Effective Management of Virtual and Office Appointments in Chronic Care

by

Xiao Yu

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering (Industrial and Systems Engineering) in the University of Michigan-Dearborn 2020

Master's Thesis Committee:

Assistant Professor Armagan Bayram, Chair
Professor Onur Ulgen
Assistant Professor Xi Chen

Acknowledgements

I want to take a moment and thank those who have helped me through the completion of my thesis and master's program.

I wish to express my sincere gratitude to Dr. Armagan Bayram for her inspiration, patient instruction, and timely guidance for this research. It was her supports and advice that helped me finish this work.

I am also very thankful to the professors and lecturers who taught me classes during the two-year graduate study for the knowledge I learned and for the questions they helped me to solve. I am also very appreciative to my family for their steady support, encouragement, and love.

Last but not least, I want to extend my appreciation to those who could not be mentioned here but give me a lot of inspiration.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Patients suffering from a chronic disease often require regular appointments and treatments. Due to the constraints on the availability of office appointments and the capacity of physicians, access to chronic care can be limited; consequently, patients may fail to receive the recommended care suggested by clinical guidelines. Virtual appointments can provide a cost-effective alternative to traditional office appointments for managing chronic conditions. Advances in information technology infrastructure, communication, and connected medical devices are enabling providers to evaluate, diagnose, and treat patients remotely. In this study, we first build a capacity allocation model to study the use of virtual appointments in a chronic care setting. We consider a cohort of patients receiving chronic care and model the flow of the patients between office and virtual appointments using an open migration network. We formulate the planning of capacity needed for office and virtual appointments with a newsvendor model to maximize long-run average earnings. Moreover, we develop two optimization models to determine the optimal follow-up rate for patients and a two-stage stochastic programming model to investigate the capacity allocation decisions along with the patients' scheduling decisions under uncertainty. We consider differences in treatment and diagnosis effectiveness for office and virtual appointments. We derive optimal policies and perform numerical experiments. With the model developed, capacity allocation, follow-up rate determination and patient scheduling decisions for office and virtual appointments can be made more systematically with the consideration of patients' disease progressions.

Chapter 1: Introduction

Chronic care involves the treatment and monitoring of pre-existing and long-term diseases such as diabetes, high blood pressure, asthma, Alzheimer's disease, and cardiovascular disease (Bodenheimer 2002). In the U.S., 45% of the population has at least one chronic disease, and the cost of chronic care contributes to over 75% of the entire health care spending in the U.S. (Wu 2000, Heffler 2002). Given that the population is increasing and aging, the need for chronic care in the future will increase faster. Current care processes are insufficient to address the coming mismatch in supply and demand for chronic care (Gupta 2008). To improve patient access to chronic care and to reduce their burden, health care providers increasingly rely on virtual appointments as a new alternative way to provide effective and consistent long-term care. Virtual appointments, consisting of e-mail, phone, and online consultations, can improve patient access and ensure continuity of care and, consequently, better health outcomes (Perednia 1995, Caceres 2006).

Virtual appointments can be used as a substitute for, or complementary to office appointments, and they can take many different forms. For example, virtual appointments can be used for diagnosis only, for treatment only, and for both treatment and diagnosis similar to the office appointments (Bayram 2019). More specifically, through virtual appointments that provide diagnosis only, chronic care patients can be monitored in real time remotely and updates regarding the patients' status can be obtained (Marcin 2013, Association 2020). Through virtual appointments that provide treatment only, educational support and reliable resources can be provided to patients without diagnosing patients' status (Goodarzi 2012). Finally, virtual appointments can be also used to provide both diagnosis and treatment in which both patients' health statuses are diagnosed, and proper treatment is provided (Bayram 2019). Since virtual appointments are provided remotely, they can enhance the delivery of health care to geographically disadvantaged and medically underserved populations (Ackerman 2010). In addition, patients who are unable or unwilling to leave their homes to seek medical treatments or are in the poor physical conditions can also benefit from the virtual appointments (Bashshur 1995, Bedi 2009). Virtualappointments have the potential to enhance primary care delivery by enabling both health-delivery and travel-cost reductions and larger panel sizes without sacrifices in the quality of health care (Russo 2016). Parallel to its benefits, more patients are willing to receive care through this convenient way. Thus, the demand for virtual appointments is increasing quickly. The total number of virtual consultations is growing by around 10% a year, with growth projected to reach around 25 million in 2020 (Wu 2018).

The focus of this thesis is on the integration of virtual appointments with traditional office appointments. We combine the advantages of virtual appointments and office appointments by considering that virtual appointments are more cost-effective than office appointments while office appointments can have better treatment effectiveness than virtual appointments. In Chapter 2, we review related literature in the area of capacity planning in chronic care. In Chapter 3, we develop a migration network to simulate the clinic system with both virtual and office appointments. Then, we develop a newsvendor-type optimization model to determine the capacity of office and virtual appointments that maximize the average long-term profit of the clinic. With the model developed, we perform numerical experiments to present the application of the mathematical model. In Chapter 4, we use the migration network model that we build in Chapter 3 and investigate the optimal follow-up rate for a given capacity of office and virtual appointments. We develop a linear and nonlinear programming models to determine the optimal follow-up rate to help healthcare providers in their decision-making process. In Chapter 5, we consider capacity allocation decisions and patient scheduling decisions for office and virtual appointments simultaneously under uncertainty. We develop a two-stage stochastic programming model to investigate capacity allocation and patient scheduling decisions that maximize patients' overall health condition. We consider that patients' health states are uncertain, and this information is realized over time. Our results provide managerial insights for clinics in allocating capacity and patient scheduling for varying parameter values.

Chapter 2: Literature Review

Our study builds on the literature of decision models in community-based chronic care delivery. Related to this area, Kucukyazici, et al. (2013) present and analyze three representative examples of prevailing quantitative decision models for managing community-based chronic care. For each example, they analyze the background of the problem, present the methodology, and show their findings and implications. Among these examples, Batun, et al. (2013) study the optimization-based problems in healthcare delivery, such as workforce scheduling, operating room scheduling, appointment scheduling, capacity planning, and some other practical problems. They refer to recent studies and present detailed examples with the use of optimization methods, especially stochastic programming, discrete convex analysis, and approximate dynamic programming. Kucukyazici, et al. (2011) propose a Markov decision process to model multiple care-provider visit patterns for stroke patients, while Deo, et al. (2013) combine a Markovian disease progression model with a capacity allocation model to determine revisit intervals for childhood asthma care. A major difference of our study from the listed literature is that we consider different types of appointments (i.e., virtual and office appointments) and investigate the optimal capacity allocation decisions among the different types of appointments.

Related to the virtual appointment setting, studies that investigate the management of virtual and office appointments are limited. In a relevant study, Liu et al. (2018) build an optimization model to design effective checkup plans (i.e., phone calls, office visits) for individual patients to monitor after hospital discharge. Their study considers only the diagnosis impact of the virtual appointments, whereas we include both the treatment and the diagnosis impact of the virtual appointments. Among the studies considering both treatment and diagnosis impact of virtual appointments, Bavafa, et al. (2019) develop a Markovian model to determine the patients' revisit intervals in primary care by incorporating virtual appointments into an office appointment setting. In another study, Bayram, et al. (2019) develop a stochastic dynamic programming model to determine the follow-up rates for virtual and office appointments, and they investigate the value of virtual appointments in patients' health outcomes. In these related papers, the capacity of the appointments is assumed to be given. Different from these studies, we investigate the optimal capacity allocation of office and virtual appointments for different settings.

Another stream of literature that is relevant to our study is on the capacity planning problem in health care, which addresses the issue of allocating limited resources to satisfy the demand of the patients. There are several studies in this area, and Hulshof, et al. (2012) provide a comprehensive review of resource allocation and capacity planning in health care. Among this literature, the following papers are more relevant to our methodology. Bretthauer, et al. (1998) develop an optimization/queuing network model for optimal planning of resource allocations (e.g., beds and nurses) and apply it to a blood bank and a health maintenance organization. Lee and Zenios (2009) develop a multi-class migration network model as an optimization model to determine the optimal capacity that maximizes the overall profit of a dialysis clinic. Li, et al. (2016) present a long-term care network model to determine the optimal capacity for nursing homes and community-based services. Different from the above literature, we consider both patient flow and patients' disease progression to determine optimal capacity allocations. Moreover, our study focuses on two different appointments with varying effectiveness in both diagnosis and treatment.

Chapter 3: Capacity Planning Using the Newsvendor Model

In this chapter, we build mathematical models to understand the patient flows and to provide managerial insights for capacity allocation decisions among office and virtual appointments. This chapter is organized as follows. Section 3.1 provides introduction and motivation related to the capacity planning of virtual and office appointments in chronic care. Section 3.2 presents the migration network model to understand patients' flow. Section 3.3 presents newsvendor models and algorithms to allocate the capacity among office and virtual appointments that maximizes clinics' average earnings. Section 3.4 presents numerical experiments, estimates parameters, and shows sensitivity analysis results to illustrate the application of the model. Section 3.5 outlines the conclusion of the chapter and provides some future research directions.

3.1 Introduction

Despite the increased usage of virtual appointments and their observed benefits in chronic care, the integration of virtual appointments with office appointments can be operationally challenging for the clinics. One of the reasons of this challenge is that virtual and office appointments can have differences in their treatment/diagnosis effectiveness and in their costs. More specifically, although virtual appointments can provide cost-effective treatments, they can result in similar (Craig 2000) or worse patient-related outcomes (Leggett 2001, McKinstry 2010) compared to the office appointments which makes it harder to decide how to allocate the available capacity among different appointments. Moreover, with the integration of virtual appointments, the patients' flow dynamics become complex, and it gets difficult to identify the expected number of patients that can be scheduled for office and virtual appointments. Indeed, faced with rising costs and patient populations, managers of health facilities, like clinics, strive to determine an appropriate capacity to meet the needs of the patients and avoid the opportunity cost and over-utilization cost as much as possible. Thus, it is important to develop strategies to determine the expected number of patients and allocate available capacity efficiently by considering the patients' flow dynamics. To address the need for capacity allocation policies, in this chapter, we study a chronic care setting in which patients are scheduled for virtual or office appointments. We consider that similar to office appointments, virtual appointments can also provide both treatment and diagnosis, and parallel to previous studies (Leggett 2001, McKinstry 2010, Bayram 2019), we assume that virtual appointments can be less effective than office appointments. We develop a modeling framework to determine the optimal allocation of the capacity for both office and virtual appointments and aim to answer the following operational questions:

1. What is the expected number of patients scheduled for office and virtual appointments for the given follow-up, service, arrival, and departure rates?
2. How should the available capacity be allocated among office and virtual appointments?

To address these questions, we develop a migration network model to analyze patients' flow and disease progressions. Using the migration network model, we first analytically investigate the number of patients in the steady state who are scheduled for virtual and office appointments. Second, we develop a newsvendor-type model to maximize the long-run average earnings of a health clinic. We further propose an algorithm to find the optimal capacity allocations among virtual and office appointments. Third, we analytically investigate how limited capacity impacts the proposed algorithm and the optimal capacity allocation decisions. Finally, through our numerical studies, we analyze the effect of model parameters on the allocation of the capacity of the office and virtual appointments by analyzing different scenarios.

3.2 Migration Network Model for Office and Virtual Appointments

In this section, we consider a cohort of patients receiving chronic care via both office and virtual appointments. In this network, two types of patients are served (i.e., new patients and returning patients), and physicians provide both office and virtual appointments. We use a continuous-time open migration network (Kelly 1979, p.48-p.57) to simulate the population dynamics (i.e., patient flows and disease progression) in which patients' arrivals are considered as Poisson process and the time intervals between patient transitions are independently and exponentially distributed.

We illustrate our migration network model in Figure 3.1, and we describe nodes and flows of the network in this section. We use $i \in \{o, v\}$, where "o" corresponds to office appointments and "v" corresponds to virtual appointments, to denote the type of appointments. New patients with office and virtual appointments arrive with Poisson arrival rate λ_i, $i \in \{o, v\}$. We define the "service" as the diagnosis and the treatment of a patient, and we consider that virtual and office

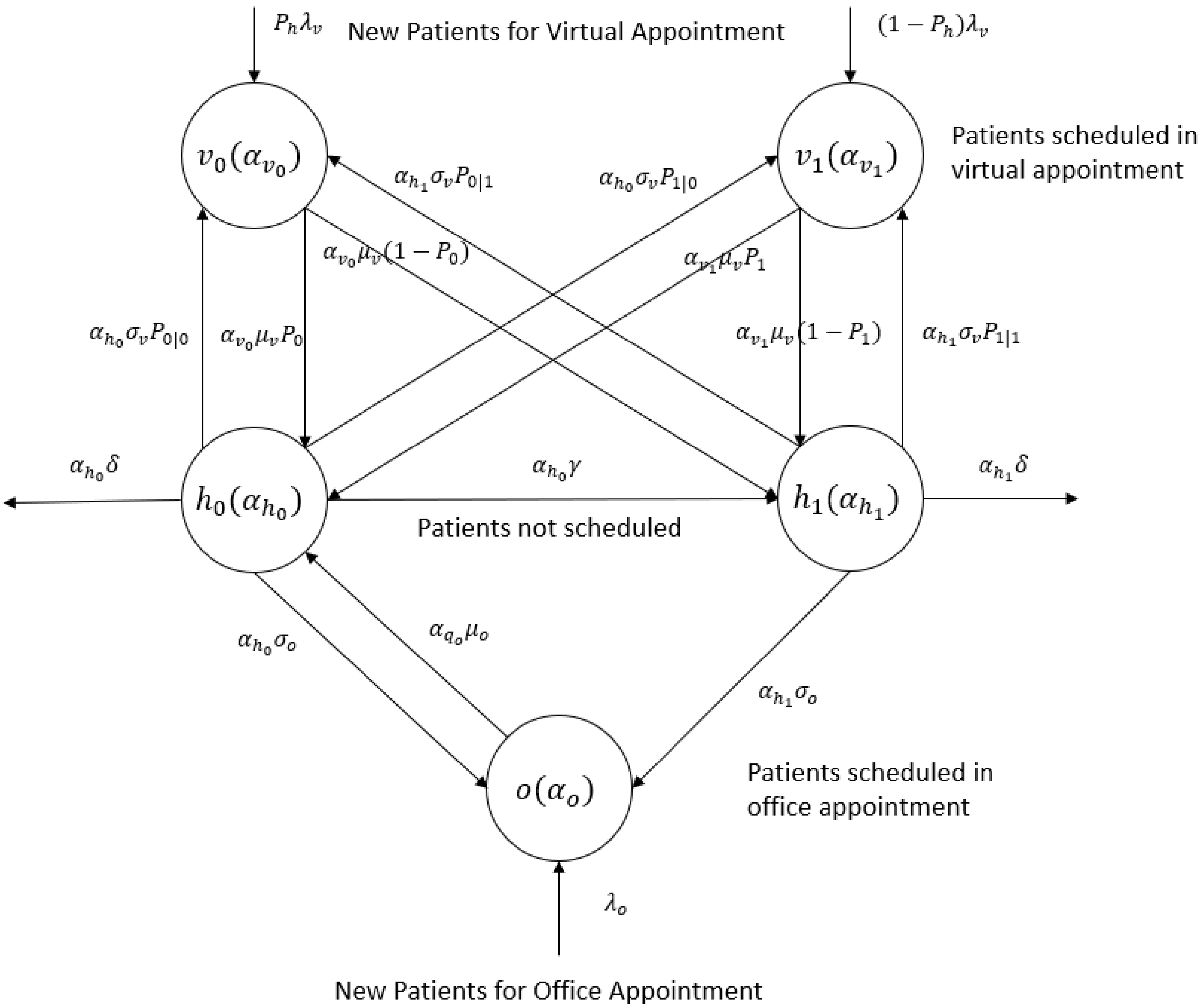

Figure 3.1: Migration network with imperfect diagnosis and treatment

appointments provide both diagnosis and treatment during the appointment. More Specifically, service time corresponds to the duration of an appointment, and service times of patients are exponentially and independently distributed. We use $\mu_i,\ i \in \{o, v\}$ to denote the service rate of office and virtual service, respectively. We define follow-up time (i.e., revisit interval) as the time between the current visit and the next time the patient initiates an appointment. We consider that after each appointment, the physician recommends to the patient the type and the time of the next visit. Hence, based on the physician's recommendation, patients are scheduled for appointments. Patients' follow-up time are assumed to be independently and exponentially distributed with an average $1/\sigma_i,\ i \in \{o, v\}$. Hence, σ_i represents the rate of patients' revisits. Patients may depart from the physicians' panel before making another appointment (i.e., change the physician). Patients' departure times are independently and exponentially distributed with a mean of $1/\delta$.

