different patient categories, subject to various resource allocation strategies as well as exogenous variables such as the transmission rate of the disease. The goal of the model is to find the optimal partitioning method for medical resources in terms of lowest overall mortality under different simulated scenarios. The paper will first construct a theoretical framework, then simulate the dynamic process based on Mexico data from 2020 using MATLAB, and finally manipulate some model parameters for corresponding outputs to draw qualitative conclusions and policy implications.

This paper finds that first, reserving resources for non-pandemic patients not only balances access to treatment but also reduces overall mortality rate from the observed 0.763% to 0.720% under the fixed reserve system. Further, a flexible reserve system that adapts to the increasing severity of a pandemic yields an even lower mortality rates at 0.661%, a significant improvement from the observed rate, which makes it a more favorable allocation strategy for policymakers. Integrating the findings with the multi-principle priority point system can also enhance the practical implications of this study. However, limitations of this study include the abstract nature of the model, the need for empirical validation of some model parameters, and ethical considerations surrounding resource allocation. It is thus imperative to conduct further research to explore these aspects thoroughly before implementing the practical implications derived from this study.
2 Literature Review

A lot of relevant literature arose between 2020 and 2022. Much of this body of work discuss how different types of medical resources should be allocated among COVID patients. A foundational work on this topic was proposed by the National Academies of Sciences, Engineering, and Medicine (NASEM) in 2020. The NASEM framework incorporates a traditional tiered priority system that categorizes individuals into different priority groups based on factors such as their risk of severe illness, occupation, and essential role in maintaining critical infrastructure [17]. By assigning individuals to specific tiers, the system helps guide decision-makers in prioritizing resource allocation with the aim of maximizing societal benefits, considering the unique needs and vulnerabilities of different populations [17].

A later study by Pathak et al. proposed modifying the tiered priority framework into a “reserve system” framework to build equity into the system [18]. In this system, medical resources would be partitioned into different reserve categories, which use different criteria for allocation of units. This enables a multi-principle priority point system as opposed to a single-principle point system widely discussed in previous literature. Patients are assigned a priority ranking in each reserve category, and they receive medical resource based on the set of priority rankings across all reserve categories [6]. The proposed theoretical system was then recommended or adopted in a number of practical settings including therapeutics and vaccine allocation.

The present paper is related to the reserve system in its design of the Markov chain framework. Instead of pooling all resources together and allocating them centrally, our Markov chain framework partitions patients into three categories – those only infected with the pandemic disease, those with other underlying health conditions, and those with both pandemic and non-pandemic related health conditions – and allocates resources to different patient categories from independent reserves. This is reflected in the framework by setting a constraint on the capacity of resources in each category and defining the transition probabilities in relation to the proportion of patients eligible for medical resource based on the capacity. The capacities stay fixed throughout the process to demonstrate a planned approach to resource allocation – it recognizes the needs of all patient categories from the very beginning and avoids pooling all resources into one specific category and over-straining the other patient categories in the long term.

The framework in this paper is a new extension from the reserve system in literature because it addresses a larger patient population. In previous literature, reserve categories and
priority rankings are designed for fairness purpose among patients with different demographics who are either infected with COVID or in need of protection from COVID [6]. This paper jumps out of the pandemic context and deals with the larger patient body. Some scholars published an article in 2022 and argued for a similar strategy, discussing how the reserve priority system may be used to ensure a basic level of access to medical resources (such as hospital beds) for non-COVID patients and thus facilitate fair distribution of scarce medical capacity between COVID and non-COVID patients [7]. However, this strategy was only proposed and has not been investigated on a deeper level. This paper would fill in the gap and explore resource allocation strategies that aims to preserve a certain level of fairness between the pandemic and non-pandemic patient body. As the overall mortality, direct and indirect from the pandemic, could be reduced, we hope that this study could shed light on how to improve the overall well-being of the patient population during a pandemic outbreak.

In terms of the theoretical methodology of the model framework, non-homogenous Markov chain has been used in a number of studies to simulate different aspects of a pandemic outbreak. A study conducted by Barraza et al. models the early growth stage of the COVID pandemic in Latin American countries based on non-homogenous Markov chain [8]. The study chooses immunization rate, the infection/immunization ratio, and the mean time between infections as the three key indicators to build the model upon and accomplishes a good fit between the model prediction and observed positive cases and deaths [8]. Stochastic continuous-time Markov chain has also been used by Al-Zoughool et al. to examine an alternative timing and duration of the COVID-19 lockdown in Kuwait [9]. The study employs a social contact matrix and a home contact matrix to capture the impact of lockdown in the model and concludes from the model an optimal lockdown time and duration for Kuwait, retrospectively. While the use of non-homogenous Markov chain in a pandemic setting is established in relevant literature, its use on modeling the impact of different medical resource allocation strategies and the different health outcomes in this paper would be a novel approach.
3 The Non-Homogenous Markov Chain Model

3.1 Setting up an Absorbing Markov Chain Framework

We use discrete-time, discrete-state-space, non-homogenous Markov chain to set up the framework of the pandemic outbreak model. The states of the Markov chain represent the different health conditions of the population at each time step. As we are interested in both the pandemic disease as well as other underlying health conditions, the states are characterized by the presence or absence of these two scenarios, defined as follows:

**State H** – where individuals are completely healthy with no illness that requires treatment.

**State C** – where individuals are only infected with the pandemic disease.

**State S** – where individuals have underlying health conditions that require treatment but are not infected with the pandemic disease.

**State B** – where individuals have both an infection of the pandemic disease and underlying health conditions.

**State R** – where patients recovered from the pandemic disease, with or without other underlying health conditions.

**State D** – where patients died from any health conditions (the pandemic disease or other illnesses).

Note that we constructed the Markov chain as an absorbing Markov chain where states R and D are absorbing states while the other four states are transient states. Absorbing states are defined as states where units once entered, cannot be left. In this case, it means that individuals cannot leave states R or D once entered. Hence, this model assumes that the system describes one infection cycle - patients who have been infected with the pandemic disease cannot be reinfected within the cycle. Moreover, it assumes that once the individuals recover from the pandemic disease, they don’t develop persisting symptoms or acquire new health conditions that require medical attention within the modeling period. While this is not always the case in real-life circumstances, for the sake of modeling, we assume its impact to be negligible. Figure 1 is a visualization of this absorbing Markov chain framework, marked with transition probability variables.
With this setup, depending on the stage of the pandemic we take as the starting point of our analysis, we would get an initial state vector with population distribution in states H, C, S, B, R, D. We thus define the initial state vector as

\[ \mathbf{K}_0 = \begin{bmatrix} H_0 & C_0 & S_0 & B_0 & R_0 & D_0 \end{bmatrix} \]

where \( H_0, C_0, S_0, B_0, R_0, D_0 \) represent the proportion of the total population in each state, and it follows that \( H_0 + C_0 + S_0 + B_0 + R_0 + D_0 = 1 \) by definition.

Following the transition relationships among the 6 states detailed in Figure 1, we define the transition matrix \( \mathbf{P} \) as a 6-by-6 matrix. The rows of \( \mathbf{P} \) represent sources, while the columns represent destinations. The sequence of states in both the row and the column is \( H, C, S, B, R, D \):

\[
\mathbf{P} = \begin{bmatrix}
    v_{HH} & v_{HC} & v_{HS} & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & f_{CR} & f_{CD} \\
    f_{SH} & 0 & f_{SS} & v_{SB} & 0 & f_{SD} \\
    0 & 0 & 0 & 0 & f_{BR} & f_{BD} \\
    0 & 0 & 0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

Note that we will further construct the Markov chain as non-homogenous in later sections. This means that unlike the constant transition probabilities in a classic Markov chain model, the probability variables within the transition matrix \( \mathbf{P} \) vary over time. They depend on some characteristics of the current state and time at which a specific transition is being made. Intuitively, for example, the transmission rate of the disease depends on the changing proportion of susceptible individuals in the population at each time step; the death rate of the population in a certain state depends on the capacity of available medical resources compared to the total population in that state at a given point in time. Non-homogenous Markov chain is common for modeling in an
epidemiological setting because it makes it easier to capture the changing dynamics of the pandemic such as availability of treatment, public health measures, etc. Non-homogeneity is crucial to the proposed framework as this paper aims to understand how the medical resource allocation strategies should change in response to the changing dynamics of a pandemic outbreak to yield the lowest overall mortality possible. Our approach to modeling the impact of medical resource allocation on health outcomes as well as how resource allocation, along with some other variables, create non-homogeneity in the transition probabilities will be discussed in detail in the next few sections.