We use the "control" measure to characterize the patient health status. "Control" measure helps to understand how well chronic-care related symptoms are currently controlled in a patient. Depending on the types of chronic diseases these categorizations may differ. For example, for asthma, four categories can be used as follows: (i) controlled, (ii) improved, (iii) unchanged, and (iv) worsened, and the last three are classified as an uncontrolled state (Deo 2013). For the sake of simplification, in our model, we consider two health states as controlled and uncontrolled to characterize the patients' health status.

We assume that patients in the network may not be scheduled for an appointment (we note that patients who are not scheduled for an appointment are the ones who are not receiving service or who are not in the queue for service) and are waiting for their next appointment time (i.e., system state, h), or are scheduled for a virtual appointment and receiving care (i.e., system state, v), or are scheduled for an office appointment and receiving care (i.e., system state, o). Let $j \in \{0,1\}$ represent the set of health states, where "0" corresponds to the controlled health state and "1" corresponds to the uncontrolled health state. To model the disease deterioration, we denote h_0 as patients who are in a controlled health state and not scheduled for an appointment, while we use h_1 to denote patients who are in an uncontrolled health state and not scheduled for an appointment. We assume that there is no transition from the uncontrolled state to the controlled state without treatment. However, due to disease progression, some of the patients in the controlled health state and not receiving care (i.e., h_0) may transition into the uncontrolled health state (i.e., h_1) within the unit time. The time for a controlled patient to progress into the uncontrolled state is assumed to follow an exponential distribution with an average $1/\gamma$.

At each type of appointment, the health state of the patient is diagnosed, and the patient is treated. We assume that office appointments can be more effective than virtual appointments (Leggett 2001, McKinstry 2010, Bayram 2019). We assume that the treatment and the diagnosis in the office appointments are perfect, while those of the virtual appointments are imperfect. Perfect treatment means that a patient's health state recovers to the best health state after treatment, while the perfect diagnosis means that a patient's health state is revealed accurately during the diagnosis. On the other hand, imperfect treatment means that a patient's health state can transit into a different health state with some probability, while the imperfect diagnosis means that a patient's health state may be revealed inaccurately during the diagnosis. Perfect diagnosis/treatment assumption is similar to the ones in the machine maintenance and repair literature as well (Block 1993, Pham 2000). Moreover, in the healthcare literature, the perfect diagnosis and treatment assumption is also used by (Deo 2013, Bayram 2019). More specifically, patients in each health state are assumed to be always diagnosed accurately if they are scheduled for an office appointment, and they will be in a controlled health state after the office appointment regardless of their initial health state before the appointment. On the other hand, patients scheduled for a virtual appointment may be diagnosed inaccurately, since virtual appointments are expected to be less precise than office appointments (Leggett 2001, McKinstry 2010, Bayram 2019). Also, since the virtual appointments are not as effective as office appointments, a patient diagnosed in an uncontrolled health state at the virtual appointment, may remain in the uncontrolled health state with probability $(1 - P_1)$ or may transition into the controlled health state with probability P_1. Similarly, a patient diagnosed in a controlled health state at the virtual appointment may remain in the controlled condition with probability P_0 after the virtual appointment or may be in the uncontrolled health state with probability $(1 - P_0)$ after the virtual appointment (since not all patients in a controlled health state may be diagnosed accurately). Thus, to capture these effects, we use v_0 to denote patients who are scheduled for a virtual appointment and diagnosed in a controlled health state after the virtual appointment and v_1 to denote patients who are scheduled for a virtual appointment and diagnosed in an uncontrolled health state after the virtual appointment. We use the conditional probability to define the perfect diagnosis probability for the virtual appointments. We denote $P_{j|j'},\ j, j' \in \{0,1\}$ as the probability that the patient in health state j' is diagnosed in health state j at virtual appointment. We have $P_{0|j} + P_{1|j} = 1,\ j \in \{0,1\}$. We further assume that the new patients scheduled for virtual appointments will be diagnosed in controlled health state with probability P_h.

Overall, we consider five nodes in the network, and we use $k \in \{h_0, h_1, o, v_0, v_1\}$ to represent the set of nodes in the migration network. In Figure 3.1, we illustrate the described flow of patients between each node through arcs. The arcs in between nodes represent the process of a patient that flows from one node to another. For example, the arc from node "o" to "h_0" represents the flow of patients from office appointments to their home after they have their appointment. We also show the inflow and outflow for each node next to each arc. For example, there are two out flows from node "v_1" where uncontrolled patients can improve to the controlled health state or can remain in uncontrolled health state after receiving the virtual appointment.

We define α_k to denote the expected number of patients at node $k \in \{h_0, h_1, o, v_0, v_1\}$ in the steady-state condition. The number of patients at node k satisfy the following balance equations, which are derived from Figure 3.1 (Kelly 1979, p.49):

$$\mu_v \alpha_{v_0} - \sigma_v P_{0|0} \alpha_{h_0} - \sigma_v P_{0|1} \alpha_{h_1} = P_h \lambda_v \tag{3.1}$$

$$\mu_v \alpha_{v_1} - \sigma_v \left(1 - P_{0|0}\right) \alpha_{h_0} - \sigma_v \left(1 - P_{0|1}\right) \alpha_{h_1} = (1 - P_h) \lambda_v \tag{3.2}$$

$$-\mu_v P_0 \alpha_{v_0} - \mu_v P_1 \alpha_{v_1} + (\sigma_v + \sigma_o + \delta + \gamma) \alpha_{h_0} - \mu_o \alpha_o = 0 \tag{3.3}$$

$$-\mu_v (1 - P_0) \alpha_{v_0} - \mu_v (1 - P_1) \alpha_{v_1} - \gamma \alpha_{h_0} + (\sigma_v + \sigma_o + \delta) \alpha_{h_1} = 0 \tag{3.4}$$

$$-\sigma_o \alpha_{h_0} - \sigma_o \alpha_{h_1} + \mu_o \alpha_o = \lambda_o \tag{3.5}$$

These equations represent that the inflow to node i must to be equal to outflow from node i. Equations (3.1-3.5) are five equations with five unknowns, then we can solve the traffic equations and obtain the average number of patients in each node at steady-state. The result for each α_i is in the Appendix. We use α_k, $\forall k \in \{h_0, h_1, o, v_0, v_1\}$ to define the steady-state distribution π_k, $\forall k \in \{h_0, h_1, o, v_0, v_1\}$. Hence, let x_k denote the number of patients at node k. It shows that in steady state, the nodes states are independent and the steady-state distribution for each node k is a Poisson distribution (Kelly 1979, p.53) and given by

$$\pi_k(x_k = x) = \frac{\alpha_k^x}{x!} \Big/ \sum_{n=0}^{\infty} \frac{\alpha_k^n}{n!} = e^{-\alpha_k} \frac{\alpha_k^x}{x!}, \quad k \in \{h_0, h_1, o, v_0, v_1\} \tag{3.6}$$

The steady-state distribution defines the probability of having x_k number of patients at each node k. We use these probabilities to define the probabilistic capacity allocation model in Section 3.3.

3.3 Capacity Allocation Optimization Model

In this section, we build newsvendor-type capacity allocation models to find the optimal capacity for office and virtual appointments to maximize the long-run average earnings of a clinic. As described in the previous section, we consider that the node capacities of the migration network are unlimited where the number of patients at each node is unlimited. However, we consider a threshold capacity for office and virtual appointments (Li 2016). The capacity that we assign describes the number of patients that can be served under the regular cost, and the actual number of patients in office and virtual appointments can exceed this threshold capacity. When the number of patients in office and virtual appointments exceeds this threshold capacity, we consider that a penalty cost due to patient overflow occurs. Hence, we aim to find the optimal threshold capacity for office and virtual appointments for the clinics under the assumption that node capacities are unlimited. In the following sections, we first introduce the capacity allocation model without constraints. Then, we modify the unconstrained model by adding constraints on the optimal office and virtual appointment capacities.

3.3.1 Base Capacity Allocation Model

We consider a clinic that provides both virtual and office appointments with office appointment capacity of M_o, virtual appointment capacity of M_v, and the total capacity of $M = M_o + M_v$. Since in our migration network, we split the virtual appointments into two parts to reflect the imperfect diagnosis and treatment, we define M_{v_0} and M_{v_1}, which denote the virtual appointment capacity for controlled and uncontrolled patients, respectively (i.e., $M_v = M_{v_0} + M_{v_1}$). Defining different types of capacities for virtual appointments ensures more flexibility in the model definition and policy development, and it does not necessarily mean to split the virtual appointment capacity for clinics. Clinics still can consider the total capacity in their decision making, and they do not need to split the capacity for controlled and uncontrolled patients. Also, patients' health status cannot be known with certainty without diagnosing patients. However, these insights can be helpful for clinics when they are making patient scheduling decisions. In practice, although there is not a direct application of splitting capacity, health care providers may tend to schedule patients according to patient needs leading to scheduling the uncontrolled patients for appointments more often than the controlled patients based on the physicians' beliefs (Deo 2013). Hence, knowing the expected number of patients in each health state may help clinics in their patient scheduling decisions.

In our capacity allocation model, we use $r_k,\ k \in \{o,\ v_0,\ v_1\}$ to denote the marginal profit for each patient treatment through office and virtual appointments, respectively; the marginal profit is the difference between revenue and variable cost (variable costs are the type of costs that can change depending on the number of patients served, such as hourly labor cost or the cost of materials or supplies). We note that there is one type of virtual appointment and $r_v = r_{v_0} = r_{v_1}$. Similarly, each unit of capacity for office and virtual appointments is associated with a fixed cost $c_k,\ k \in \{o, v_0,\ v_1\}$ per unit of time, where unit capacity cost is the fixed cost of allocating capacity which can be employee salaries, building-related costs and equipment. Thus, the cost of capacity,

$c_kM_k,\ k \in \{o, v_0, v_1\}$, is independent of the patient flow. We assume that $r_k > c_k$ (Lee 2009) and $c_v = c_{v_0} = c_{v_1}$. By assuming $r_k > c_k$, we ensure that the optimal capacity M_k is greater than 0. More specifically, if the unit capacity cost is larger than or equal to the marginal profit, it will be optimal to provide no service and $M_k = 0$. We assume that the number of patients at the office and virtual appointments can exceed the allocated capacity, and in this case, the clinic provides the corresponding appointment but at a higher total cost. To reflect the cost of patient overflow, we define $f_k,\ k \in \{o,\ v_0,\ v_1\}$ to represent the unit net penalty cost of the overflow, where $f_v = f_{v_0} = f_{v_1}$. The definition is similar to the definition of the overbooking cost used by (Lee 2009). It is the net cost of meeting the overflow demand, which is the difference between the total variable cost of meeting the extra demand and the revenue earned for that appointment. The clinic still earns the marginal profit r_k for the overflow patients, but the extra variable cost of meeting this excess demand is more than the marginal profit. Let $x_k(t)$ denote the current number of patients at the node $k, \forall k \in \{o,\ v_0,\ v_1\}$ at time t. Then, our base capacity allocation model can be defined as follows:

$$\max \quad A(M) \tag{3.7}$$

$$s.t. \quad M_k \geq 0 \quad \forall k \in \{o,\ v_0,\ v_1\} \tag{3.8}$$

Where

$$A(M) = \lim_{T\to\infty} \frac{1}{T} \left\{ \sum_{k\in\{o,v_0,v_1\}} \int_0^T r_k min\,[x_k(t), M_k]dt - \int_0^T c_k M_k dt - \int_0^T f_k (x_k(t) - M_k)^+ dt \right\} \tag{3.9}$$

As noted before, the objective function is defined as the function of allocated capacity and the number of patients at each node. Since the number of patients at each node is uncertain, we use the steady-state probabilities π_k defined in Section 3.2. In objective function (3.9), the first term represents the marginal profit generated from office and virtual appointments, the second term represents the fixed capacity cost, and the third term represents the penalty costs associated with the capacity shortage. Equation (3.8) defines the non-negativity setting.

Let $\mathbb{E}_{\pi_k}(x_k), k \in \{o, v_o, v_1\}$ be the expected number of patients at node k under the steady state distribution $\pi_k, k \in \{o, v_o, v_1\}$. Due to the ergodicity of the open migration network (Kelly 1979, p.49), we can define the following equations:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T x_k(t)dt = \mathbb{E}_{\pi_k}(x_k) = \alpha_k \tag{3.10}$$

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T [M_k - x_k(t)]^+ dt = \mathbb{E}_{\pi_k}(M_k - x_k)^+ \tag{3.11}$$

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T [x_k(t) - M_k]^+ dt = \mathbb{E}_{\pi_k}(x_k - M_k)^+ \tag{3.12}$$

Then, we reformulate the objective function (3.9) with the following equation, and we include the detailed steps of the reformulation in the Appendix.

$$A(M) = \sum_{k\in\{o,v_0,v_1\}} \left[(r_k - c_k)\alpha_k - c_k E_{\pi_k}(M_k - x_k)^+ - (f_k + r_k - c_k)\mathbb{E}_{\pi_k}(x_k - M_k)^+\right] \tag{3.13}$$

Similar to equation (3.9), in equation (3.13), the first term is the difference between the marginal profit and the fixed cost. The second term is the opportunity cost for unutilized capacity, and the last term represents the cost due to patient overflow.

We define $A_k(M_k)$ as the individual objective function for appointment $k \in \{o, v_0, v_1\}$, and it can be defined as follows:

$$A_k(M_k) = (r_k - c_k)\alpha_k - c_k E_{\pi_k}(M_k - x_k)^+ - (f_k + r_k - c_k)\mathbb{E}_{\pi_k}(x_k - M_k)^+ \tag{3.14}$$

To maximize the objective $A(M)$, each sub-objective $A_k(M_k)$ can be maximized separately. We derive the optimal capacity for this unconstrained capacity planning model through Proposition 1.