3.2 Determining Health Outcomes through Resource Eligibility

Prior to a pandemic outbreak, all medical resources within a health system would be allocated to individuals in state S. However, with a disease outbreak, this equilibrium would be broken – the limited medical resources are allocated to serve individuals in states C, S and B. Many health systems have seen an overcrowding of medical resources during the pandemic due to a sudden increase in demand for medical care, and thus suffered from excess death due to resource insufficiency among the entire patient body – not only pertaining to patients infected with the pandemic disease but also those with underlying health conditions that require continuous medical treatment. Hence, it is important to understand how the absence of medical resources could affect the health conditions of patients in different states and in turn, find an optimal strategy to reduce mortality among the entire patient body. In this model, we account for the impact of resource eligibility on health outcomes through the following variables:

- $\alpha$ – the share of population in State C eligible for medical resources at each time step.
- $\beta$ – the share of population in State S eligible for medical resources at each time step.
- $\gamma$ – the share of population in State B eligible for medical resources at each time step.
- $\nu_{SB}$ – the share of population in State S who transitions to State B (infected with the pandemic disease) at each time step.
- $\rho_S$ – the recovery rate of population with underlying diseases given sufficient medical resources at each time step.
- $\lambda_C/\lambda_S/\lambda_B$ – the death rates of the eligible population at their corresponding states.
- $\theta_C/\theta_S/\theta_B$ – the multiples on death rates of the ineligible population at their corresponding states.
Note that variables $\alpha, \beta, \gamma, v_{SB}$ are time-dependent variables endogenous to the system, as they are expressed as proportions subject to the changing population of the states. For the sake of highlighting the different health outcomes associated with different eligibilities at each state, we only define the framework directly related to health outcomes in this section – further details on creating time-dependent functions for certain variables will be discussed later in section 3.3. On the other hand, $\rho_{S}, \lambda_{C}, \lambda_{S}, \lambda_{B}, \theta_{C}, \theta_{S}, \theta_{B}$ are all exogenous variables that do not vary over time. They collectively determine the death rates of different patient bodies with the presence or absence of medical resources. Intuitively, patients eligible for medical resources would take on a death rate of $\lambda$, while patients ineligible for medical resources will take on a death rate of $\theta \cdot \lambda$. Since we assume positive health outcome associated with resource eligibility, it follows that $\theta_{C}, \theta_{S}, \theta_{B} > 1$.

Figure 2 demonstrates the transition probabilities from states C, S, B that are subject to resource eligibility. It is important to highlight two assumptions associated with these transitions:

a) We assume that individuals who leave states C and B will only enter the absorbing states R or D and will not enter state S. Intuitively, this means that individuals do not get infected with another illness when they are in the middle of a single pandemic disease infection. This is a reasonable assumption for simplicity of the model because the turnaround time for pandemic infections is usually short, and it tends to be far more prevalent than other illnesses. This also means that for individuals infected with multiple illnesses including the pandemic disease, the model does not distinguish the cause of death for individuals entering State D from State B. This
is a reasonable setup because the goal of the model is to understand how to minimize overall mortality rate and it is not prone to this specific distinction as long as it is established that individuals are more at risk of death with multiple infections, i.e., \( \lambda_B > \lambda_S \).

b) We also assume that in state S, only eligible individuals can possibly re-enter state H. This is a reasonable assumption because it distinguishes the health outcomes of the eligible individuals from the ineligible individuals through the necessity of medical resources to recover. In real life scenarios, there could be cases where individuals self-recover without treatment. However, as these nuances fail to reflect the positive impact of resources, it would be reasonable to leave them out for simplicity.

With these assumptions in mind, by adding up the transition probabilities to each state, the \( f \)'s in the transition matrix as functions of the variables defined earlier in this section are as follows:

\[
\begin{align*}
    f_{CR} &= 1 - \lambda_C \alpha - \theta_C \lambda_C + \theta_C \lambda_C \alpha, \\
    f_{CD} &= \lambda_C \alpha + \theta_C \lambda_C - \theta_C \lambda_C \alpha, \\
    f_{SH} &= \rho_s \beta, \\
    f_{SD} &= \lambda_S \beta + \theta_S \lambda_S - \theta_S \lambda_S \beta, \\
    f_{SS} &= 1 - \rho_S \beta - \lambda_S \beta - \theta_S \lambda_S + \theta_S \lambda_S \beta - \nu_{SB}, \\
    f_{BR} &= 1 - \lambda_B \gamma - \theta_B \lambda_B + \theta_B \lambda_B \gamma, \\
    f_{BD} &= \lambda_B \gamma + \theta_B \lambda_B - \theta_B \lambda_B \gamma.
\end{align*}
\]

Note that these expressions are time-dependent, i.e., non-homogenous across each time step, since they all involve endogenous variables \( \alpha, \beta, \gamma, \nu_{SB} \). We should thus discuss the non-homogeneity of the variables in the transition matrix and how exactly they evolve over time in the next section.

### 3.3 Modeling Non-Homogeneity of the Transition Matrix

It has been briefly discussed in section 3.1 that we constructed the model framework based on non-homogenous Markov chain to capture the changing dynamics of a pandemic. Two main dynamic aspects of a pandemic outbreak that we want to capture in the model is:

a) The susceptibility of the general population

b) The share of eligible patients in each state given fixed resource pools

Intuitively, for the former aspect, since more and more individuals are infected and recovered at each time step, the transmission rate of the disease would slow down due to the increasing immunity from previous infection. For the latter aspect, the total population in each state would
change at each time step, and thus the share of patients eligible for a fixed capacity of resources would change accordingly. It is then evident that certain variables that determine the transition probabilities are dependent on the current distribution vector at a given point in time. In section 3.2 we defined the initial state vector as

$$K_0 = [H_0, C_0, S_0, B_0, R_0, D_0].$$

Hence, we define the current distribution vector at time $t$ as

$$K_t = [H_t, C_t, S_t, B_t, R_t, D_t],$$

where

$$K_t = K_{t-1} \cdot P_{t-1} = K_0 \cdot P_0 \cdot \ldots \cdot P_{t-1}.$$  

For the purpose of the larger model, we are most interested in the last element of the current distribution vector at the last stage of the modeling period – the value of $D_t$ represents the share of population who died in the modeling period, and we aim to minimize this value with optimal resource allocation. Nevertheless, at any time $t$, all elements in the current distribution vector $K_t$, along with some exogenous parameters, collectively determine some elements in the transition probability matrix at time $t + 1$ (denoted as $P_{t+1}$). We thus want to bring in some additional parameters and discuss time dependency in these two aspects.

### 3.3.1 The Changing Susceptibility of the Population

A key term in epidemiology to describe the contagiousness of an infectious disease is $r_0$, which stands for the basic reproduction rate. In simpler terms, this is the number of secondary infections that could result from one index case in a completely susceptible population. A term closely related to $r_0$ is $r_e$, which is defined as the effective reproduction rate. Essentially, this captures the role of public health measures such as case isolation and general contact reduction in the observed transmission rate of an infectious disease. With these measures, even in a completely susceptible population, we would observe the effective reproduction rate to be smaller than the basic reproduction rate defined in literature ($r_e < r_0$). Nevertheless, the transmission rate of an infectious disease could go further below $r_e$ – as the population becomes increasingly immune to the disease due to previous infections, $r_e$ would further decrease by the proportion of the immune population.

In the case of our model, we assume that the entire population is susceptible to the disease at the initial stage. This essentially means we assume $R_0 = 0$ in $K_0$. After at least one time step,
we would get a current distribution vector \( K_t \) where \( C_t + S_t + R_t \neq 0 \), and thus we get a proportion of the population immune to the disease within the modeling period. The proportion of the population still susceptible to the disease at time \( t \) would thus be given by \( \frac{H_t + S_t}{H_t + C_t + S_t + B_t + R_t} \). This would bring down the effective reproduction rate at time \( t \) to \( r_e \cdot \frac{H_t + S_t}{H_t + C_t + S_t + B_t + R_t} \). Assuming that the total population within the medical system is \( p \), we would then get the total number of contagious individuals at time \( t + 1 \) to be \( (C_t + B_t)p \cdot r_e \cdot \frac{H_t + S_t}{H_t + C_t + S_t + B_t + R_t} \). Therefore, the transition probability from susceptible states to infected states is given by 

\[
\frac{(C_t + B_t)p \cdot r_e \cdot \frac{H_t + S_t}{H_t + C_t + S_t + B_t + R_t}}{(H_t + S_t)p} = \frac{C_t + B_t}{H_t + C_t + S_t + B_t + R_t} \cdot r_e
\]

Note that in this model, we assume that there is no significant difference in isolation methods between the healthy population and those with underlying health conditions, and thus we take \( v_{HC} = v_{SB} \) at each time step. Hence, we define these two transition probabilities at time \( t + 1 \) as

\[
v_{HC} = v_{SB} = \frac{C_t + B_t}{H_t + C_t + S_t + B_t + R_t} \cdot r_e
\]

Given that \( P(1,2) = v_{HC} \), \( P(3,4) = v_{SB} \) (\( P(i,j) \) represent the element in the \( i^{th} \) row and \( j^{th} \) column in the transition matrix \( P \)),

\[
P_{t+1}(1,2) = P_{t+1}(3,4) = \frac{C_t + B_t}{H_t + C_t + S_t + B_t + R_t} \cdot r_e
\]

As this function is time-dependent, this makes up the first aspect of non-homogeneity in this pandemic outbreak model.

### 3.3.2 The Changing Share of Resource-Eligible Population

In section 2, we briefly mentioned that the resource management strategy in this paper is built upon the foundational work by Pathak et al. [6]. Essentially, we want to reserve a share of total resources for patients in states C, S, B respectively. Optimizing resource allocation for mortality reduction in this case would mean finding the optimal partitioning of resources to each of the categories so that these resource pools could yield the lowest overall mortality possible. As discussed previously in section 3.2, we define \( \alpha, \beta, \gamma \) as the proportion of the population in states C, S, B that are eligible for medical resources at a certain time step. Given that the total capacity of resources for each category is fixed over the modeling period, while number of individuals in
each category varies with each time step, this means that the eligibility variables $\alpha, \beta, \gamma$ are time-dependent. We want to define some additional parameters to understand the relationship between the eligibility variables and allocated resource capacity.

Let $m$ be the total unit of medical resources within a medical system. This could be defined as any type of resources with a capacity limit. For example, in the context of a pandemic outbreak, $m$ could be defined as the total number of medical professionals, total number of ventilators or hospital beds, etc.

Let $n_C, n_S, n_B$ be the number of individuals one unit of medical resource can satisfy subject to the duration of one time step in the model. We can further define $n_C, n_S, n_B$ using the following parameters:

$\mu_C/\mu_S/\mu_B$ – the percentage of patients in the corresponding category in need of the resource type.

d_C/d_S/d_B – the average duration of usage of the resource type for the corresponding category.

$\varphi_C/\varphi_S/\varphi_B$ – the number of patients this resource type can serve simultaneously.

Suppose the duration of one time step is denoted as $t_0$, we can calculate $n_C, n_S, n_B$ using formula:

$$n = \frac{\varphi \cdot t_0}{\mu \cdot d}$$

For example, let’s take the duration of one time step as 14 days ($t_0 = 14$). Suppose for state B, the hospitalization rate is 45%, the physician to patient ratio is 1:10, and an average patient stays hospitalized for 14 days. Then $n_B$ (with medical professionals being the resource type) is given by

$$n_B = \frac{10 \cdot 14}{0.45 \cdot 14} \approx 22$$

Finally, let $a, b, c$ be the share of total resource allocated to state C, S, B respectively. The total number of resource units allocated to each category would then be $am, bm, cm$. The total resource capacity for each category is in turn $amn_C, bmn_S, cmn_B$. $a, b, c$ would be variables set as constant from the initial stage of the Markov chain to reflect a fixed resource reserve for each patient category. These are also variables we want to optimize over to yield the lowest mortality rate possible at the end stage of the Markov chain.

The eligibility variables $\alpha, \beta, \gamma$ are determined collectively by the current distribution vector at each time step as well as the constant exogenous parameters defined above. For a given time step at time $t$, the eligibility variables are given by:

$$\alpha_t = \frac{a \cdot m \cdot n_C}{C_t \cdot p}$$
\[ \beta_t = \frac{b \cdot m \cdot n_S}{S_t \cdot p} \]

\[ \gamma_t = \frac{c \cdot m \cdot n_B}{B_t \cdot p} \]

These variables are essentially given by total capacity divided by total number of individuals in each category at a given time step. The eligibility variables at time \( t \) in turn determine certain elements in the transition matrix \( P \) at time \( t + 1 \). Intuitively, this means that resource eligibility at a certain time period impacts the health outcomes of the next time period. Indicated by formulas defined in section 3.2, \( P_{t+1} \) is dependent on \( \alpha_t, \beta_t, \gamma_t \), i.e., \( P_{t+1} \) is given by \( P_{t+1}(\alpha_t, \beta_t, \gamma_t) \).

To sum up, our dynamic model is defined by a set of exogenous variables (empirically given) and a set of endogenous variables (changing at each time step of the model). The goal of the model is to minimize overall mortality rate \( d \) at some fixed time \( t = T \), where \( T \) is the duration of the pandemic cycle that we choose as our modeling period. We would iteratively run the model for \( T \) times, retrieve \( d = D_T \) in the current distribution vector at time \( T \), and look for the optimal set of parameters or allocation methods that yield the lowest \( d \).
4 MATLAB Simulation

4.1 Methodology

In section 3, we discussed the setup of the Markov chain model and defined both exogenous parameters and time-dependent variables to construct the transition matrix. The goal of our analysis is to minimize $D_T$ (equal to overall mortality rate $d$) over a set of $a, b, c$ (partitioning of medical resources, defined in section 3.3.2) and find the optimal set of $a, b, c$ under different exogenous circumstances. Given that the Markov chain we constructed is non-homogenous, it would be difficult for us to navigate a mathematical solution to this optimization. Therefore, in this section, we transform this theoretical setup into a MATLAB simulation model to verify its validity and test the model under a realistic context. The three main parts to the simulation process are empirical setting, data calibration, and optimization, discussed as follows.

4.1.1 Empirical Setting

Given that the pandemic experiences in different medical systems differ from one another, it would be too big of a generalization to run the simulation on a global scale. Hence, we want to choose one type of medical resource in one specific medical system during the COVID-19 pandemic outbreak as an empirical setting to discuss the status quo and the optimization potential of medical resource allocation.

We choose the Mexico medical system during the COVID-19 pandemic outbreak as a sample case to run with our model, for two reasons. Firstly, Mexico is one of the seven countries in the world that experienced the most excess mortality during the COVID-19 pandemic outbreak in years 2020-2021. In 2020 alone, Mexico reported 325,415 excess deaths, 34.4% of which were not directly related to COVID-19 but rather due to the lack of access to medical treatment [4]. Running our model on this medical system could thus yield retrospective conclusions as to what would have been an optimal strategy to allocating resources to reduce excess mortality. Secondly, among all countries with available pandemic data, Mexico is one of the countries with the most comprehensive data collection and analysis on the excess deaths as well as the COVID death rates related to patients with underlying health conditions. These data enable us to take proxies that conform to real-life scenarios more accurately. Note that not all proxies in the input values (discussed later in section 4.2) are based on data in Mexico due to limited data availability, but the
key parameters related to death rates and country demographics are all based on Mexico data. We assume that the other proxies bear universal applicability and are appropriate to apply under the context of Mexico.

We set the timeframe of our model between March and December 2020, as the confirmed cases of COVID first accumulated to two-digits in early March, yielding approximately 44 weeks till the end of the year [10]. We define each time step in the Markov chain model as a 2-week interval, and thus we iteratively run the model for 22 time steps in every simulation, i.e., T=22. Note that input parameters to the model, such as death rates ($\lambda$), resource capacity ($n$), all elaborated later in section 4.2.1, are scaled for 14 days.

Furthermore, we choose hospital bed as the type of medical resource to simulate on. This is because hospital beds are usually the bottleneck that prevents patients from receiving necessary medical care in the sudden outbreak of a pandemic, as needs for hospitalization sharply increase. And since the demand for hospital beds could be reflected directly through hospitalization rate, parameters related to this type of resource are easier to proximate on compared to other resources such as ventilators and medical professionals, as they seem to involve more nuances in determining their demand from patients.