Proposition 1. The optimal solution of the base capacity allocation model, denoted by $M^* = \left(M_o^{min}, M_{v_0}^{min}, M_{v_1}^{min}\right)$, is given by

$$M_k^{min} = min\left\{M_k \geq 0 : \pi_k(x_k \leq M_k) \geq \frac{f_k + r_k - c_k}{f_k + r_k}\right\} \tag{3.15}$$

where $\pi_k(x_k \leq M_k)$ is the cumulative probability that x_k is less than and equal to M_k and x_k follows Poisson distribution with parameter as α_k.

The proof of Proposition 1 is in the appendix. According to equation (3.15), the optimal capacity is influenced by both the parameter of Poisson distribution, α_k, and profit-related parameters, $\frac{f_k+r_k-c_k}{f_k+r_k}$. For a fixed Poisson parameter α_k, the cumulative probability function is a non-decreasing function of M_k. So, as $\frac{f_k+r_k-c_k}{f_k+r_k}$ increases, the new optimal capacity is no smaller than the old optimal capacity. For the value of $\frac{f_k+r_k-c_k}{f_k+r_k}$, increasing f_k, r_k, and decreasing c_k leads to the increase of $\frac{f_k+r_k-c_k}{f_k+r_k}$. Hence, we conclude that to some extent, increasing the marginal profit,

r_k, penalty cost on the overflow patients, f_k, and decreasing the fixed cost of each unit of capacity, c_k has a potential effect on increasing the optimal capacity for node k.

3.3.2 Capacity Allocation Model with Capacity Constraint

In practice, due to the limited resources of the clinic, the number of regular office and virtual appointments may be limited. In this section, we extend the base capacity allocation model presented in Section 3.3.1 and investigate the impact of adding a constraint on the optimal capacity allocation decisions. This change does not impact the balance equations of the migration network and patients' flows, and the objective function $A(M)$ remain the same as with the base model. More specifically, node capacities in the migration network are unlimited, and it is still allowed to have more than the M_k number of patients. Hence, equations (3.1) - (3.6) still hold when the constraint (3.17) is added. We use TC to denote the limited total capacity. Then, the capacity allocation model can be updated as follows:

$$\max A(M) \tag{3.16}$$

$$s.t. \sum_{k \in \{o, v_0, v_1\}} M_k \leq TC \tag{3.17}$$

$$M_k \geq 0 \forall k \in \{o, v_0, v_1\} \tag{3.18}$$

In the model, equation (3.17) states that allocated capacity should be less than or equal to the total available capacity TC, and equation (3.18) defines non-negativity constraints. Due to the capacity constraint, this problem becomes a resource allocation problem to optimally allocate the capacity to office and virtual appointments. Let M_{TC}^* be the capacity allocation decision when the total capacity is limited. Recall that $M^* = \left(M_o^{min}, M_{v_0}^{min}, M_{v_1}^{min}\right)$ is the optimal capacity for the base capacity allocation model given in Proposition 1. It is clear that if $\sum_{i \in \{o, v_0, v_1\}} M_i^{min} \leq TC$, then $M_{TC}^* = M^*$. This means that the clinic has enough capacity (resources), which maximizes their overall average earnings, and the clinic may consider not to have excess capacity. When the clinic doesn't have enough space, M_o, M_{v_0} and M_{v_1} are no longer independent. We provide an algorithm based on the partial differential as shown in Algorithm 1. Let M_k^t, $k \in \{o, v_0, v_1\}$ represent the capacity of node k at t^{th} iteration. In addition, we define the partial differential of the objective function for $k \in \{o, v_0, v_1\}$ as follows:

$$A'(M_k) = \frac{\Delta A(M)}{\Delta M_k} = A(M_k + 1) - A(M_k) = f_k + r_k - c_k - (f_k + r_k)\pi_k(x_k \leq M_k) \tag{3.19}$$

Then, Algorithm 1 can be stated as follows:

Algorithm 1 Capacity Allocation Algorithm based on the Partial Differential

Input: $TC > 0, TC \geq 0$

Output: M_{TC}^{*}

$t = 0$

$M_k^t \leftarrow 0$ for any $k \in \{o, v_0, v_1\}$

While $\sum_{k \in \{o, v_0, v_1\}} M_k^t < TC$

  $A'(M_k^t) = f_k + r_k - c_k - (f_k + r_k)\pi_k(x_k \leq M_k^t)$ for any $k \in \{o, v_0, v_1\}$

  if $A'(M_k^t) \leq 0$ for any $k \in \{o, v_0, v_1\}$ **then**

    break

  end if

  $x = argmax_k\, A'(M_k^t)$ for $k \in \{o, v_0, v_1\}$

  $M_x^{t+1} \leftarrow M_x^t + 1$

  $t \leftarrow t + 1$

  $M_{TC}^t \leftarrow \left(M_o^t, M_{v_0}^t, M_{v_1}^t\right)$

end while

$M_{TC}^{*} \leftarrow M_{TC}^t$

In Algorithm 1, we calculate the marginal gain of having one more unit of capacity of the office and virtual appointments. At each step, we compare the marginal gain of having one office and one virtual appointment and increase the capacity of the appointment with the highest gain by one. The algorithm stops when the allocated capacity reaches the available capacity or when adding one more capacity for all appointments yields a negative profit gain. Through Algorithm 1, $M_{TC}^{*} = \left(M_o^{*}, M_{v_0}^{*}, M_{v_1}^{*}\right)$ is the optimal solution to this problem. We adopt Algorithm 1 in our numerical studies for the capacity allocation model with capacity constraint.

3.3.3 Capacity Allocation Model with Time Constraint

In this section, we take into account the total time required for providing each type of appointment. Section 3.3.2 assumes that office and virtual appointments both take an equal amount of time. However, virtual appointments are expected to be shorter than office appointments. Hence, we update equation (3.17) by considering the total available time and service time of office and virtual appointments.

We use T_w to denote the average total available time for the clinic. As we define in Section 3.3.1, $\mu_i, i \in \{o, v\}$ represents the service rate of office and virtual service, respectively. And we assume virtual appointments for controlled and uncontrolled patients have the same service rate, μ_v (i.e., $\mu_{v_0} = \mu_{v_1} = \mu_v$). Hence, $\frac{1}{\mu_i},\ i \in \{o, v\}$ represents the average service time of one office and virtual service, respectively. Then the average long-run earnings maximization problem under time constraint becomes:

$$\max A(M) \tag{3.20}$$

$$s.t. \quad \frac{1}{\mu_o} M_o + \frac{1}{\mu_v}\left(M_{v_0} + M_{v_1}\right) \leq T_w \tag{3.21}$$

Where the objective function remains the same. Recall that $M^* = \left(M_o^{min}, M_{v_0}^{min}, M_{v_1}^{min}\right)$ is the optimal capacity for the base capacity allocation model given in Proposition 1. We use $M_{T_w}^*$ to denote the optimal capacity to the problem with average limited working time, T_w. It is clear that if $\frac{1}{\mu_o} M_o^{min} + \frac{1}{\mu_v}\left(M_{v_0}^{min} + M_{v_1}^{min}\right) \leq T_w$, then $M_{T_w}^* = M^*$. It is the case that the physicians have enough time to address the total number of both office and virtual appointments. When physicians don't have enough working time, M_o, $M_{v_0}$and M_{v_1} are no longer independent, they compete for the physicians' working time. Li et al. (2016) provide an approximation algorithm based on marginal analysis to solve this problem as shown Algorithm 2. We use M_{T_w} to denote the solution from Algorithm 2. we still define three intermediate variables, $M_k^t,\ k \in \{o, v_0, v_1\}$, to represent the capacity of node k at t^{th} iteration. We use $Z_k(M_k) = A'(M_k)/\frac{1}{\mu_k} = \mu_k A'(M_k)$ to denote the marginal profit of node k under unit of time.

Through Algorithm 2, We get the solution to this problem, M_{T_w}. Algorithm 2 is similar to Algorithm 1, but since the coefficients of the decision variables are not the same, Algorithm 2 cannot make sure the optimal solution to the problem under time constraint. Hence, we analyze the relative error of the solution from Algorithm 2 to the optimal solution.

Algorithm 2 Approximation algorithm based on marginal analysis

Input: $T_w > 0$

Output: M_{T_w}

$t = 0$

$M_k^t \leftarrow 0$ for any $k \in \{o, v_0, v_1\}$

While $\frac{1}{\mu_o} M_o^{min} + \frac{1}{\mu_v}\left(M_{v_0}^{min} + M_{v_1}^{min}\right) \le T_w$ **do**

  $Z_k(M_k^t) = \mu_k A'(M_k^t)$ for any $t \in \{o, v_0, v_1\}$

  for any $k \in \{o, v_1, v_0\}$ **do**

    if $\sum_{j \in \{o, v_0, v_1\}} \frac{1}{\mu_j} M_j^t + \frac{1}{\mu_k} > T_w$ **then**

      $Z_k = -1$

    endif

  end for

  if $Z_k(M_k^t) \le 0$ for any $k \in \{o, v_0, v_1\}$ **then**

    break

  endif

  $x = argmax_k\ Z_k$ for $k \in \{o, v_0, v_1\}$

  $M_x^{t+1} \leftarrow M_x^t + 1$

  $t \leftarrow t + 1$

  $M_{T_w}^t \leftarrow \left(M_o^t, M_{v_0}^t, M_{v_1}^t\right)$

end while

$M_{T_w} \leftarrow M_{T_w}^t$

Proposition 2. The relative error by using the solution from Algorithm 2, M_{T_w}, as an approximation of $M_{T_w}^*$ is no greater than $\frac{max\ A'\left(M_i^k\right)}{A\left(M_{T_w}\right)}$. We state the corresponding equations as follows:

$$\frac{A\left(M_{T_w}^*\right) - A\left(M_{T_w}\right)}{A\left(M_{T_w}^*\right)} \le \frac{A\left(M_{T_w}^*\right) - A\left(M_{T_w}\right)}{A\left(M_{T_w}\right)} < \frac{max\ A'(M_k^t)}{A\left(M_{T_w}\right)} \tag{3.22}$$

where M_k^t represents the capacity in the final iteration of the algorithms. Also, $A'(M_k^t)$ is the marginal gain of appointment type $k \in \{o, v_0, v_1\}$ in the final iteration. Proposition 2 ensures that the percent profit gap between the optimal solution and the proposed algorithm solutions is not greater than the percent marginal profit gain in the final iteration. The proof of Proposition 2 is in the Appendix, we use Algorithm 2 in our numerical studies for the capacity allocation with time constraint.

3.3.4 Capacity Allocation Model with Imperfect Diagnosis

In this section, we extend our base capacity allocation model present in Section 3.3.1 and study the effect of imperfect diagnosis (i.e., $P_{0|0}$ and $P_{0|1}$) on the optimal capacity and average long-run earnings for the migration network in Figure 3.1.

The model incorporates three decision variables: the capacity of office appointments, M_o, capacity of virtual appointments for controlled patients, M_{v_o}, and the capacity of virtual appointments for uncontrolled patients, M_{v_1}. Let $M = \left(M_o,\ M_{v_0},\ M_{v_1}\right)$. To consider the imperfect diagnosis in the model, we define e_{id} to denote the penalty cost on each imperfect diagnosis. Hence, for the patient that receives an inaccurate diagnosis, the physicians can still earn the revenue from the patient but face a penalty cost on his imperfect diagnosis. We still let $x_k(t)$ denote the current number of patients at the node k at time t. Our objective is to determine the optimal capacity M that the network long-run average earning is maximized, that is

$$\max A(M) \tag{3.23}$$

$$s.t. \quad M \in N_+ \tag{3.24}$$

where

$$A(M) = \sum_{k \in \{o, v_0, v_1\}} A_k(M_k) - \lim_{T \to \infty} \frac{1}{T} \left\{ \int_0^T e_{id} \left(\sigma_v (1 - P_{0|0}) x_{h_0}(t) + \sigma_v P_{0|1} x_{h_1}(t) \right) dt \right\} \tag{3.25}$$

The first term represents the total average long-run earnings of nodes o, v_0 and v_1 before considering penalty cost on imperfect diagnosis, the second term represents the average long-run penalty cost on the imperfect diagnosis.

We reformulate the objective function (3.24),

$$A(M) = \sum_{k \in \{o, v_0, v_1\}} A_i(M_i) - W_{id} \tag{3.26}$$

We denote the second term, which is the average long-run penalty cost on imperfect diagnosis, as W_{id}. The decision variables of the capacity allocation model with imperfect diagnosis are the capacity of office appointments, virtual appointments for controlled patients, and virtual appointments for uncontrolled patients (i.e., M_o, M_{v_0}, and M_{v_1}). From function (3.25), we can see that the value of W_{id} is determined by e_{id}, σ_v, $P_{0|0}$, $P_{0|1}$, α_{h_0}, and α_{h_1}. The values of these parameters are defined in the migration network, which is not related to the decision variables in the capacity allocation model. Hence, in this capacity allocation model, W_{id} is constant. The optimal solution to the capacity allocation model with imperfect diagnosis can still obtained through Proposition 1. The optimal solution is denoted by $M^* = \left(M_o^{min}, M_{v_0}^{min}, M_{v_1}^{min}\right)$. Since $P_{0|0}$ and $P_{0|1}$ affect the optimal capacity through affecting the steady condition of the migration network and also affect the average long-run earnings through capacity allocation model, we do a sensitivity analysis in the numerical experiments to find the effect of imperfect diagnosis.

3.4 Numerical Studies

In this section, we perform numerical experiments to analyze how the optimal capacity allocation decision varies under different scenarios. To this end, we first describe the model parameter estimation process. Then, we investigate the change in the optimal capacity allocation with respect to the follow-up rate and capacity constraint. Finally, we compare some common policies in practice with the proposed solutions.

3.4.1 Parameter Estimation

In this section, we describe how the model parameters are obtained. We note that our data are based on the literature, and we use several sources to find the parameter values. The parameter values we obtain represent different characteristics. Due to the variation of the parameters' characteristics, fluctuation in the results can be expected. The parameters that we present in this section are our initial setting and results that are obtained based on a single setting may not be generalized. To overcome these issues, from Section 3.4.2 to Section 3.4.6, we define several scenarios and investigate the change in the proposed office and virtual appointment capacities by considering possible fluctuations in the parameter values. We present minimum, average and maximum values of the proposed capacity values to provide a range for the decision-makers.