4.1.2 Data Calibration Approach

Most input parameters for the model could be approximated through empirical literature based on the COVID-19 pandemic outbreak either in Mexico or other medical systems. However, some model parameters such as death rates in absence of medical treatment, share of medical resources allocated to each patient category, average effective reproduction rate, etc., lack any empirical data collection or analysis. We thus use data calibration to estimate the values of the unknown parameters. This involves iteratively adjusting these unknown parameters within a reasonable range of values and comparing the simulated results with the observed outcome data. We then determine the baseline model based on the set of parameters that minimize the discrepancy between the simulated and observed results. The parameters and outcomes of the baseline model would then serve as the foundation and benchmark for further optimization analyses.
4.1.3 Optimization Approach

In the baseline model, we would have an estimated resource partitioning method (indicated by the calibrated parameters) that yields outputs closest to empirically observed outcomes. By running the model again based on the known parameters and calibrated parameters and iteratively testing on difference partitioning methods, we expect that we can get an optimal set of partitions that yields the lowest mortality rate possible through fixed reserves over the modeling period. We can then compare this optimal partition with the one in the baseline model to conclude qualitative policy implications in terms of how a mortality-minimizing counterfactual could be achieved in the given empirical circumstance.

Additionally, we also want to understand what the optimal resource allocation method could look like given changing circumstances. By altering some exogenous parameters from the baseline model to output different sets of optimal partitions, we expect the model could give insights on the best strategy to allocating resources in future pandemic outbreaks that bear different characteristics from the one we just experienced. For example, we will run the model with different effective reproduction rates ($r_e$), different total resource capacity, and different resource pooling methods (fixed / flexible reserves) to draw further conclusions and policy implications.

4.2 Data Calibration

4.2.1 Parameters and Assumptions

Following the methodology detailed in section 4.1, we first conduct data calibration to estimate the unknown parameters. The majority of the input parameters to the simulation model can be proxied through empirical literature, listed in Table 1. The empirical evidence and calculations leading up to these values can be found in detail in Appendix 1.

*Table 1 – Known Input Parameters to the Simulation Model*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$</td>
<td>Share of healthy population at the initial stage</td>
<td>0.624</td>
</tr>
<tr>
<td>$S_0$</td>
<td>Share of population with underlying health conditions at the initial stage</td>
<td>0.376</td>
</tr>
<tr>
<td>$\lambda_c$</td>
<td>Death rate of the pandemic patients given sufficient medical resource in 14 days</td>
<td>0.03</td>
</tr>
</tbody>
</table>
\[ \lambda_B \] Death rate of pandemic patients with underlying diseases given sufficient medical resource in 14 days 0.12

\[ \lambda_S \] Death rate of patients with underlying diseases given sufficient medical resource in 14 days 0.0006

\( m \) Total units of medical resources (hospital beds) 126,000

\( p \) Total population 126,000,000

\( n_C \) The number of individuals one unit of medical resource can satisfy in state C in 14 days (subject to probability of hospitalization) 14

\( n_B \) The number of individuals one unit of medical resource can satisfy in state B in 14 days (subject to probability of hospitalization) 3

\( n_S \) The number of individuals one unit of medical resource can satisfy in state S in 14 days (subject to probability of hospitalization) 400

\( v_{HC} \) Share of population in state H infected with the pandemic disease from \( t = 0 \) to \( t = 1 \) 0.000002

\( v_{SB} \) Share of population in state S infected with the pandemic disease from \( t = 0 \) to \( t = 1 \) 0.000002

On the other hand, there are a couple of input parameters to the model that lack empirical evidence for us to reach a proximate value. Therefore, we need to calibrate these parameters within a reasonable range of values so that we can get the best fit between the simulation output and the observed data. Table 2 is the list of unknown parameters along with the range of values that we want to calibrate them for. The justifications for the value range can also be found in Appendix A.

\textit{Table 2 – Unknown Input Parameters to the Simulation Model and Their Calibration Value Range}

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_e )</td>
<td>Average effective reproduction rate</td>
<td>[1.1, 3.0], 0.1 increment</td>
</tr>
</tbody>
</table>
Note that there are three important assumptions associated with the calibration process:

**Assumption a. Sufficient resources are allocated to COVID-19 patients in Mexico in 2020.**

As reported by a study on excess mortality in Mexico in 2020-2021, more than 80% of COVID-19 deaths (82.9%) occurred in medical units [4]. This is much higher than the percentage of all-cause deaths that occurred in medical units (45.1%) during the pre-pandemic years 2015-2019 [4]. Moreover, during the pandemic outbreak, only 32.8% of the deaths that occurred due to remaining causes other than COVID-19 took place in the medical units [4]. Although no evidence is available to indicate the actual resource distribution strategy in Mexico in 2020, the large difference in the places of death between COVID and non-COVID patients demonstrates an imbalance of resource allocation among the two patient categories. To differentiate how the eligibility of resources leads to different health outcomes between the two categories, it is a reasonable assumption to hold that sufficient resources are allocated to COVID-19 patients in the baseline simulation. In other words, we assume all patients with a COVID-19 infection are eligible for hospital beds, while patients with other health conditions might turn out ineligible for hospital care due to shortage of resources in the baseline model.

This assumption has two implications regarding the data calibration. Firstly, we are only able to calibrate for the proportion of medical resources allocated to patients in state S (denoted by \( b \)), but not the proportions allocated to states C and B (denoted by \( a \) and \( c \)). This is because sufficient resources for COVID patients indicate that the associated health outcomes do not differ among all possible allocation strategies of \( a, c \) that ensures resource sufficiency, given the fixed reserve assumption (will discuss in Assumption b). On the contrary, for different values of \( b \),
eligibility of patients in state S would differ, leading to different health outcomes and thus flexibility in output values. This makes it possible to find the best fit $b$ that yields the closest output to observed values. Secondly, since we assume that all patients with COVID-19 received sufficient medical treatment, we would not be able to calibrate for $\theta_C$ and $\theta_B$, which stand for the multiples on death rates for patients with COVID-19 (with / without underlying health conditions) in absence of medical treatment. Given limited empirical evidence to rely on to find proxies on these two parameters, we take $\theta_C = 1.2$ and $\theta_B = 2.5$ as reasonable assumptions to these two values moving forward. For the purpose of drawing qualitative conclusions, it is sufficient to set two reasonable values that satisfy $\theta_C < \theta_S < \theta_B$, reflecting the different levels of demand for medical treatment in seeking recovery under these three conditions. However, for future research, it would be helpful to conduct empirical data collection and research in this aspect to reach more accurate predictions on health outcomes.

**Assumption b.** *Mexico adopted the fixed resource reserve system during the outbreak in 2020.*

As mentioned in section 2, this paper is built upon a framework discussing how fixed reserve systems could facilitate fairness in distributing resources among individuals with different demographics [6]. This paper aims to expand the application of this concept and explore an optimal method in setting separate reserves for pandemic and non-pandemic patients, in an effort to reduce overall mortality. As discussed in *Assumption a*, it is empirically observed that a disproportionately higher percentage of COVID deaths occurred in medical units in comparison to deaths by other causes. This is consistent with practices where COVID infections are prioritized for treatment regardless of the severity of patients’ symptoms, leading to resource overcrowding for non-pandemic patients demanding medical attention. Hence, it is reasonable to assume a fixed reserve system heavily tilting toward pandemic patients for Mexico during the outbreak, which is characterized by reserving resources to suffice pandemic needs, and potentially neglecting the needs from other medical conditions. While the medical system in Mexico did not strictly apply reserve system in allocating medical resources, holding this assumption could capture the essence of prioritizing pandemic needs and make it easier to analyze optimal partitioning method under the reserve system framework in later analyses.

**Assumption c.** *Before the pandemic, medical resources are saturated and in equilibrium for the population in Mexico.*
We assume that medical resources are saturated before the pandemic. Although medical resources could have been insufficient even in absence of the pandemic, we are more interested in how the pandemic has contributed to the excess deaths in comparison to deaths in a normal year. Hence, for simplicity, we assume that in a normal year without a pandemic, $\beta = 1$. We thus approximate the value of $n_s$ by taking the average number of sick individuals that one unit of medical resource serves in a normal year, associated with a death rate in 14 days scaled from the annual natural death rate due to illnesses (see Table 1 and Appendix A).

In addition, we assume that in the absence of a pandemic, at any given point in time, the share of population in state S is constant (i.e., in equilibrium). This means that the number of individuals transitioning between state H and state S for any given time period should be equal. Given the previous assumption that $\beta = 1$, the transition probabilities between state H and state S should satisfy:

$$v_{HS} \cdot H_0 = \rho_s \cdot S_0$$

As we calibrate for $v_{HS}$, the value of $\rho_s$ should follow directly from this expression above ($H_0$ and $S_0$ are known parameters). Note that when we are modeling the case of a pandemic, if $\beta \neq 1$, we would have $v_{HS} \cdot H_0 > \rho_s \cdot \beta \cdot S_0$. This simulates the accumulation effect of the demand for medical resources from patients with underlying health conditions. Intuitively, when certain patients fail to get treated immediately when they need medical attention, their demands for treatment do not disappear but rather transfer to the next time period in the forms of, for instance, more severe symptoms or prolonged treatment cycle. In essence, the calibration process is calibrating for a value of $\rho_s$ that most accurately captures this accumulation effect.