Flow Parameters. Based on a survey of American physicians (Foundation 2018), on average, doctors see 20.2 patients per day, and physicians work on-average 51.40 hours per week (including all clinical and non-clinical duties). Of these, physicians work average 11.37 hours per week on non-clinical (paperwork) duties only. Hence, we obtain the average service time for each patient as $(51.4 - 11.37)/(5 \times 20.2) = 0.396$ hours. Thus, the office service rate is estimated as $\mu_o = 1/0.396 = 2.525/\text{hour}$. In addition, the average appointment time of the virtual appointments is less than that of the office appointments, and it is reported as around 12 minutes (Valero 2000). Thus, the virtual service rate is estimated as $\mu_v = 60/12 = 5/\text{hour}$. To calculate the new patient arrival rate, we consider the state of Michigan. The population of Michigan is 9.976 million in 2017, and 47.9% of them suffered chronic diseases (Disease 2017). While in 2018 the Michigan population increases to 9.996 million and 48.1% of them suffered chronic disease (Bureau 2018). Then the increasing number of chronic patients in Michigan can be calculated as $(9.996 \times 48.1\% - 9.976 \times 47.9\%) \times 10^6 = 29{,}572$. There are 278 clinics in the state of Michigan (Clinics.com 2019). Then, the total monthly new arrival rate is estimated as $\lambda_o + \lambda_v = 29572/(278 \times 12) = 8.865/\text{month}$. Besides, about 10.4% percent of the patients occur through virtual appointments (Foundation 2018). Thus, $\lambda_o = 8.865 \times (1 - 10.4\%) = 7.943/$ month, $\lambda_v = 8.865 - 7.943 = 0.922/$ month. According to CPT code 99490 in Chronic Care Management (CCM) (Services 2016), a patient should receive at least 20 minutes of clinical staff time directed by a physician or other qualified health care professional per calendar month. Considering the service rates of office and virtual appointments, the follow-up rate can be estimated as $\sigma_o = \frac{20\text{minutes/month}}{0.396\text{hour}} = 0.842/\text{month}$, and $\sigma_v = \frac{20\text{minutes/month}}{12\text{minutes}} = 1.667/\text{month}$. Based on a CDC report (Prevention 2009), generally incurable and ongoing, chronic diseases affect approximately 133 million Americans in 2009, representing more than 40% of the total population of this country. In 2009, 7 out of 10 deaths in the U.S. were due to chronic diseases, and the death population due to chronic disease is 1.706 million (Kochanek 2011). Thus, the monthly departure/death rate is estimated as $\delta = 1.706/(133 \times 12) = 0.00107/\text{month}$. In addition, the disease progression (i.e., transferring from a controlled health state to an uncontrolled health state) is estimated as $\gamma = 0.5$ /week (C. f. Prevention 2010). To estimate the new patients' health status, we consider a study that considers diabetes patients. Dall et al. (2011) present that 43% of new patients among their analytic sample had indications of uncontrolled diabetes. Hence, we assume $P_h = 1 - 43\% = 57\%$, which is the probability that the new arrival patient in virtual appointment is in controlled condition. We have no historical data to rely on in terms of the probability that the patients diagnosed in a controlled condition stay in controlled condition after a virtual treatment and the probability that the patients diagnosed in an uncontrolled condition improve in controlled condition after a virtual treatment. We note that for the remaining parameters (P_0, P_1, $P_{0|0}$, and $P_{1|1}$), we perform sensitivity analysis to investigate their effects on the capacity allocation.

Revenue and Cost. We assume that the workday for a clinic is 20 days per month. For the revenue and cost parameters of the office appointment, we refer to the study of (Lee 2009). Then, the marginal profit of office appointments is estimated as $r_o = \$131\ /\text{day} \times 20\ \text{day/month} = \$2620/\text{month}$, the fixed capacity cost is estimated as $c_o = \$\,84.6/\text{day} \times 20\ \text{day/month} = \$\,1692/\text{month}$, and the penalty cost is estimated as $f_o = 50/\text{day} \times 20\text{day/month} = 1000/\text{month}$ (Lee 2009). For the virtual appointments, there are no direct historical data to refer to. Thus, the marginal profit and penalty cost are estimated as the same as those of office, $r_{v_0} = r_{v_1} = r_o$, and $f_{v_0} = f_{v_1} = f_o$. But the total cost of virtual appointments care was 32 percent less than traditional hospital care (A. H. Association 2016). Therefore, the fixed capacity cost is estimated as $c_{v_0} = c_{v_1} = 1692/\text{month} \times (1 - 32\%) = 1150.56/\text{month}$. We have no information regarding to imperfect diagnosis penalty cost. But Pinnacle (2016) reports that 30% of annual healthcare spending in the United States is wasted due to unnecessary services and other inefficiencies. Here we assume that $e_{id} = 30\%\ r_v = 30\% \times 2620/\text{month} = 786/\text{month}$. Table 3.1 summarizes the value of the patients flow and profit-related parameters together with the sources from which they are estimated. We note that the parameter values listed in Table 3.1 are used in one of the scenarios. Then, we analyze several scenarios by considering the possible fluctuations in the parameter values.

3.4.2 Impact of Follow-up Rate

In the following sections, we investigate the change in the optimal capacity and average earnings, as some of the key parameters in the model change. The parameters we use in the model are obtained from the literature which are not specific to any healthcare organization. For that reason, we perform sensitivity analysis to investigate the optimal capacity for varying parameter values to ensure that the changes in the parameter values due to the different clinics' characteristics can be addressed. We show how the optimal capacity can change with respect to a change in other parameter values. Through our results, we show not only how the optimal capacity changes but also the range of the change in optimal capacity and average earnings.

Table 3.1: List of flow and profit-related parameters

Parameters	Values	Sources
Office service rate (μ_o)	2.525/hour	(Foundation 2018)
Virtual service rate (μ_v)	5/hour	(Valero 2000)
Office arrival rate (λ_o) Virtual arrival rate (λ_v)	7.943/month 0.922/month	(Bureau 2018, Disease 2017, Foundation 2018, Clinics.com 2019)
Office follow-up rate (σ_o)	0.842/month	(Services 2016)
Virtual follow-up rate (σ_v)	1.667/month	
Departure/death rate (δ)	0.00107/month	(C. f. Prevention 2009, B. a. Kucukyazici 2011)
Transfer rate from controlled to uncontrolled (γ)	0.5/week	(C. f. Prevention 2010)
Probability that the new arrival patient in virtual appointment is in controlled condition (P_h)	0.57	(Dall 2011)
Probability that patients diagnosed in a controlled condition stay in controlled condition (P_0)	0.9	
Probability that patients diagnosed in an uncontrolled condition improve in controlled condition (P_1)	0.7	
Probability controlled patient is diagnosed as controlled ($P_{0\|0}$)	0.9	
Probability uncontrolled patient is diagnosed as controlled ($P_{0\|1}$)	0.2	
Office marginal profit (r_o)	\$2620/month	(Lee 2009)
Virtual marginal profit (r_{v_0}, r_{v_1})	\$2620/month	
Office overflow penalty cost (f_o)	\$1000/month	
Virtual overflow penalty cost (f_{v_0}, f_{v_1})	\$1000/month	
Office fixed capacity cost (c_o)	\$1692/month	
Virtual fixed capacity cost (c_{v_0}, c_{v_1})	\$1150.56/month	(A. H. Association 2016)
Penalty cost of imperfect diagnosis (e_{id})	\$786/month	(Care 2016)

First, we study the impact of follow-up rate (i.e., σ_o, σ_v) on the optimal capacity. It has important relevance since reducing or increasing the follow-up rate implies less or more frequency of patients' visits. We vary the follow-up rate in a range between $0.5\sigma_o$ and $1.5\sigma_o$, and present the corresponding optimal capacities in Figure 3.2. The results show that as we increase the follow-up rate of office appointments, the optimal capacity of office appointments increases monotonically, while the total optimal capacity for virtual appointments does not change. This is reasonable because the increase in the office follow-up rate would result in an increase in the expected number of patients in the office appointments at the steady state, but not in the expected number of patients in the virtual appointments. It is also observed that the expected number of patients in virtual appointments is not a function of the office follow-up rate as stated with equation (A.11) in the Appendix. There occurs a slight increase in the optimal capacity of controlled patients, the reason for which is that the increasing follow-up rate of office appointments transfers more patients from the uncontrolled condition into the controlled condition. Similarly, as the follow-up rate of virtual appointments increases, it is observed that the optimal capacity for controlled health status increases more than the optimal capacity for uncontrolled health status.

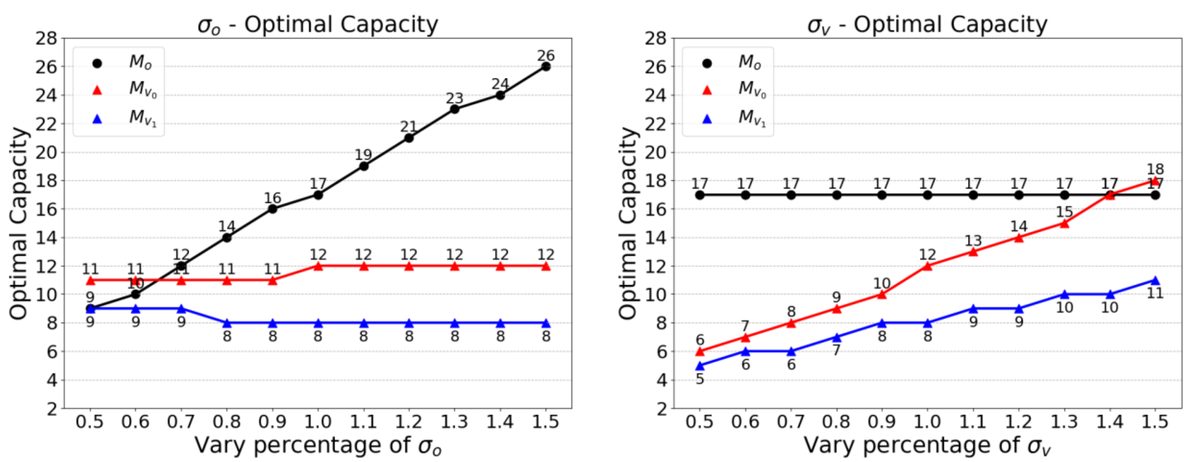

Figure 3.2: The impact of follow-up rate on the optimal capacity

Although in practice there is no distinction between capacity for uncontrolled and controlled conditions, this is an indication that the increasing follow-up rate of virtual appointments relatively reduces the number of patients in an uncontrolled health state. Thus, increasing the follow-up rate of either office or virtual appointments can improve the health condition of the patients, but it simultaneously increases the demand for office and virtual appointments, which is a challenge for the clinic. This result is consistent with the practice that as the average follow-up rates of the patients increase, the panel size of one physician would decrease. Thus, to serve the same number of patients, more physicians are needed for the clinic.

3.4.3 Impact of Capacity on Capacity Allocation Model with Capacity Constraint

Next, we analyze the impact of limited total capacity (TC) on the optimal capacity and the average earnings of the clinic. To this end, we first determine $\boldsymbol{M}^*$, which is the optimal capacity allocation vector for the unconstraFigur

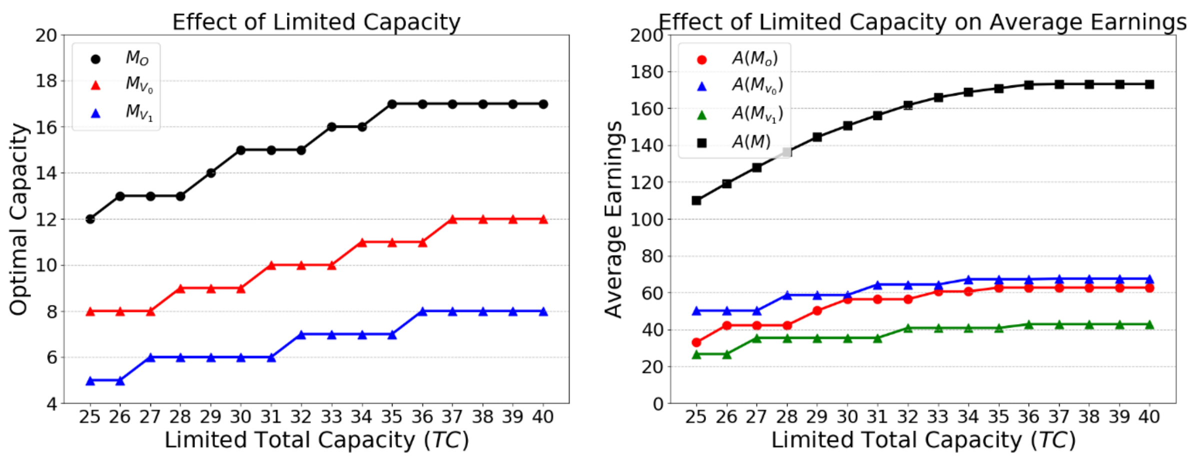

Figure 3.3: The impact of limited capacity (TC) on the optimal capacity and average earnings

In addition, as we increase the limited total capacity, the average long-run earnings of office and virtual appointments increase, and the marginal profit gain decreases, which is consistent with equation (3.19) that the first-order differential of the long-run average earnings is a monotonic decreasing function.

We also notice that the increasing ratio of the total long-run average earnings slows down. This is because of Algorithm 1. Algorithm 1 makes sure that in each iteration, one more capacity is added to the allocated appointments which can bring the most marginal earnings. In the next iteration, there are two options: first, if the capacity addition in the current iteration is not the same as the last iteration, the marginal profit that it brings to the clinic must be less than that of appointments added one more capacity in the last iteration; second, if capacity addition in the current iteration is the same as last iteration, the attribute that the first-order differential of the long-run average earnings' function is a monotonic decreasing function makes sure that the marginal profit that it brings to the clinic must be less than that of appointments added one more capacity in the last iteration. Hence, we can say that the benefits of adding one more capacity to the clinical system are diminishing marginally and the stopping rule for adding one more capacity to the clinic system is when the marginal earnings are not positive.

Table 3.2: The impact of limited capacity (TC) on the optimal capacity and average earnings

TC	M_o^*	$M_{v_0}^*$	$M_{v_1}^*$	$A_o(M_o^*)$	$A_{v_0}(M_{v_0}^*)$	$A_{v_1}(M_{v_1}^*)$	$A(M)$
25	12	8	5	32.981	50.219	26.708	109.907
26	13	8	5	42.289	50.219	26.708	119.216
27	13	8	6	42.289	50.219	35.450	127.957
28	13	9	6	42.289	58.686	35.450	136.425
29	14	9	6	50.208	58.686	35.450	144.343
30	15	9	6	56.410	58.686	35.450	150.545
31	15	10	6	56.410	64.393	35.450	156.252
32	15	10	7	56.410	64.393	40.839	161.641
33	16	10	7	60.634	64.393	40.839	165.866
34	16	11	7	60.634	67.276	40.839	168.749
35	17	11	7	62.721	67.276	40.839	170.836
36	17	11	8	62.721	67.276	42.858	172.854
37	17	12	8	62.721	67.534	42.858	173.112
38	17	12	8	62.721	67.534	42.858	173.112
39	17	12	8	62.721	67.534	42.858	173.112
40	17	12	8	62.721	67.534	42.858	173.112

3.4.4 Impact of Work Time on Capacity Allocation Model with Time Constraint

We also study the impact of limited working time (T_w) on the optimal capacity and the average earnings of the clinic. Recall that the optimal capacity for the unconstrained model is $\boldsymbol{M}^* = \left(M_o^{min} = 17, M_{v_0}^{min} = 12, M_{v_1}^{min} = 8\right)$. By considering the service rates (i.e., $\mu_o = 2.525/\text{hour}$ and $\mu_v = 5/\text{hour}$), the optimal capacity for the unconstrained model is around 11 hours. Hence, we vary the range of the limited working time from 8 to 12 hours. Then, we use Algorithm 2 to obtain the allocation decision for this problem. As shown in Figure 3.4 and Table 3.3, when $T_w = 8$ hours, the optimal capacities for office and virtual appointments for controlled and uncontrolled patients are 11, 11, and 7, and when $T_w = 10$ hours, the optimal capacities for office and virtual appointments for controlled and uncontrolled patients are 16, 11, and 7.