### 4.2.2 The Baseline Model and Interpretations

Using the known parameters specified in Table 1, we calibrate for the unknown parameters within the value range and increment specified in Table 2. The calibration process involves testing among all combinations of the 20 possible values for each of the four parameters, for a total of 160,000 combinations, and choosing the best-fit combination using the least square method. Specifically, we choose $d_C$, $d_S$, and $d_B$ as the outputs to calibrate on, shown in Table 3. The third column shows the empirically observed values of these three outputs (see Appendix A for justifications), while the fourth column shows the predicted outputs from the best-fit combination of the four unknown parameters. Table 4 shows the final calibrated values of the unknown
parameters, which yield the outputs shown in Table 3 with the least sum of square differences ($S = 1.884 \times 10^{-8}$) among all combinations.

### Table 3 – Calibration Results: Observed and Predicted Output Values

<table>
<thead>
<tr>
<th>Output</th>
<th>Definition</th>
<th>Observed</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_c$</td>
<td>Proportion of population who died of COVID</td>
<td>0.05%</td>
<td>0.037%</td>
</tr>
<tr>
<td>$d_s$</td>
<td>Proportion of population who died of underlying health conditions</td>
<td>0.61%</td>
<td>0.609%</td>
</tr>
<tr>
<td>$d_B$</td>
<td>Proportion of population who died of both diseases</td>
<td>0.11%</td>
<td>0.116%</td>
</tr>
<tr>
<td>$d$</td>
<td>Overall proportion of population who died of diseases</td>
<td>0.77%</td>
<td>0.763%</td>
</tr>
</tbody>
</table>

*d* not used in the calibration process, shown here for reference only

### Table 4 – Best-fit Calibrated Values of the Unknown Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Calibrated Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_e$</td>
<td>Average effective reproduction rate</td>
<td>1.5</td>
</tr>
<tr>
<td>$v_{HS}$</td>
<td>Share of healthy population diagnosed with an underlying health condition or those already diagnosed requiring continuous medical treatment in a 14-day period</td>
<td>0.005</td>
</tr>
<tr>
<td>$\theta_S$</td>
<td>Multiple on death rate for individuals with underlying health conditions in absence of treatment</td>
<td>1.4</td>
</tr>
<tr>
<td>$b$</td>
<td>Share of medical resources allocated to patients in state S in the baseline model</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Since the model is a high-level abstraction of the nuanced realistic circumstances, the calibrated values bear limited empirical significance and should not be taken alone as evidence to derive factual implications. However, we can interpret the calibrated results as follows: in the context of this non-homogenous Markov chain model alone, with the empirically known parameters as inputs, it is most consistent with empirically observed health outcomes that:
a) The COVID-19 virus had an average effective reproduction rate \( (r_e) \) of 1.5 in Mexico in 2020. This means that on average, one confirmed index case of COVID-19 transmits the disease to 1.5 people during its infection cycle. This is a smaller number than the empirically recognized basic reproductive number \( (r_0) \) of around 3 [11], which corroborates Mexico’s efforts in slowing down disease transmission through public health measures such as physical distancing [12].

b) The share of individuals in state S would increase by \( v_{HS} \cdot H_0 = 0.005 \cdot 0.624 = 0.00312 \) if no medical resource is allocated to state S \( (\beta = 0) \) within a 14-day period. This indicates that the medical resources demanded by this patient category would increase by \( 0.00312 \div 0.376 = 0.8\% \) in the next time period compared to the initial stage if no medical resource is reserved to support their needs.

c) Patients with underlying health conditions who did not get infected with COVID-19 were 1.4 times more likely to die if they were ineligible for medical resources. This would yield a death rate of \( \lambda_S \cdot \theta_S = 0.0006 \cdot 1.4 = 0.00084 \) over every 14-day period for this group of individuals and thus contribute to excess deaths if insufficient medical resources are allocated. This is a reasonable value as it satisfies our previous assumption in section 4.2.1 that \( \theta_c < \theta_S < \theta_B \).

d) On average, only half of the medical resources are allocated to non-COVID patient during the modeling period. This implies that half of the patient body with underlying health conditions were not treated properly and thus experienced higher risk of death due to a shift of resources to treating COVID patients.

The known parameters, the calibrated parameters, and the predicted outputs based on these parameters collectively make up the baseline model, summarized in Table 5. We will navigate approaches to minimize mortality compared to this baseline model in the next section.

**Table 5 – The Baseline Model**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0 )</td>
<td>0.624</td>
<td>( v_{HC} )</td>
<td>0.000002</td>
</tr>
<tr>
<td>( S_0 )</td>
<td>0.376</td>
<td>( v_{SB} )</td>
<td>0.000002</td>
</tr>
<tr>
<td>( \lambda_c )</td>
<td>0.03</td>
<td>( r_e )</td>
<td>1.5</td>
</tr>
<tr>
<td>( \lambda_B )</td>
<td>0.12</td>
<td>( v_{HS} )</td>
<td>0.005</td>
</tr>
<tr>
<td>( \lambda_S )</td>
<td>0.0006</td>
<td>( \theta_S )</td>
<td>1.4</td>
</tr>
<tr>
<td>$m$</td>
<td>126,000</td>
<td>$b$</td>
<td>0.5</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>$p$</td>
<td>126,000,000</td>
<td>$d_c$</td>
<td>0.037%</td>
</tr>
<tr>
<td>$n_c$</td>
<td>14</td>
<td>$d_S$</td>
<td>0.609%</td>
</tr>
<tr>
<td>$n_B$</td>
<td>3</td>
<td>$d_B$</td>
<td>0.116%</td>
</tr>
<tr>
<td>$n_S$</td>
<td>400</td>
<td>$d$</td>
<td>0.763%</td>
</tr>
</tbody>
</table>

4.3 Simulation-Based Optimization

To find out what partition of medical resources could minimize mortality rate under different circumstances, we iteratively run the model with different combinations of $a, b, c$ in 0.01 increments, for a total of 5050 combinations, and select the combination with the lowest overall mortality rate to be the optimal partition under the given circumstance. In the following subsections, we first navigate the optimal partition under the fixed parameters in the baseline model. This would retrospectively inform us of how excess deaths in Mexico could have been mitigated using optimal allocation method under fixed reserves. Next, we relax two fixed assumptions in the baseline model to explore how the optimal partition evolves in response to flexibility in those assumptions. These assumptions include fixed effective reproduction rate and fixed reserve system.

4.3.1 Optimization subject to the Baseline Model

Using the method described above, we run the simulation model using parameters specified in Table 5 (except for $b$) and tune for the optimal set of $a, b, c$. A comparison between the results from the baseline model and the optimized model is shown below in Table 6.

Table 6 – Baseline vs. Optimized Resource Partition and Mortality Rate Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline</th>
<th>Optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>$b$</td>
<td>0.5</td>
<td>0.89</td>
</tr>
<tr>
<td>$c$</td>
<td>0.4</td>
<td>0.11</td>
</tr>
<tr>
<td>$d_c$</td>
<td>0.037%</td>
<td>0.044%</td>
</tr>
<tr>
<td>$d_S$</td>
<td>0.609%</td>
<td>0.504%</td>
</tr>
<tr>
<td>$d_B$</td>
<td>0.116%</td>
<td>0.173%</td>
</tr>
<tr>
<td>$d$</td>
<td>0.763%</td>
<td>0.720%</td>
</tr>
</tbody>
</table>
*The values of $a$ and $c$ in the baseline column is one possible set of partitions that satisfy resource sufficiency for all COVID patients, assumed in the baseline model. They are not the only possible set and are shown here for the sole purpose of comparing with the optimized partitions.

The results from the table indicate that the optimized solution under the fixed reserve assumption partitions medical resources by a ratio of 89:11 between non-COVID and COVID patients, contrasting the 50:50 partitioning ratio in the baseline model. Additionally, resources should be prioritized for patients with underlying health conditions in the COVID patient category as the optimal partition yields $a = 0$ and $c = 0.11$. Under this partitioning method, the predicted overall mortality rate drops to 0.720%, a roughly 0.04 percentage point decrease from the baseline mortality rate (0.763%).

It is worth noting that such an allocation method leads to an increase in mortality from COVID patients – the mortality rate occurring in state C increases from 0.040% to 0.044%, and the mortality rate occurring in state B increases from 0.12% to 0.17%. However, as the mortality rate occurring in state S decreases by 1 percentage point from 0.61% to 0.50%, the positive impact of this strategy on excess deaths due to other health conditions seem to outweigh the increase in mortality due to the pandemic disease.

The results imply that retrospectively speaking, if the medical system in Mexico seeks overall mortality reduction as their main objective in policy setting during the pandemic, they should reserve more medical resources for non-COVID patients and prioritize COVID patients with underlying health conditions for the remaining resources. This strategy differs drastically from the strategy employed by some countries where public health policies prioritize reducing mortality directly related to the pandemic disease and thus allocate as much medical resources as possible to the pandemic patients. The simulated outcomes suggest that by reasonably shifting focus from the pandemic and reserving part of the resources for non-pandemic patients, the decrease in excess deaths from patients in state S could potentially outweigh the negative impact of resource insufficiency among pandemic patients.

4.3.2 Optimization subject to Varying Effective Reproduction Rates

In previous analyses, we consider the effective reproduction rate ($r_e$) as a fixed parameter in the baseline model that reflects the average transmission rate of the pandemic disease over the modeling period. However, unlike parameters such as total resource capacity that remain rather inelastic in face of a sudden pandemic outbreak, the effective reproduction rate is not solely
determined by the type of the virus prevalent, but also collectively determined by public health measures in mitigating means of transmission. Hence, it is valuable to explore how optimal partition strategy could vary to match varying levels of effort in public health measures toward disease prevention.

Applying the same iterative process of testing for 5050 combinations of \(a, b, c\) to each value of \(r_e\) between 1.0 and 3.0 in 0.1 increments, we derive the optimal partition of resources as a function of the effective reproduction rate, shown in Figure 3. We further simulate the death rates by patient category as well as the overall death rates under each optimized scenario and demonstrate these death rates as functions of effective reproduction rates, shown in Figure 4.
From Figure 3, we can see that the optimal resource partition allocated to patients in state S (non-pandemic patients with underlying health conditions) and state B (pandemic patients with underlying health conditions) is non-monotone in the rate. That is, as average effective reproduction rate increases, optimal resource partition allocated to pandemic patients first increases, then decreases. This could be explained by two opposing forces. On one hand, as the disease becomes increasingly contagious, patients become more likely to transition from state S to state B, resulting in much higher rate of death if too many patients in state B remain untreated, jacking up overall mortality. This would make the case for shifting more resources toward state-B patients. On the other hand, higher disease transmission rate would amplify the negative health consequence associated with the accumulation effect (discussed in section 4.2.1 Assumption c). Intuitively, if insufficient resource for treating patients in state S leaves a large number of patients more vulnerable to the pandemic disease infection due to exacerbating pre-existing conditions (and thus at larger risk of death once infected), it could potentially be more resource efficient to shift resources to state S to prevent further compromising this group of patients. The non-monotone function in the graph indicates that the former force prevails the latter at lower transmission rates, up until $r_e = 1.6$. For higher transmission rates than this threshold, the latter force prevails the former.

Note that the optimal resource partitioning remains remarkably constant for effective reproduction rates between 2.0 and 3.0. This is because at higher transmission rates, the spread of the disease is characterized by a relatively slow build-up followed by a sharp peak of infections, after which the vast majority of people in the population have been infected. As we simulate the process over a relatively long timeframe of 44 weeks and assume fixed reserve for the entire timeframe, the inflexibility of the resource reserves necessitates that the allocation method prioritizes patients in state S over those in state B and state C. Such a disadvantage of an inflexible reserve system under the fast-changing dynamic of a pandemic outbreak is also demonstrated through Figure 4 – we can see that the mortality rate of pandemic patients (denoted by the orange line and the purple line) increases sharply as the effective reproduction rate increases, whereas the mortality rate of non-pandemic patients (denoted by the green line) appears to be a slightly decreasing function, given that most patients in that category get infected with the pandemic disease over the modeling period and are attributed to the pandemic category. Hence, in the next section, we seek to explore whether a flexible reserve system, as opposed to fixed reserve, could
further reduce overall mortality while allowing resource allocation to shift focus in response to the dynamic development of a pandemic.

4.3.3 Optimization subject to Flexible Reserve System

In this section, we further relax the assumption of fixed reserve and adopt a flexible reserve system framework. This means that we allow one switching point over the modeling period and derive different optimal resource partitions before and after the switching point to yield the lowest overall mortality rate possible. For each value of \( r_e \), we simulate results from switching at every possible time step \( t \) (\( t \in [1,21] \)) and choose the optimal switching point \( t^* \) that yields the lowest mortality rate under optimal resource partitions before and after \( t^* \). The optimal switching points, optimal partitions, and predicted mortality rates corresponding to each value of \( r_e \) between 1 and 3 (in 0.1 increments) can be found in Appendix B.

We first examine the baseline case where \( r_e = 1.5 \). The simulation gives us an optimal switching point \( t^* = 18 \). Table 7 shows the optimal partitions and associated death rates under fixed reserve and flexible reserve, in comparison to the baseline scenario. We denote the optimal resource partitions before the optimal switching point as \( a_1, b_1, c_1 \), and the optimal partitions after the optimal switching point as \( a_2, b_2, c_2 \).

Table 7 – Optimal Partitions and Death Rates by Category under Different Allocation Strategies at Baseline \( r_e = 1.5 \)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline</th>
<th>Optimal Fixed Reserve</th>
<th>Flexible Reserve (before ( t^* ))</th>
<th>Flexible Reserve (after ( t^* ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( b )</td>
<td>0.5</td>
<td>0.89</td>
<td>0.93</td>
<td>0.19</td>
</tr>
<tr>
<td>( c )</td>
<td>0.4</td>
<td>0.11</td>
<td>0.07</td>
<td>0.81</td>
</tr>
<tr>
<td>( d_C )</td>
<td>0.037%</td>
<td>0.044%</td>
<td>0.044%</td>
<td></td>
</tr>
<tr>
<td>( d_S )</td>
<td>0.609%</td>
<td>0.504%</td>
<td>0.522%</td>
<td></td>
</tr>
<tr>
<td>( d_B )</td>
<td>0.116%</td>
<td>0.173%</td>
<td>0.095%</td>
<td></td>
</tr>
<tr>
<td>( d )</td>
<td>0.763%</td>
<td>0.720%</td>
<td>0.661%</td>
<td></td>
</tr>
</tbody>
</table>

*The values of \( a \) and \( c \) in the baseline column is one possible set of partitions that satisfy resource sufficiency for all COVID patients, assumed in the baseline model. They are not the only possible set and are shown here for the sole purpose of comparing with the optimized partitions.
From the last two columns of Table 7, we can see that under flexible reserve, it is optimal to reserve the majority of resources for non-COVID patients during the early stage of the pandemic \((b_1 = 0.93, \ c_1 = 0.07)\), and then shift resources to COVID patients when the disease begins to transmit among a larger share of the population \((b_2 = 0.19, \ c_2 = 0.81)\). An optimal switching point \(t^* = 18\) means that it is mortality-minimizing to switch resource partitioning at Week 36 of the modeling period, which translates to early November 2020 in empirical context of Mexico. This corroborates the observed spike of COVID confirmed cases between early November and late December in Mexico in 2020 [10].

With regards to the predicted mortality rate by patient category, given minimization on overall mortality rate, the mortality occurring in state C remains the same under flexible reserve as under optimal fixed reserve \((d_C = 0.044\%)\), both slightly higher than their counterpart in the baseline scenario \((0.037\%)\). The mortality occurring in state S under flexible reserve lies in between the other two scenarios at \(d_S = 0.522\%\), lower than the baseline scenario \((0.609\%)\) and slightly higher than the optimal fixed reserve scenario \((0.504\%)\). We also see that mortality occurring in state B under flexible reserve is lower than both the baseline and optimal fixed reserve scenario \((0.116\% / 0.173\%)\), yielding \(d_B = 0.095\%\). The overall mortality rate further drops from 0.720\% under the optimal fixed reserve strategy to 0.661\% under flexible reserve, a 0.06 percentage point decrease. It is worth highlighting that this is also a significant improvement from the baseline model, with 0.1 percentage point decrease in overall death rate.