We observe that the change in the limited time affects the office appointment capacity more than the virtual appointment capacity. This is because the average service time of office appointments is nearly twice that of virtual appointments. Hence, if the limited time decreases, it becomes more profitable to reduce the office appointment capacity by one unit rather than reducing the virtual appointment capacity by two units. As it is shown, if the limited time is greater than 11 hours, the actual working time remains constant, which is consistent with the analysis in the Section 3.3.3 that as $\frac{1}{\mu_o} M_o^{min} + \frac{1}{\mu_v}\left(M_{v_0}^{min} + M_{v_1}^{min}\right) \leq T_w$, then $M_{T_w}^* = M^*$. It is the case that the physicians have enough time/resources to satisfy the need of patients through both office and virtual appointments.

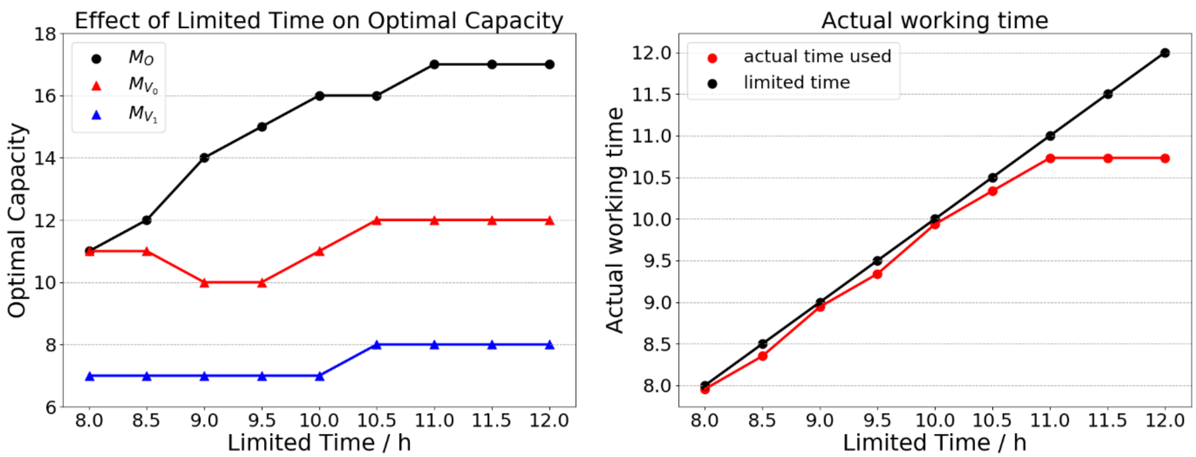

Figure 3.4: The impact of limited working time (T_w) on the optimal capacity

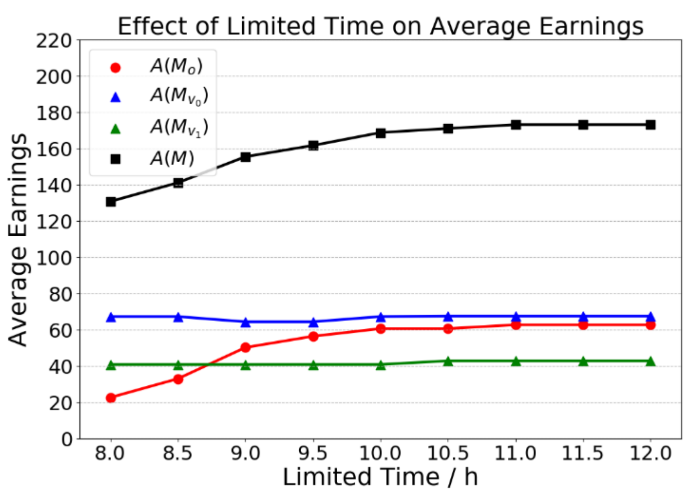

Figure 3.5: The impact of limited working time (T_w) on the average earnings

Table 3.3: The impact of limited working time (T_w) on the optimal capacity and average earnings

T_w	T_{use}	M_o^*	$M_{v_0}^*$	$M_{v_1}^*$	$A_o(M_o^*)$	$A_{v_0}(M_{v_0}^*)$	$A_{v_1}(M_{v_1}^*)$	$A(M)$
8	7.956	11	11	7	22.627	67.276	40.839	130.742
8.5	8.352	12	11	7	32.981	67.276	40.839	141.095
9	8.945	14	10	7	50.208	64.393	40.839	155.439
9.5	9.341	15	10	7	56.410	64.393	40.839	161.641
10	9.937	16	11	7	60.634	67.276	40.839	168.749
10.5	10.337	16	12	8	60.634	67.534	42.858	171.025
11	10.733	17	12	8	62.721	67.534	42.858	173.112
11.5	10.733	17	12	8	62.721	67.534	42.858	173.112
12	10.733	17	12	8	62.721	67.534	42.858	173.112

Table 3.4: The values of the $Z_i(M_i)$ in Algorithm 2

M_i	$Z_o(M_i)$	$Z_{v_0}(M_i)$	$Z_{v_1}(M_i)$
0	30.43	77.17	77.07
1	30.43	77.12	76.37
2	30.43	76.91	73.91
3	30.42	76.18	68.15
4	30.42	74.31	58.00
5	30.39	70.50	43.71
6	30.33	63.99	26.95
7	30.17	54.49	10.09
8	29.82	42.34	-4.74
9	29.14	28.53	
10	27.98	14.42	
11	26.14	1.29	
12	23.50	-9.89	
13	19.99		
14	15.66		
15	10.67		
16	5.27		
17	-0.22		

In addition, since we apply Algorithm 2, the physicians cannot fully spend the total limited working time as shown in the second graph in Figure 3.4. From Table 3.4, we can see that $Z_o(M_o^k = 11) > Z_{v_0}(M_v^k = 10)$, but when $T_w = 8$ hours, the optimal capacity for office and virtual appointments for controlled patients are 11 and 11, not 12 and 10, respectively. This is because in Algorithm 2, when adding one more capacity to the office appointments, the time exceeds the limited working time ($\frac{1}{\mu_o} \times 11 + \frac{1}{\mu_v} \times (10+7) + \frac{1}{\mu_o} = 8.152$ hours $> T_w = 8$ hours). In that case, Even through $Z_o(M_o^k = 11)$ is greater than $Z_v(M_v^k = 10)$, Algorithm 2 still adds one more capacity to virtual appointments if adding one capacity to virtual appointments does not exceed the limited working time and the marginal profit is positive (i.e., $\frac{1}{\mu_o} \times 11 + \frac{1}{\mu_v} \times (10+7) + \frac{1}{\mu_v} = 7.956$ hours $< T_w = 8$ hours). Also, we notice that as the T_w changes from 8.5 hours to 9 hours, we observe a decrease in the capacity of virtual appointments for controlled patients. This is because when $T_w = 9$ hours, there is enough time to add two more slots to the office appointments, such as from 12 to 14 (i.e., $\frac{1}{\mu_o} \times 12 + \frac{1}{\mu_v} \times (10+7) + \frac{2}{\mu_o} = 8.945$ hours$<$ $T_w = 9$ hours). From Table 3.4, we notice that $Z_o(M_o^k = 12) > Z_v(M_v^k = 10)$ and $Z_o(M_o^k = 13) > Z_v(M_v^k = 10)$. Based on the value of $Z_i\left(M_i^k\right)$, the algorithm adds two more slots to the office appointments. After that, there are not enough resources (i.e., time) to add one more capacity for virtual appointments ($9 - 8.945 = 0.055$ hour $< \frac{1}{\mu_v} = 0.2$ hour). Hence, $M_v^k = 10$, which accounts for the decreasing when T_w changes from 8.5 hours to 9 hours.

3.4.5 Imperfect Diagnosis Effect on Optimal Capacity and Average Earnings

In this section, we study the impact of the imperfect diagnosis on the optimal capacity and the average earnings of the clinic based on the capacity allocation model presented in Section 3.3.4. First, we have two separate analysis to study the effect of $P_{0|0}$ and $P_{0|1}$ independently. We let $P_{0|1} = 0.2$, and vary $P_{0|0}$ from 0.5 to 1, to find out the impact of $P_{0|0}$ on the optimal capacity (i.e., depicted in the left Figure 3.6); then, we let $P_{0|0} = 0.9$, and vary $P_{0|1}$ from 0 to 0.5, to find out the impact of $P_{0|1}$ on the optimal capacity (i.e., depicted in the right Figure 3.6). From Figure 3.6, we can see that as $P_{0|0}$ or $P_{0|1}$ increases, capacity of virtual appointments for controlled patients increases. Because the probability that a patient is diagnosed in the controlled condition increases no matter what his real condition is, the number of patients at node v_0 increases. Hence, more capacity should be allocated to node v_0 to provide treatment, improve earnings, and prevent patients overflow. While $P_{0|0}$ or $P_{0|1}$ decreases, capacity of virtual appointments for uncontrolled patients increases, and the reason is similar to the former. In addition, as $P_{0|0}$ or $P_{0|1}$ changes, the total capacity of virtual appointments (i.e., $M_{v_0} + M_{v_1}$) is stable in most of the cases, which shows that the accuracy of the diagnosis does not have a significant impact on the number of patients scheduled for virtual appointments. The accuracy of the diagnosis impacts the proportion of patients diagnosed in controlled and uncontrolled conditions.

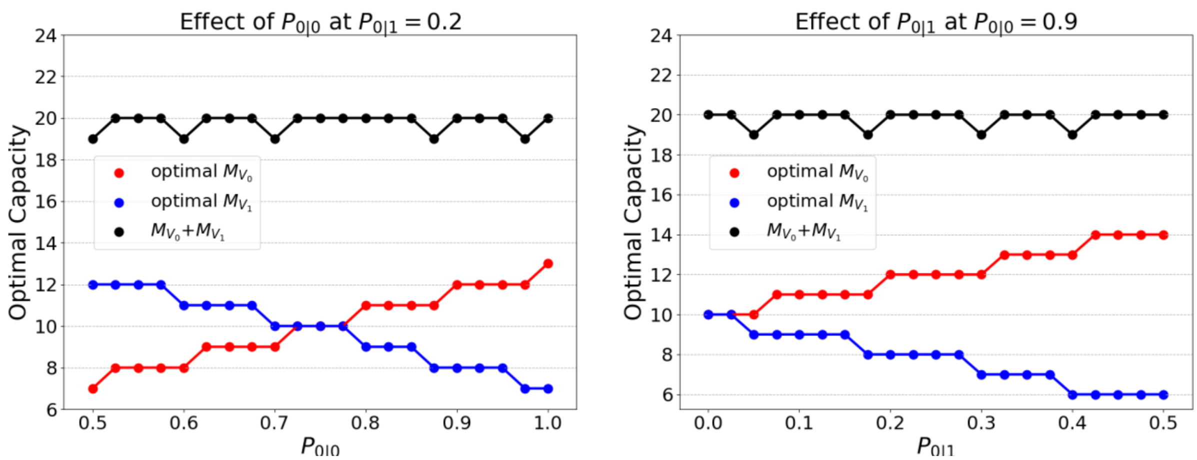

Figure 3.6: Effect of imperfect diagnosis on optimal capacity

In another analysis, we set $P_{0|0} \in \{0.7, 0.8, 0.9\}$, and $P_{0|1} \in \{0.3, 0.2, 0.1\}$, then we have 9 different settings for the sensitivity analysis. In each setting, we obtain the optimal capacity for office appointments, and the optimal capacity for virtual appointments in controlled and in uncontrolled states according to Proposition 1 (i.e., M_o^{min}, $M_{v_0}^{min}$, $M_{v_1}^{min}$). We also calculate average long-run earnings for each type of appointments and the total average long-run earnings (i.e., $A_o(M_o^{min})$, $A_{v_0}(M_{v_0}^{min})$, $A_{v_1}(M_{v_1}^{min})$, $A(M)$) for each setting. Besides, we calculate the average penalty cost on imperfect diagnosis (W_{id}) to find how the accuracy of diagnosis affects its cost. The result is shown in Table 3.5. We can see that the imperfect diagnosis of virtual appointment does not affect the optimal capacity and average long-run earnings of office appointments. However, it affects the capacity allocation of virtual appointments, and the impact of the imperfect diagnosis on the optimal virtual appointment capacity is analyzed in Figure 3.6. Its effect on the average long-run earnings is consistent with its impact on the optimal capacity. More specifically, as $P_{0|0}$ or $P_{0|1}$ increases, average long-run earnings of virtual appointments for controlled patients increases, while that of virtual appointments for uncontrolled patients decreases. In addition, we notice that, as the diagnosis becomes more accurate, i.e., $P_{0|0}$ increases and $P_{0|1}$ decreases, total average long-run earnings increases and the average penalty cost on imperfect diagnosis decreases as expected.

Table 3.5: Effect of imperfect diagnosis on optimal capacity and average earnings

$(P_{0\|0}, P_{0\|1})$	M_o^{min}	$M_{v_0}^{min}$	$M_{v_1}^{min}$	$A_o(M_o^{min})$	$A_{v_0}(M_{v_0}^{min})$	$A_{v_1}(M_{v_1}^{min})$	$A(M)$	W_{id}
(0.7, 0.3)	17	10	9	62.721	58.034	51.965	45.505	127.214
(0.7, 0.2)	17	9	10	62.721	51.640	58.384	64.590	108.155
(0.7, 0.1)	17	9	11	62.721	45.268	64.900	84.133	88.756
(0.8, 0.3)	17	11	8	62.721	65.835	44.339	69.402	103.493
(0.8, 0.2)	17	11	9	62.721	59.388	50.645	87.944	84.809
(0.8, 0.1)	17	10	10	62.721	53.135	57.059	107.127	65.789
(0.9, 0.3)	17	12	7	62.721	73.818	36.699	93.888	79.350
(0.9, 0.2)	17	12	8	62.721	67.534	42.858	112.069	61.043
(0.9, 0.1)	17	11	9	62.721	61.211	49.146	130.673	42.405

Finally, we perform a sensitivity analysis on the accuracy of the diagnosis on the capacity allocation model with an imperfect diagnosis under the time constraint. We set total limited working time $T_w \in \{8,10,12\}$ and the time unit is hour. Then, we vary $P_{0|0}$ from 0.5 to 1 and vary $P_{0|1}$ from 0 to 0.5 respectively, to find out the effect of the accuracy of diagnosis and the time constraint as shown in . There are six graphs in Figure 3.7. The independent variable is $P_{0|0}$ in the first column and $P_{0|1}$ in the second column. From the first to the third row, the total limited working time T_w equals to 8, 10, and 12, respectively. In terms of each column, where the total limited working time is fixed, we can see that as $P_{0|0}$ or $P_{0|1}$ increases, capacity of virtual appointments for controlled patients increases. While $P_{0|0}$ or $P_{0|1}$ decreases, capacity of virtual appointments for uncontrolled patients increases. Hence, no matter what total limited working time is, the effect of the accuracy of diagnosis is similar, just like their influence on the model without time constraint in Figure 3.6.