The results indicate that as a pandemic outbreak progresses, given a shortage of medical resources, overall mortality rate may be further reduced by adopting flexible reserve and switching resource partitioning from non-pandemic tilted to pandemic tilted. The mortality rate predictions also suggest that the flexible reserve strategy could significantly reduce mortality from severe pandemic disease infections (state B) without excessively hurting non-pandemic patients, as the mortality rate from state S only slightly increases. This is a more ideal allocation strategy than fixed reserve since it counters the disadvantage discussed at the end of section 4.3.2 where reserving too many resources for non-pandemic patients, especially in the case of high disease transmission rate, could result in failure to adjust for sudden spike in pandemic infections at a later stage of the pandemic and disproportionately hurt pandemic patients.
This conclusion is further validated by examining the mortality rates given optimal partitioning under flexible reserves versus fixed reserves, subject to varying effective reproduction rates, shown in Figure 5. Mortality rates from state C under the two strategies roughly overlap. Mortality rates from state S under flexible reserve always lies slightly above the line representing fixed reserve. Mortality rates from state B under flexible reserve always lies below the line representing fixed reserve. And finally, the overall mortality rate is always lower under the flexible reserve strategy compared to the fixed reserve strategy for any value of the effective reproduction rate. This means that regardless of the transmission rate of the disease, it is more ideal to adopt the flexible reserve strategy over the fixed reserve strategy in order to reduce overall mortality rate.
5 Discussion

This study establishes a theoretical framework based on non-homogenous Markov chain to simulate the dynamics of a pandemic outbreak and to predict mortality rates subject to different resource allocation strategies. To ensure that the pandemic response does not overwhelm other medical aims [7] and to reduce overall mortality among both pandemic patients and non-pandemic patients, our Markov chain framework categorizes individuals based on the presence or absence of the pandemic disease infections as well as underlying health conditions, and models the transitions between these categories with distinct, time-dependent probabilities toward different health outcomes (recovery / death). We then calibrate the model parameters under the empirical setting of Mexico’s pandemic outbreak in 2020 for a baseline model, and run simulations with the model to conclude insights on optimal medical resource allocation.

Our conclusions from the simulation-based analyses are detailed in the previous sections, but there are two important findings we want to highlight as key takeaways from this study:

First, as we compare the predicted health outcomes from the optimal fixed reserve strategy with the observed outcomes in the baseline scenario, besides noticing a smaller difference in mortality rates between pandemic and non-pandemic patients, we also see a lower overall mortality rate from the optimal fixed reserve strategy. This implies that reserving a subset of resources for non-pandemic patients is not only favorable in terms of balancing the access to medical treatment between different patient groups, but also, it should be preferred on the sole basis of minimizing overall mortality.

Second, while a fixed reserve mechanism could be a basic strategy to start with, it is not the most ideal given the fast-changing dynamic of a pandemic outbreak. A flexible reserve system where the partition of reserves switches at an optimal point from non-pandemic tilted to pandemic tilted yields even lower mortality rates at varying levels of disease transmission rates. This implies that policymakers should refrain from putting all resources into combating the pandemic at the early stage of an outbreak. Rather, they should always reserve a portion of the resources for non-pandemic patients, and only shift those reserves toward pandemic purposes when the pandemic progresses to a certain degree where infections start to spike and demands from pandemic patients become utterly significant. By taking in parameters that align with an empirical scenario, our model can also shed light on the optimal switching point using the approach from section 4.3.3.
Additionally, in a practical setting, the conclusions from this study can potentially be adopted in conjunction with the allocation strategy proposed in Pathak et al. [6]. Specifically, in this study, we conclude that subsets of medical resources should be reserved for pandemic patients with underlying health conditions (state B) and non-pandemic patients with underlying health conditions (state S). Pandemic patients without any pre-existing conditions (state C) are de-prioritized, as optimal partitions examined in this study assigns close to zero share to this patient category under most scenarios. This is because we artificially constructed the risk of death for patients in state C to be lower than the other two categories. In reality, of course, we cannot definitively assume that all patients in state C experience mild symptoms and do not require as much medical attention as those in the other two states. It is still standard practice to follow a baseline priority order of treating patients requiring more urgent medical attention. Hence, to adapt our conclusions into a more practical strategy, we can adopt the multi-principle priority point system proposed by Pathak et al. [6] – that is, we set up two reserve systems, one characterized by underlying health conditions and the other characterized by pandemic infections. Within each reserve, patients with the corresponding characteristic are assigned priority rankings based on the severity of their conditions (the baseline priority order). Patients then receive medical resource based on their higher ranking between the two priority lists, given that the higher ranking lies above the cutoff equilibrium. Under this setting, our results on optimal partitions, i.e., \( b \) and \( c \), can be interpreted as the size of these two reserves and thus serve as a benchmark in determining the cutoff equilibrium.

Limitations to this study are manifold. Firstly, this model is highly abstract with limited consideration toward practical nuances. Although we can draw qualitative conclusions on potential policy implications based on the simulations, these conclusions are not empirically proven in practice. Policymakers should use caution and collect more comprehensive empirical data to validate the model before implementing the strategies. Moreover, this study is built upon the objective to reduce overall mortality. Although we mentioned earlier that this is one of the preferred objectives for both medical professionals and the public [1][5], there are always controversies around the prioritization criteria when the eligibility of one group of patients entails compromise from another group. Hence, the ethical ground for the reserve mechanism must be fully discussed and explored before such a mechanism is deployed.
Appendix A – Justifications for the Parameter Values used in Model Calibration

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value (Range)</th>
<th>Justification</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$</td>
<td>0.624</td>
<td>Proportion of individuals without underlying health conditions among 566,602 COVID patients in data from the Mexican government. Assuming no significant difference in probability of COVID infection among all people, we generalized this proportion to the entire population.</td>
<td>[2]</td>
</tr>
<tr>
<td>$S_0$</td>
<td>0.376</td>
<td>Proportion of individuals with at least one underlying health conditions among 566,602 COVID patients in data from the Mexican government. Assuming no significant difference in probability of COVID infection among all people, we generalized this proportion to the entire population.</td>
<td>[2]</td>
</tr>
<tr>
<td>$\lambda_C$</td>
<td>0.03</td>
<td>Study reports probability of death due to COVID given no underlying health conditions to be 3.02%. For simplicity, we approximate this number to 0.03.</td>
<td>[2]</td>
</tr>
<tr>
<td>$\lambda_B$</td>
<td>0.12</td>
<td>Study reports probability of death due to COVID given underlying health conditions to be 11.97%. For simplicity, we approximate this number to 0.12.</td>
<td>[2]</td>
</tr>
<tr>
<td>$\lambda_S$</td>
<td>0.0006</td>
<td>Study reports total deaths from all diseases in Mexico to be 731,250 in 2019 (pre-pandemic), approximately 0.58% of the total population. This translates to 1.45% of the population in $S_0$, and scaling to 0.0006 of that population per 14 days.</td>
<td>[6][13]</td>
</tr>
<tr>
<td>$m$</td>
<td>126,000</td>
<td>Number of hospital beds in Mexico reached 126,449 in 2020. For simplicity of the model, we approximate this number to 126,000.</td>
<td>[14]</td>
</tr>
<tr>
<td>$p$</td>
<td>126,000,000</td>
<td>The population in Mexico was 125,998,302 in 2020. For simplicity of the model, we approximate this number to 126,000,000.</td>
<td>[13]</td>
</tr>
<tr>
<td>$n_C$</td>
<td>14</td>
<td>Hospitalization rate of COVID patients without underlying health conditions is around 7%. Assuming average hospitalization duration to be 14 days, 1 hospital bed can satisfy the need of approximately 14 people in state C.</td>
<td>[15][16]</td>
</tr>
<tr>
<td>$n_B$</td>
<td>3</td>
<td>Hospitalization rate of COVID patients with underlying health conditions is around 40%. Assuming average hospitalization duration to be 14 days, 1 hospital bed can satisfy the need of approximately 3 people in state B.</td>
<td>[15][16]</td>
</tr>
<tr>
<td>$n_S$</td>
<td>400</td>
<td>Given Assumption c in section 4.2.1, we assume this number to be the average number of people in state S that 1 hospital bed serves in pre-pandemic times. This gives us 126,000,000*0.04/126,000 = 400.</td>
<td>[13][14]</td>
</tr>
<tr>
<td>$\nu_{HC}$</td>
<td>0.000002</td>
<td>Mexico had 228 confirmed COVID cases in the first two weeks of March 2020, which yields to around 0.0002% of the population. We assume no significant difference in probability of infection between state H and state S, so we take this probability for both $\nu_{HC}$ and $\nu_{SB}$.</td>
<td>[10]</td>
</tr>
<tr>
<td>$\nu_{SB}$</td>
<td>0.000002</td>
<td>Mexico had 228 confirmed COVID cases in the first two weeks of March 2020, which yields to around 0.0002% of the population. We assume no significant difference in probability of infection between state H and state S, so we take this probability for both $\nu_{HC}$ and $\nu_{SB}$.</td>
<td>[10]</td>
</tr>
<tr>
<td>$\tau_e$</td>
<td>[1.1, 3.0]</td>
<td>Different studies give different estimates on the basic/effective reproduction rate of COVID, but a</td>
<td>[11][16]</td>
</tr>
</tbody>
</table>
A reasonable range for effective reproduction rate over a 14-day interval, especially given quarantine measures, would be below 3.0 and above 1.0, taking reference from multiple relevant studies.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v_{HS} )</td>
<td>([0.005, 0.1])</td>
<td>This parameter serves to capture the “accumulation effect” defined in section 4.2.1. It is not an empirically established measure, so we assume a reasonable range by common sense in the simulation.</td>
</tr>
<tr>
<td>( \theta_s )</td>
<td>([1.1, 3.0])</td>
<td>There is no empirical literature that studies the increase in death rate due to lack of access to medical treatment for COVID patients. We thus take a reasonable range of values larger than 1, assuming positive impact in health outcomes from access to medical treatment.</td>
</tr>
<tr>
<td>( b )</td>
<td>([0.41, 0.60])</td>
<td>During the pandemic, the proportion of deaths by underlying health conditions such as diabetes and ischemic heart disease that occurred in a hospital decreased by around 50% compared to the pre-pandemic period. We conjecture that general access to medical resources decreased by a similar rate, and test for a reasonable range around 0.5.</td>
</tr>
<tr>
<td>( d_C )</td>
<td>0.05%</td>
<td>The proportion of population that died from COVID yields ( \frac{202,141}{126,000,000} = 0.16% ) according to data in 2020. Among the deaths, study shows approximately 30% do not have underlying health conditions, which scales the rate to ( 0.16% \times 0.3 \approx 0.05% ).</td>
</tr>
<tr>
<td>( d_S )</td>
<td>0.61%</td>
<td>We subtract the number of COVID deaths from total deaths due to diseases in 2020 and scale the value for March to December using deaths due to diseases in a pre-pandemic year (2019). The number yields 764,800, which is approximately 0.61% of the total population.</td>
</tr>
<tr>
<td>( d_B )</td>
<td>0.11%</td>
<td>The proportion of population that died from COVID yields ( \frac{202,141}{126,000,000} = 0.16% ) according to data in 2020. Among the deaths, study shows approximately 70% have underlying health conditions, which scales the rate to ( 0.16% \times 0.7 \approx 0.11% ).</td>
</tr>
<tr>
<td>( d )</td>
<td>0.77%</td>
<td>The sum of the approximated values of ( d_C ), ( d_S ), and ( d_B ).</td>
</tr>
</tbody>
</table>
Appendix B – Optimal Flexible Reserve Strategies and Predicted Mortality Rates subject to Varying Effective Reproduction Rates ($r_e$)