3.4.6 Comparison of Policies

In this section, we compare the total profits of some common benchmark policies with our proposed policies (i.e., optimal policy, Algorithm 1, and Algorithm 2). As benchmark policies, we consider three varying ratios of office appointment capacity to virtual appointment capacity (i.e., M_o/M_v): (i) Policy-1: $M_o/M_v = 2$, (ii) Policy-2: $M_o/M_v = 1$, and (iii) Policy-3: $M_o/M_v = 1/2$.

Initially, we consider that the virtual appointment capacities allocated for controlled and uncontrolled patients are equal to each other. For comparison, we analyze several scenarios by varying the parameter values. As the number of varying parameters increases, the number of scenarios and the complexity of the analyses increase. Hence, considering that the impact of the δ

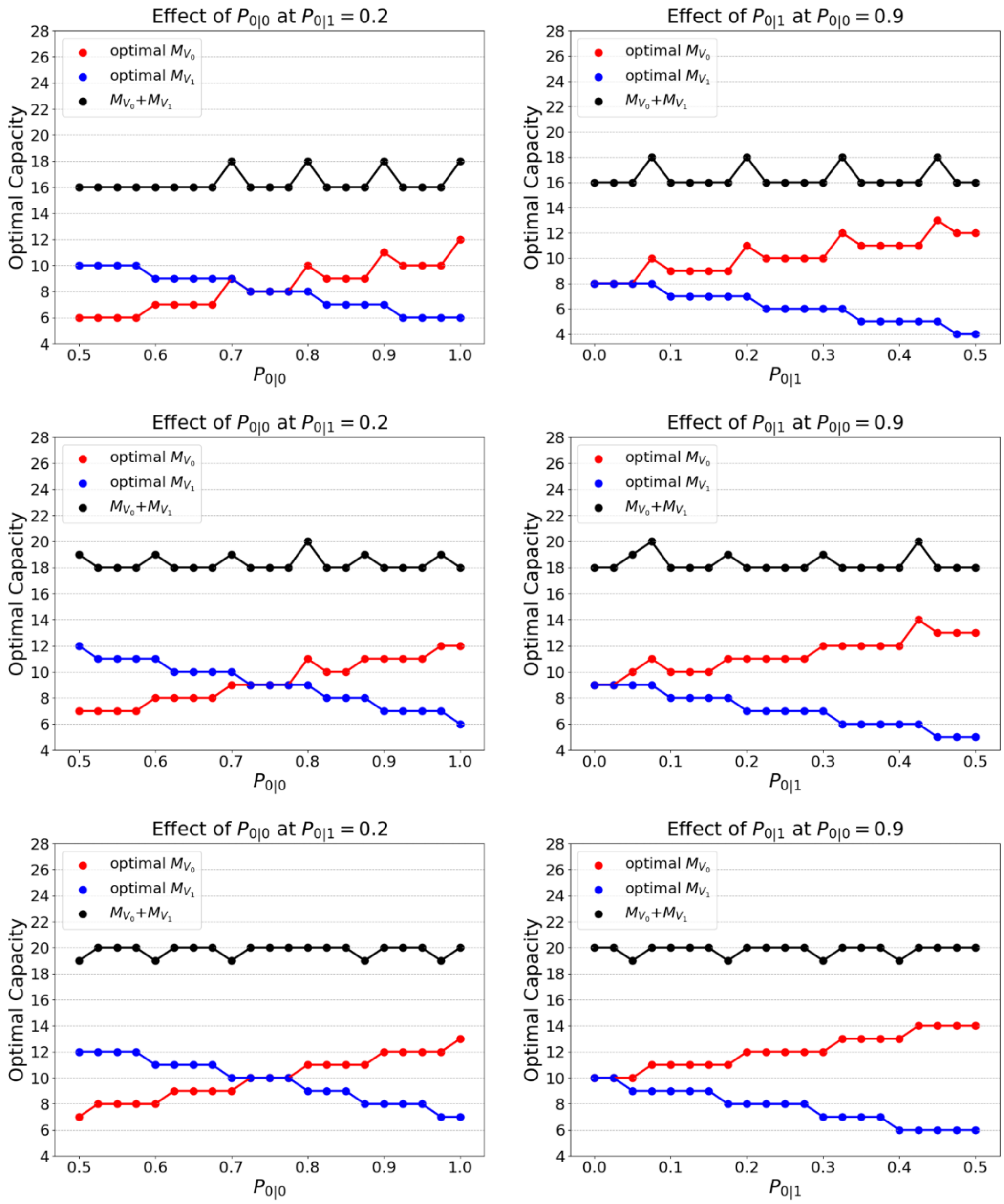

Figure 3.7: Effect of imperfect diagnosis and time constraint on optimal capacity

and γ variables on the capacity allocation decisions can be small and that preserving the relationship of $\mu_o \leq \mu_v$ is important, we keep these variables constant. For all 16 remaining parameters, we use two possible values (i.e., low, high). We use the following formulas to calculate the low and high levels for each parameter:

$$\text{Low Value of a Parameter} = (1 - \text{Fluctuation Rate}) \times \text{Orginal Parameter Value} \quad (3.27)$$

$$\text{High Value of a Parameter} = (1 + \text{Fluctuation Rate}) \times \text{Orginal Parameter Value} \quad (3.28)$$

By considering all possible combinations, we evaluate $2^{15} = 32{,}768$ scenarios for the base capacity allocation model, and for the model with the capacity and time constraint, we analyze $2^{16} = 65{,}536$ scenarios as we also change the parameter TC and T_w. In Table 3.6 and Table 3.7, we present the solutions obtained from the proposed algorithms (i.e., the ratio of M_o/M_v, M_{v_0}/M_{v_1}, and $A(M)$) and the comparison results of the policies (i.e., Policy-1, Policy-2, and Policy-3) with respect to the proposed algorithms for the base capacity allocation model, the capacity-constrained model, and the time-constrained model. Table 3.6 shows the results for a fluctuation rate of 5% while Table 3.7 shows the results for a fluctuation rate of 10%. To calculate the percent gap between the profit function of the policies and of the proposed algorithms, we use the following formula:

$$\text{\% Profit Gap} = \frac{\text{Profit of Proposed Algorithm} - \text{Profit of Policy}}{\text{Profit of Proposed Algorithm}} \times 100\% \qquad (3.29)$$

In Table 3.6 and Table 3.7, we present the average, maximum, and minimum values obtained overall scenarios, and we observe that the optimal capacities for office and virtual appointments fluctuate as parameters change. When the three common policies are compared, we can see that Policy-2 (i.e., $M_o/M_v = 1$) is the best, even though the result becomes worse as fluctuation increases. For the base capacity allocation model, the optimal capacity allocation ratio (i.e., M_o/M_v) varies between 0.71 and 1.11 when the fluctuation rate is 5%, while it varies between 0.62 and 1.27 when the fluctuation rate is 10%. As expected, when the uncertainty in the parameter values increases, the optimal capacity allocation ratio varies more. It also shows that even if the fluctuation rate is high, it is not reasonable to use a capacity allocation ratio of less than 0.62 or more than 1.27. Similar to the base capacity allocation model, in the model with the capacity and time constraint, Policy-2 performs the closest to the proposed solutions, but the variation is more compared to the unconstrained model where the model with time constraint has the highest variability. In the model with capacity constraint, the suggested M_o/M_v ratio varies between 0.7 and 1.14 when the fluctuation rate is 5%, and it varies between 0.5 and 1.5 when the fluctuation rate is 10%. In the time-constrained model, the proposed M_o/M_v ratio changes between 0.64 and 1.11 when the fluctuation rate is 5%, while it changes between 0.54 and 1.29 when the fluctuation rate is 10%. According to the results of the proposed policies (i.e., optimal, Algorithm 1, and Algorithm 2), the average capacity allocation ratio M_o/M_v should be 0.89 for the unconstrained model, 0.92 for the capacity-constrained model, and 0.86 for the time-constrained model.

Table 3.6: Comparison of policies with benchmark policies when the fluctuation rate is 5%

	Optimal						% Profit Gap		
Unconstrained Model	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-1**	**Policy-2**	**Policy-3**
Average	17.45	19.75	0.89	11.59	8.16	1.43	47.61%	5.12%	28.60%
Max	20.00	23.00	1.11	14.00	10.00	1.86	83.31%	18.31%	53.08%
Min	15.00	17.00	0.71	9.00	6.00	1.11	25.40%	0.00%	12.81%

	Algorithm-1						% Profit Gap		
Model with Capacity Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-1**	**Policy-2**	**Policy-3**
Average	15.28	16.71	0.92	9.94	6.76	1.49	43.25%	6.20%	27.94%
Max	18.00	20.00	1.14	12.00	9.00	2.20	82.79%	24.33%	53.14%
Min	13.00	14.00	0.70	8.00	5.00	1.11	16.31%	0.00%	7.93%

	Algorithm-2						% Profit Gap		
Model with Time Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-1**	**Policy-2**	**Policy-3**
Average	16.50	19.25	0.86	11.30	7.95	1.43	47.58%	9.97%	28.57%
Max	20.00	23.00	1.11	14.00	10.00	2.00	86.77%	31.28%	58.69%
Min	14.00	16.00	0.64	9.00	6.00	1.11	25.20%	0.95%	10.88%

We also compare the distribution of total virtual appointment capacity for all models in Table 3.6 and Table 3.7. For the base capacity allocation model, the optimal M_{v_0}/M_{v_1} ratio varies from 1.11 to 1.86 when the fluctuation is 5%, and it varies between 0.88 and 2.43 when the fluctuation is 10%. For the model with capacity constraint the suggested M_{v_0}/M_{v_1} ratio ranges between 1.11 and 2.2 when the fluctuation rate is 5%, while this ratio ranges between 0.88 and 3.33 when the fluctuation rate is 10%. Finally, for the time-constrained model, the optimal M_{v_0}/M_{v_1} ratio changes between 1.1 and 2 for the fluctuation rate of 5%, while this change is between 0.88 and 2.6 for the fluctuation rate of 10%. For all models, we observe that the average suggested M_{v_0}/M_{v_1} ratio is greater than 1 which suggests that more capacity should be allocated for the patients in the controlled state. These results also suggest that virtual appointments can be used to complement office appointments where the patients in controlled health state can be scheduled for virtual appointments more frequently than the others.

Table 3.7: Comparison of policies with benchmark policies when the fluctuation rate is 10%

	Optimal						% Profit Gap		
Unconstrainted Model	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-1**	**Policy-2**	**Policy-3**
Average	17.41	19.74	0.89	11.61	8.12	1.47	50.69%	7.88%	31.41%
Max	22.00	25.00	1.27	17.00	12.00	2.43	136.90%	50.33%	101.14%
Min	13.00	15.00	0.62	7.00	5.00	0.88	12.34%	0.00%	3.56%

	Algorithm-1						% Profit Gap		
Model with Capacity Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-1**	**Policy-2**	**Policy-3**
Average	15.10	16.61	0.92	9.90	6.71	1.54	47.83%	9.71%	31.07%
Max	20.00	21.00	1.50	14.00	10.00	3.33	155.82%	65.40%	101.58%
Min	10.00	12.00	0.50	6.00	3.00	0.88	3.43%	0.00%	0.00%

	Algorithm-2						% Profit Gap		
Model with Time Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-1**	**Policy-2**	**Policy-3**
Average	16.15	19.05	0.86	11.24	7.80	1.49	52.66%	15.00%	33.45%
Max	21.00	25.00	1.29	17.00	12.00	2.60	170.39%	101.08%	136.46%
Min	13.00	14.00	0.54	7.00	5.00	0.88	11.44%	-0.77%	3.08%

In terms of the profit, as the fluctuation rate increases from 5% to 10%, the % Profit Gap under Policy-2 for the base capacity allocation model increases from 5.12% to 7.88%, which shows that the optimal policy brings more profit to the clinic than Policy-2, and this gap is larger under more fluctuation. The similar tendency also works for model with capacity and time constraint. Another important result is that as the fluctuation rate is 10%, the % Profit Gap under all three policies with time constraint are greater than 100%, which means the long-run average earnings under these policies are negative. We check the parameters settings for these scenarios and find that the marginal profit for office appointments, $r_o' = 0.9r_o = 2358/$month, the fixed capacity cost for office appointments, $c_o' = 1.1c_o = 1861.2/$month, and the overflow penalty cost for office appointments, $f_o' = 1.1f_o = 1100/$month. The virtual appointments have the similar scenario setting, which is the worst situation for the clinic. In this system, the long-run average earnings are much more sensitive to the allocation of the capacity and the three fixed-ratio policies are not suitable in this situation and lead to negative long-run average earnings, which results in % Profit Gap is greater than 100%. In addition, we notice that the minimum % Profit Gap under "Algorithm 2 vs Policy-2" with time constraint as the fluctuation rate is 10% is negative, which shows that Algorithm 2 cannot ensure the optimal allocation of the capacity just as we explain in the Section 3.3.3, but the difference to the optimal is not much, which is accepted.

Next, we note that the capacity allocation ratio of the proposed policies (i.e., optimal, Algorithm 1, and Algorithm 2) varies for each scenario and it is not fixed. However, we consider the average proposed capacity allocation ratios and present two more policies by fixing them to test its performance. For the base capacity allocation model, we analyze two more policies, namely, Policy-4 and Policy-5, where $M_o/M_v = 0.89$ and $M_{v_0}/M_{v_1} = 1$for Policy-4 and $M_o/M_v = 0.89$ and $M_{v_0}/M_{v_1} = 1.5$ for Policy-5. In Policy-5, which is based on Algorithm 2, we further consider varying capacity allocations among controlled and uncontrolled virtual appointments. We present our results in Table 3.8 and Table 3.9 for fluctuation rates of 5% and 10%, respectively. According to our results, both Policy-4 and Policy-5 perform better than the previous policies. For a fluctuation rate of 5%, the variation between the smallest and largest gap is 0%-18.85% for Policy-4 and 0%-4.71% for Policy-5. Also, on average Policy-4 deviates from the proposed policy by 5.26% while Policy-5 deviates by 1.04%. On the other hand, when the fluctuation rate is 10%, both the average percent gap and the range between the maximum and the minimum percent gap increase as expected. Among the suggested fixed policies, Policy-5 is the best fixed policy which suggests that more virtual appointment capacity should be allocated for controlled patients, compared to uncontrolled patients. The reason for this is that treatment and diagnosis from virtual appointments are imperfect. Since the treatment and diagnosis effectiveness of virtual appointments is not as good as that of office appointments, they can be mostly used to follow-up, controlled patients. Office appointments, on the other hand, can be used to treat both controlled and uncontrolled patients. Although Policy-5 performs very well, as the uncertainty in parameter values occurs, a policy with a fixed allocation ratio becomes worse and less stable. The long-run average earnings are much more sensitive to the allocation of the capacity. Hence, applying the optimal policy dynamically would be better for the clinics. We note that the comparison results of the Policy-4 and Policy-5 with the proposed algorithms for the capacity-constrained model and time-constrained model are similar to the unconstrained model, and we included those results in Appendix.

Table 3.8: Comparison of Policy-4 and 5 with benchmark policies when the fluctuation rate is 5%

	Optimal						**% Profit Gap**	
Unconstrained Model	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-4**	**Policy-5**
Average	17.45	19.75	0.89	11.59	8.16	1.43	5.26%	1.04%
Max	20.00	23.00	1.11	14.00	10.00	1.86	18.85%	4.71%
Min	15.00	17.00	0.71	9.00	6.00	1.11	0.00%	0.00%

Table 3.9: Comparison of Policy-4 and 5 with benchmark policies when the fluctuation rate is 10%

	Optimal						**% Profit Gap**	
Unconstrained Model	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-4**	**Policy-5**
Average	17.41	19.74	0.89	11.61	8.12	1.47	8.11%	4.01%
Max	22.00	25.00	1.27	17.00	12.00	2.43	51.71%	21.66%
Min	13.00	15.00	0.62	7.00	5.00	0.88	0.00%	0.00%

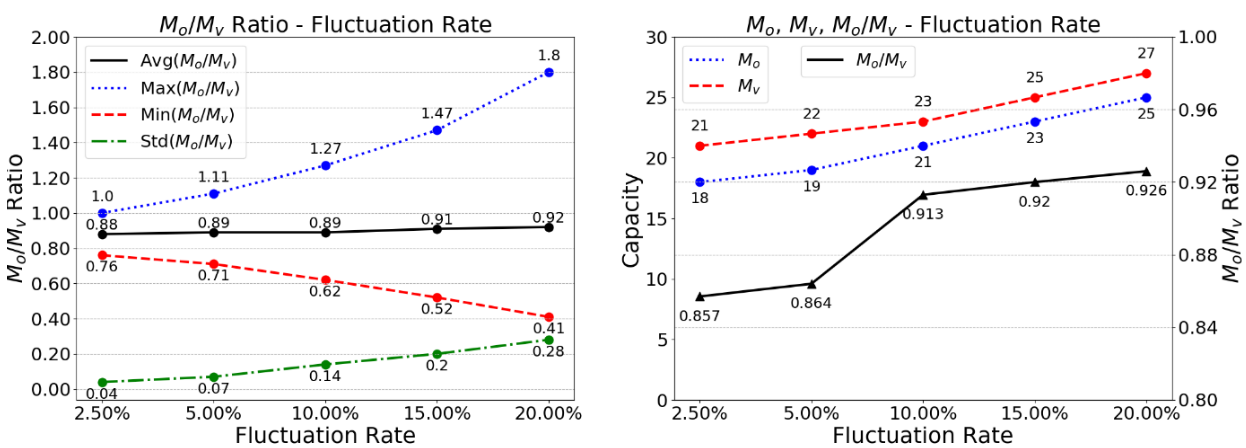

Figure 3.8: The impact of fluctuation rate to the office to virtual appointment ratio, M_o/M_v

After the comparison of different policies under 5% and 10% of fluctuation of parameters' values, it shows that a fixed policy might lead to negative long-run average earnings with high fluctuation of the values of the parameters. Hence, we next analyze the impact of the fluctuation rate on the optimal policy. We vary the fluctuation rate between 2.5% and 20%. In Figure 3.8, we present the impact of the fluctuation rate on the average, maximum, minimum, and standard deviation of the optimal M_o/M_v ratio for the base capacity allocation model. As the fluctuation rate increases, the average increases slightly, but the range of the ratio of office appointment capacity to virtual appointment capacity and the fluctuation of this ratio become larger. This indicates that the increasing uncertainty in parameter values makes the allocation decision harder for the policymakers. Although our results do not explore all the possible scenarios, we believe they are informative to the policy-makers in the clinic to better allocate the capacity of the office and virtual appointments to maximize the average long-run earnings of the clinic.

3.5 Conclusion

Virtual appointments, consisting of email, phone, cloud-electrocardiography, and online consultations, are increasingly changing our point of view of traditional office appointments, which makes the integration of the virtual appointments and the office critical for healthcare providers. Nowadays, many health clinics and hospitals are transitioning in virtual health services, which brings several operational challenges. Hence, capacity allocation between these two interventions with different effectiveness becomes an important but complex problem in the field of healthcare.

In this chapter, we use a migration network to model chronic patients' flows between controlled condition, uncontrolled condition, office appointments, and virtual appointments. Our model further reflects the varying effectiveness of office and virtual appointments in treatment and diagnosis. Then, we build a newsvendor optimization model to determine how to allocate the capacity of office and virtual appointments to maximize the network's long-run average earnings. We present a base model and three extended models with capacity and time constraints to demonstrate the potential use of this model in the clinic network under different scenarios. Through numerical studies, we present one clinic network with parameters estimated from the previous literature and open data sources. We study the use of our three optimization models under different scenarios and perform sensitivity analysis for the comparison of different allocation policies.

Our numerical studies bring us several insights about the clinic system and the application of virtual appointments. First, the follow-up rate is an important parameter in chronic care, and it represents the patient's revisit frequency. Our numerical results show that increasing follow-up rate for both office and virtual appointments tend to improve the health condition of the patients and transfer more patients from uncontrolled condition to controlled condition. However, at the same time, higher follow-up rates reduce the panel size of physicians, thus, to serve the same number of patients, the clinic should allocate more capacity of the office and virtual appointments. Thus, considering the potential number of patients that a clinic should serve, and the capacity that one clinic has, the policymakers of the clinic have to cooperate with the physicians to determine an appropriate follow-up rate. Second, the comparison of the results of capacity allocation under capacity and time constraint shows that the scarce service time makes the virtual appointments a more efficient way to provide service for the patients since they allow healthcare providers to serve more patients during the same time compared to office appointments in a more cost-effective way. Thus, the policymakers should consider providing more virtual appointments in the clinic system to improve the efficiency of the service. Finally, our comprehensive sensitivity analyses show that the average long-run earnings are sensitive to different fixed-allocation-ratio policies, but an appropriate fixed-allocation-ratio policy would be easy to apply for the policymakers in the clinic and can improve clinics' earnings. As the fluctuation rate increases significantly, the fixed-allocation-ratio policy will be no longer robust. Thus, updating the parameters frequently and applying the optimal policy dynamically is complex but better for the policymakers. Our results also suggest that virtual appointments should be used more to follow up controlled patients than for treating uncontrolled patients. Also, although virtual appointments are not as effective as office appointments, due to their lower costs, they have equal importance to office appointments.

Chapter 4: Optimization of Patient Revisit Intervals

In this chapter, we investigate the patients' revisit intervals in chronic care for virtual and office appointments. The section is organized as follows. Section 4.1 is the introduction to the optimization of patient revisit intervals in chronic care. Section 4.2 presents two different optimization models to determine the optimal follow-up rate that maximizes the average earnings of the clinics. Section 4.3 presents numerical studies for the clinic network. Section 4.4 outlines the conclusion and provides some future research directions.

4.1 Introduction

Scheduling patients for virtual and office appointments is one of the most important operational decisions for the health care providers. A good follow-up rate (i.e., patient revisit interval) leads to the efficient use of the clinics' resources, and it can ensure the regular and consistent care of patients to keep their chronic diseases under control. On the other hand, if there is not an efficient policy to determine the follow-up rates, this can cause capacity over bookings or can impact patients' health status negatively. Hence, it is important to have efficient policies to determine the follow-up rate (i.e., revisit interval) for both virtual and office appointments.

In this study, we consider the integration of office and virtual appointments and aim to answer the following operational question:

- Given the number of patients at the steady state, what should be the optimal follow-up rate decision that maximizes the clinic's overall earnings?

To address the above questions, we apply the migration network in Chapter 3 to analyze the flow of patients and determine the number of patients that need virtual and office appointments in the steady state. We further develop two follow-up rate optimization models by using the output of the migration network model to determine the optimal follow-up rates where the overall earnings are maximized. One of the models is a linear programming model that considers the linear overbooking cost, while the second one is the nonlinear model that considers a nonlinear convex overbooking cost function. With our numerical experiments, we show the change of the optimal follow-up rate for varying conditions. We note that the models that present in this chapter are different from Chapter 3 since we investigate the patients' revisit intervals.

4.2 The Mathematical Model

In this section, we consider a chronic care network with both office and virtual appointments, where physicians provide a diagnosis of the patients' condition and the required treatment through either office or virtual appointments. We describe the linear and the non-linear optimization models which are used to obtain the optimal follow-up rate decisions to maximize the clinics' average earnings.

4.2.1 The Model with Linear Overbooking Cost

In this section, we build a linear programming model to determine the optimal follow-up rate for office and virtual appointments to maximize the overall earnings of a clinic. We define $\bar{\sigma_i}, i \in \{o, v\}$ to denote the upper-bound of the follow-up rate of the office and virtual appointments, respectively. For other parameters, we use the same notation that we define in Chapter 2. We summarize that notation as follows:

$M_i, i \in \{o, v\}$: the capacity of office and virtual appointments, respectively;

$r_i, i \in \{o, v\}$: the marginal profit of each patient treatment through office and virtual appointments, respectively;

$c_i, i \in \{o, v\}$: the fixed costs of office and virtual appointments, respectively;

$f_i, i \in \{o, v\}$: the extra variable cost of office and virtual appointments due to overbooking.

In the model, we consider two decision variables: the follow-up rate of office and virtual appointments, σ_o and σ_v. We also define two intermediate decision variables to describe the overbooking cost as follows: the number of overbooked office and overbooked virtual appointments, u_o and u_v. As defined above, the expected number of patients at node i in the steady state is $\alpha_i, i \in \{o, v\}$. Since α_i is the function of σ_o and σ_v, we denote it as $\alpha_i(\sigma_i)$ in our optimization model. Then, our model with linear overbooking cost can be defined as follows:

$$\max A(\sigma_k) = \sum_{i \in \{o,v\}} r_i \alpha_i(\sigma_i) - \sum_{i \in \{o,v\}} c_i M_i - \sum_{i \in \{o,v\}} f_i u_i \tag{4.9}$$

Subject to

$$u_i \geq 0 \qquad \forall i \in \{o, v\} \qquad (4.10)$$

$$u_i \geq \alpha_i(\sigma_i) - M_i \qquad \forall i \in \{o, v\} \qquad (4.11)$$

$$\sigma_i \leq \bar{\sigma}_i \qquad \forall i \in \{o, v\} \qquad (4.12)$$

$$\sigma_i \geq 0 \qquad \forall i \in \{o, v\} \qquad (4.13)$$

Equation (4.9) states the objective function which is to maximize the overall clinic's earning. In the objective function, the first term is the marginal profit, the second term is the total cost of the appointment capacity, and the last term represents the extra cost due to patient overflow. Constraint (4.10) and (4.11) ensures that the overbooking cost occurs only when there are overbooked patients and the number of overbooked patients is the difference between the average number of patients and the capacity. Constraint (4.12) and (4.13) ensures that the follow-up rate is greater than 0 and less than the upper-bound.

4.2.2 The Model with Non-linear Overbooking Cost

In practice, when the clinic is slightly overbooked, its impact on the cost may be tolerated and the marginal profit of overbooking a few patients can be still positive. However, if the clinic is highly overbooked, its impact on cost would be high as well and the marginal profit can be negative. Hence, to reflect this change, it is better to model the overbooking cost as a nonlinear function. In the literature, there are studies that define and use the non-linear cost structure in their model (Lee 2009, LaGanga 2012). Similar to the literature and different from the model defined in Section 4.2.1, we define the overbooking cost function as a nonlinear increasing convex function and we use an exponential function to define this relationship. Thus, the marginal cost of overbooking a patient is increasing as the number of overbooked patients increases. More specifically, we redefine the overbooking function as$u_i' = max\{0, e^{\alpha_i(\sigma_i)-M_i} - 1\}$. In this relation, we ensure that the overbooking cost occurs only when the number of scheduled patients is greater than the assigned capacity. Since the smallest value of the term $e^{\alpha_i(\sigma_i)-M_i}$ can be '1', we subtract '1' from the exponential term to ensure that if the capacity equals to the number of scheduled patients no overbooking cost occurs. Then our model can be defined as follows:

$$\max A(\sigma_k) = \sum_{i\in\{o,v\}} r_i \alpha_i(\sigma_i) - \sum_{i\in\{o,v\}} c_i M_i - \sum_{i\in\{o,v\}} f_i u_i \qquad (4.14)$$

Subject to

$$u_i \geq 0 \qquad \forall i \in \{o, v\} \tag{4.15}$$

$$u_i \geq e^{\alpha_i(\sigma_i)-M_i} - 1 \qquad \forall i \in \{o, v\} \tag{4.16}$$

$$\sigma_i \leq \bar{\sigma}_i \qquad \forall i \in \{o, v\} \tag{4.17}$$

$$\sigma_i \geq 0 \qquad \forall i \in \{o, v\} \tag{4.18}$$

Since the described model is non-linear, it is not easy to find the optimal solution through standard solvers where the starting points would impact the final solution. Also, since our model does not include several constraints, we investigate the Karush–Kuhn–Tucker (KKT) necessity and sufficiency conditions to determine the optimal solution. To solve this problem, for $i \in \{o, v\}$, we define the sub-problem as follows:

$$max\, A(\sigma_i) = r_i \alpha_i(\sigma_i) - c_i M_i - f_i u_i \tag{4.19}$$

$$= r_i \left[\frac{\sigma_i(\lambda_o + \lambda_v)}{\delta \mu_i} + \frac{\lambda_i}{\mu_i}\right] - c_i M_i - f_i u_i \tag{4.20}$$

Subject to

$$g_1(\sigma_i, u_i) = -u_i \leq 0 \tag{4.21}$$

$$g_2(\sigma_i, u_i) = e^{\alpha_i(\sigma_i)-M_i} - 1 - u_i \leq 0 \tag{4.22}$$

$$g_3(\sigma_i, u_i) = \sigma_i - \bar{\sigma}_i \leq 0 \tag{4.23}$$

$$g_4(\sigma_i, u_i) = -\sigma_i \leq 0 \tag{4.24}$$

Where $g(\sigma_i, u_i)$ represents the constraints. The objective is a linear function and the constraints are all convex in the format of $g(\sigma_i, \mu_i) \leq 0$. Let $L(\sigma_i, u_i, \vec{\lambda})$ be the Lagrangian function to represent the Lagrangian relaxed objective function, and λ_i be the corresponding Lagrangian multiplier of each constraint. Then, we can use sufficiency of KKT conditions to reach the global optimum as below:

$$\max L(\sigma_i, u_i, \vec{\lambda}) = r_i \left[\frac{\sigma_i(\lambda_o + \lambda_v)}{\delta \mu_i} + \frac{\lambda_i}{\mu_i}\right] - c_i M_i - f_i u_i - \sum_{j=1}^{4} \lambda_j g_j(\sigma_i, u_i) \tag{4.25}$$

Subject to

$$\lambda_j g_j(\sigma_i, u_i) = 0 \qquad \forall j \in \{1,2,3,4\} \tag{4.26}$$

$$\lambda_j \geq 0 \qquad \forall j \in \{1,2,3,4\} \tag{4.27}$$

We set $\frac{\partial L(\sigma_i, u_i, \vec{\lambda})}{\partial \sigma_i} = 0$ and $\frac{\partial L(\sigma_i, u_i, \vec{\lambda})}{\partial u_i} = 0$ (i.e., find its derivative with respect to follow-up rate and the overbooking level and set it equal to zero) to reach the optimality by solving the equation set numerically and obtain the optimal solution for $\sigma_i\ \forall i \in \mathcal{I}$. In the numerical studies, we apply this method to solve the problem.