<table>
<thead>
<tr>
<th>$r_e$</th>
<th>Optimal Switching Point $t^*$</th>
<th>$a_1$</th>
<th>$b_1$</th>
<th>$c_1$</th>
<th>$a_2$</th>
<th>$b_2$</th>
<th>$c_2$</th>
<th>$d_C$</th>
<th>$d_S$</th>
<th>$d_B$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.94</td>
<td>0.06</td>
<td>0.01</td>
<td>0.94</td>
<td>0.05</td>
<td></td>
<td>0.0008%</td>
<td>0.4934%</td>
<td>0.00019%</td>
<td>0.49%</td>
</tr>
<tr>
<td>1.1</td>
<td>1</td>
<td>0.94</td>
<td>0.06</td>
<td>0.01</td>
<td>0.94</td>
<td>0.05</td>
<td></td>
<td>0.00024%</td>
<td>0.4933%</td>
<td>0.00057%</td>
<td>0.49%</td>
</tr>
<tr>
<td>1.2</td>
<td>1</td>
<td>0.94</td>
<td>0.06</td>
<td>0.01</td>
<td>0.94</td>
<td>0.05</td>
<td></td>
<td>0.00084%</td>
<td>0.4933%</td>
<td>0.00020%</td>
<td>0.50%</td>
</tr>
<tr>
<td>1.3</td>
<td>18</td>
<td>0.01</td>
<td>0.94</td>
<td>0.05</td>
<td>0.02</td>
<td>0.93</td>
<td>0.05</td>
<td>0.0031%</td>
<td>0.4932%</td>
<td>0.00073%</td>
<td>0.50%</td>
</tr>
<tr>
<td>1.4</td>
<td>19</td>
<td>0.94</td>
<td>0.06</td>
<td>0</td>
<td>0.79</td>
<td>0.21</td>
<td></td>
<td>0.0131%</td>
<td>0.4969%</td>
<td>0.00269%</td>
<td>0.54%</td>
</tr>
<tr>
<td>1.5</td>
<td>18</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0.19</td>
<td>0.81</td>
<td></td>
<td>0.0438%</td>
<td>0.5224%</td>
<td>0.00953%</td>
<td>0.66%</td>
</tr>
<tr>
<td>1.6</td>
<td>17</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>0.1355%</td>
<td>0.5366%</td>
<td>0.4428%</td>
<td>1.11%</td>
</tr>
<tr>
<td>1.7</td>
<td>16</td>
<td>0.92</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>0.3736%</td>
<td>0.5401%</td>
<td>1.5966%</td>
<td>2.51%</td>
</tr>
<tr>
<td>1.8</td>
<td>15</td>
<td>0.91</td>
<td>0.09</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>0.8572%</td>
<td>0.5337%</td>
<td>4.0890%</td>
<td>5.48%</td>
</tr>
<tr>
<td>1.9</td>
<td>14</td>
<td>0.92</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>1.4823%</td>
<td>0.5113%</td>
<td>7.3754%</td>
<td>9.37%</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>1.9232%</td>
<td>0.4328%</td>
<td>9.3929%</td>
<td>11.75%</td>
</tr>
<tr>
<td>2.1</td>
<td>18</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.0577%</td>
<td>0.4085%</td>
<td>10.0243%</td>
<td>12.49%</td>
</tr>
<tr>
<td>2.2</td>
<td>17</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.1165%</td>
<td>0.3870%</td>
<td>10.3145%</td>
<td>12.82%</td>
</tr>
<tr>
<td>2.3</td>
<td>16</td>
<td>0.92</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.1568%</td>
<td>0.3695%</td>
<td>10.5474%</td>
<td>13.07%</td>
</tr>
<tr>
<td>2.4</td>
<td>16</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.1962%</td>
<td>0.3475%</td>
<td>10.7550%</td>
<td>13.30%</td>
</tr>
<tr>
<td>2.5</td>
<td>15</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.2203%</td>
<td>0.3330%</td>
<td>10.8981%</td>
<td>13.45%</td>
</tr>
<tr>
<td>2.6</td>
<td>14</td>
<td>0.92</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.2365%</td>
<td>0.3223%</td>
<td>11.0332%</td>
<td>13.59%</td>
</tr>
<tr>
<td>2.7</td>
<td>14</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.2427%</td>
<td>0.3077%</td>
<td>11.0662%</td>
<td>13.62%</td>
</tr>
<tr>
<td>2.8</td>
<td>14</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.2447%</td>
<td>0.2961%</td>
<td>11.0635%</td>
<td>13.60%</td>
</tr>
<tr>
<td>2.9</td>
<td>13</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.2422%</td>
<td>0.2901%</td>
<td>11.0767%</td>
<td>13.61%</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>0.93</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>2.2443%</td>
<td>0.2803%</td>
<td>11.0708%</td>
<td>13.60%</td>
</tr>
</tbody>
</table>

*We denote the optimal resource partition before the optimal switching point as $a_1, b_1, c_1$ and optimal partition after the optimal switching point as $a_2, b_2, c_2$.

**The highlighted row denotes the optimization under flexible reserves using baseline parameters.
Bibliography


[15] “Patients with underlying conditions were 12 times as likely to die of covid-19 as otherwise healthy people, CDC finds” The Washington Post, https://www.washingtonpost.com/health/2020/06/15/patients-with-underlying-conditions-were-12-times-more-likely-die-covid-19-than-otherwise-healthy-people-cdc-finds/.