4.3 Numerical Studies

In this section, we numerically analyze a clinic network with office and virtual appointments. We conduct numerical experiments to investigate the optimal follow rate with respect to a change in different parameter values. Since the study is the extension of Chapter 3, we refer to Chapter 3 for the estimation of the model parameters.

With the flow parameters, revenue and cost parameters discussed, we study the impact of the arrival rate of office appointments and the upper-bound of the follow-up rate. (i.e., λ_o, $\bar{\sigma}_o$, and $\bar{\sigma}_v$) on the optimal follow-up rate in both linear and nonlinear models. It has important relevance since reducing (resp. increasing) the new patients' arrival rate leads to less (resp. more) more patients in the entire system. With the fixed capacity of office and virtual appointments, the clinic and the physicians need to adjust the follow-up rate to the system. In the parameter estimation, we set the original patient follow-up rates as $\lambda_o = 7.94/\text{month}$ as described in Chapter 3. We vary this rate by 50-150%, from 3.97 to 11.91, and present the corresponding optimal follow-up rate in Figure 4.1. The two black lines are the upper-bound of the follow-up rate for office and virtual appointments, respectively. In the left figure, $\bar{\sigma}_o = 2/\text{month}$ and $\bar{\sigma}_v = 4/\text{month}$ while in the right figure, $\bar{\sigma}_o = 1.6/\text{month}$ and $\bar{\sigma}_v = 4/\text{month}$, which is 20% less than the former one. In Figure 4.1, we show the results of both linear and nonlinear models where the linear model is represented by the solid line, and the nonlinear model is represented by the dashed line. As observed, the behavior of the follow-up rate change is similar in linear and nonlinear models. The virtual follow-up rate changes between 1.6 and 4.2, and the office follow-up rate changes between 0.6 and 1.8 where the range of change for the virtual appointments is more than that of office appointments. As we increase the arrival rate of office appointments, the optimal follow-up rates for both types of appointment decrease monotonically with a decreasing rate (i.e., convex). The optimal follow-up rate in the nonlinear model is greater than that of the linear model, but the difference between the optimal follow-up rate for different models is decreasing when the arrival rate increases. The result is consistent with the practice that as the average arrival rate of office appointments increases, the total number of the patients in the whole clinic system increases. To serve more patients with a fixed capacity, the healthcare providers need to decrease the frequency to see the same patient but allocate time for more patients. Thus, when the arrival rates of new patients increase, for a given capacity, the patients' follow-up rate decrease to adjust to the increasing number of patients in the whole system.

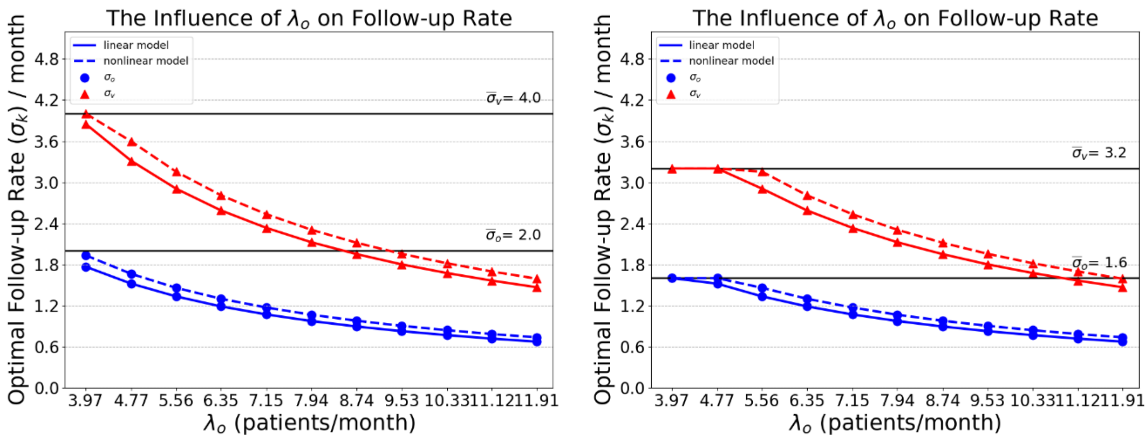

Figure 4.1: The impact of office arrival rate on the optimal follow-up rate

Next we study the impacts of marginal profit and the extra cost due to overbooking patients on the optimal follow-up rate and the overbooking level. To figure out the relationship, we define a new variable: r_i/f_i , $i \in \{o, v\}$, which is the ratio of marginal profit to the extra cost. In the linear model, the behavior of the overbooking cost is straightforward as it is a linear function. Hence, in Figure 4.2 and Figure 4.3, we analyze this change for the nonlinear model. From both figures, we can see that the relationship between the r_i/f_i ratio and optimal follow-up rate and the relationship between the r_i/f_i ratio and the overbooking function are increasing concave. More specifically, as we linearly increase the marginal profit or decrease the extra cost due to overbooking, the increase rate of the optimal follow-up rate and the overbooking function are decreasing. In addition, when the r_i/f_i ratio increases from 1 to 11, the optimal follow-up rate of virtual appointments increases by nearly 0.25/month and the overbooking level approaches to 2.5 patients. This means that the impact of office arrival rate on the optimal follow-up rate and overbooking function is more than the impact of the r_i/f_i ratio.

Finally, we study the impact of the capacity of the office and virtual appointments on the optimal follow-up rate for both linear and nonlinear models. Since increasing the capacity of appointments means that the clinic can serve more patients during the same time, it is important to investigate its impact on the optimal follow-up rate. We vary the capacity of office appointments from 14 to 26, and the capacity of virtual appointments from 16 to 28 and present the corresponding optimal follow-up rate in Figure 4.4.

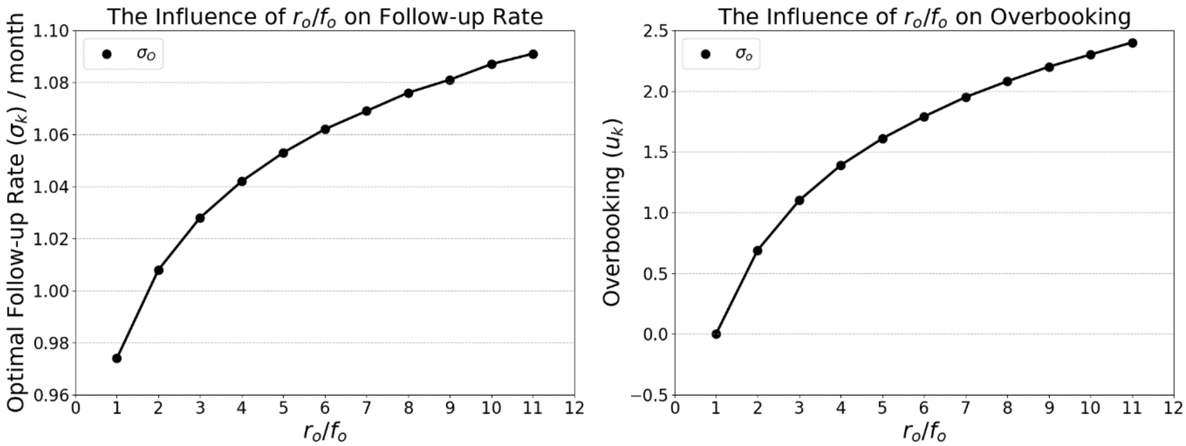

Figure 4.2: The impact of r_o/f_o on optimal follow-up rate and overbooking level of office appointments

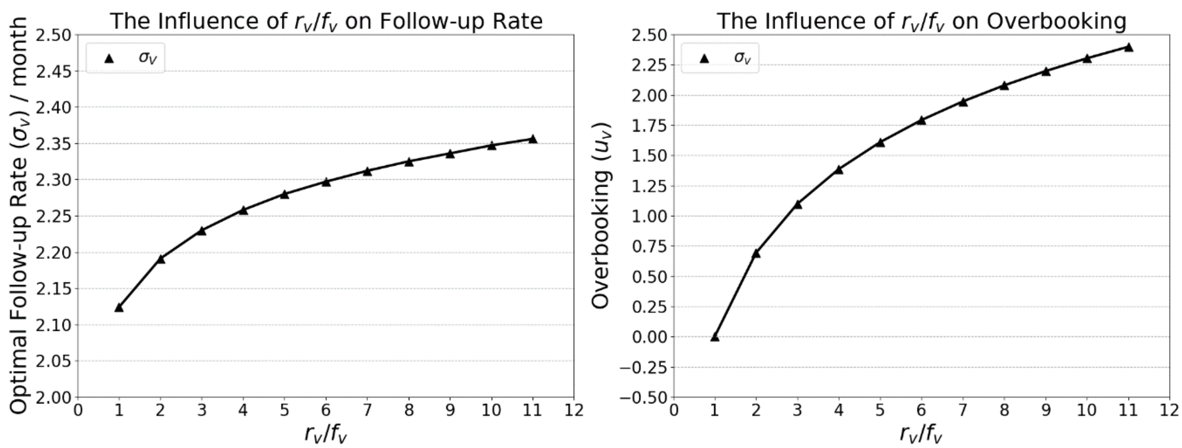

Figure 4.3: The impact of r_v/f_v on optimal follow-up rate and overbooking level of virtual appointments

In Figure 4.4, we show the results of both linear and nonlinear models where the linear model is represented by the solid line, and the nonlinear model is represented by the dashed line. The red line represents the virtual appointments while the blue line represents the office appointments. As observed, the behavior of the follow-up rate change is similar in linear and nonlinear models. The relationship between the follow-up rate and the capacity is linear for both office and virtual appointments. However, office and virtual appointments are independent of the observation. More specifically, the capacity of office appointments just affects the follow-up rate of office appointments and does not affect the follow-up rate of virtual appointments. The influence of the capacity of virtual appointments is similar. The virtual follow-up rate changes from 1.6 to 2.7 when the capacity increases 12 while the office follow-up rate changes from 0.65 to 1.25 when the office capacity increases 12. Thus, we conclude that the optimal follow-up rate of virtual appointments is more sensitive to the changes in capacity. The optimal follow-up rate in the nonlinear model is greater than that of in the linear model, but the difference between models for office appointments is less than that of virtual appointments. The result is consistent with the practice that as the capacity increases, for a given new patients' arrival rate, the physicians need to increase the follow-up rate and shorten the patients' revisit intervals to ensure the maximization of the profits of the clinic.

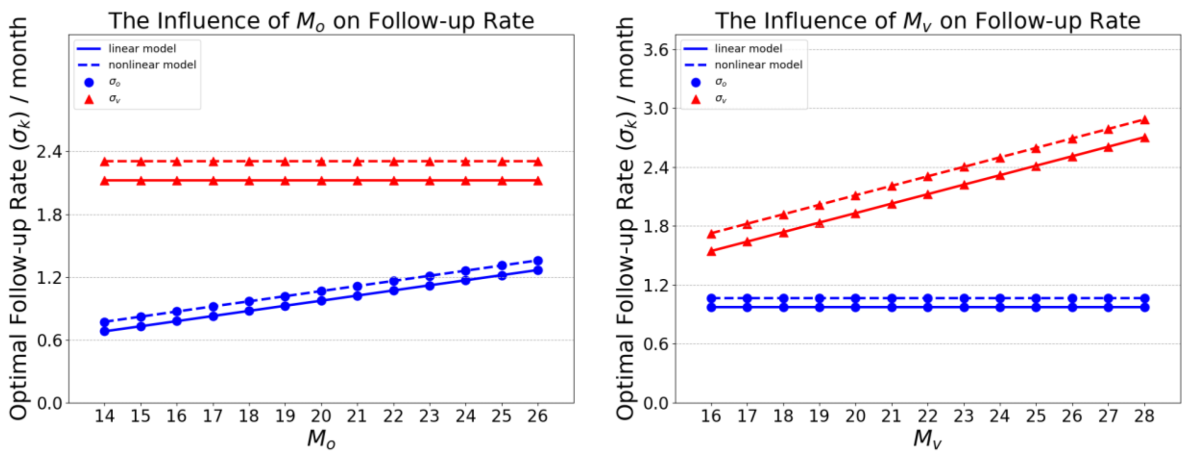

Figure 4.4: The impact of M_i on the optimal follow-up rate

4.4 Conclusion

Virtual appointments are gaining ground rapidly and they are used to complement and substitute for the office appointments. Nowadays, many health clinics and hospitals are transitioning in virtual health services and integrating virtual appointments with office appointments. However, this integration brings several operational challenges and the determination of the revisit intervals for office and virtual appointments becomes an important but complex problem.

In this chapter, with the foundation of the migration network model in Chapter 3, we build two optimization models by considering linear and nonlinear cost functions to determine the optimal follow-up rates of office and virtual appointments that maximizes the clinic's long-run average earnings. Through numerical studies, we investigate the impact of parameters on the value of the optimal follow-up rates. Based on the numerical studies, we conclude that when the arrival rate of office appointments increases, the follow-up rates of both office and virtual appointments decrease to ensure the service of more new patients. When the total overbooking cost is modeled as a nonlinear function, the marginal earnings and the marginal overbooking cost are also found to impact the optimal follow-up rate based on the mathematical model and the optimal follow-up rate is shown to be greater than that of the linear model based on the numerical studies.

Chapter 5: Capacity Planning and Patients Scheduling

In this chapter, we study patient scheduling decisions for office and virtual appointments along with capacity management decisions. Different from Chapter 3, this chapter applies stochastic programming to determine the capacity planning decisions and the patients' schedule for office and virtual appointments. The chapter is organized as follows. Section 5.1 gives an introduction. Section 5.2 presents the two-stage stochastic programming model. Section 5.3 presents numerical studies, estimates parameters, and provides the sensitivity analysis results that illustrate the application of the model. Section 5.4 outlines the conclusion of the chapter and provides some future research directions.

5.1 Introduction

In Chapter 3, we consider solely capacity planning decisions by considering the flow of patients among different states of the clinical network. However, in practice, the capacity planning decisions are also considered along with patient scheduling decisions. Moreover, patients' health statuses also impact both capacity management and patient scheduling decisions. In Chapter 3, we include the diagnosis only for the virtual appointments but not for the office appointments. In this chapter, we incorporate the diagnosis for office appointments as well.

To reflect the influence of the stochastic health status and patient scheduling decisions, we develop a two-stage stochastic programming model considering different scenarios that patients' health status is realized stochastically. The decision process is to determine the capacity of the office and virtual appointments along with the patient scheduling decisions to maximize the average patients' health status among all scenarios. In our model, we further incorporate patients' disease progression. With our numerical experiments, we show the changes in capacity allocation decisions for different settings as model parameters vary.

5.2 Two-stage Stochastic Programming Model

In this section, we build a two-stage stochastic programming model. More specifically, the decision process we consider for the capacity planning and the patients' scheduling optimization framework consists of two decision periods where $t \in T = \{1,2\}$. In the first period, the capacity allocation decisions among office and virtual appointments are made under the uncertainty of patients' health states. The physicians have a belief about patients' health statuses at the beginning of the first period. Over time due to the disease progress, patients' health statuses can change, and patients' health statuses are realized. In the second period, patients are scheduled for the office and virtual appointments which are determined in the first period. We consider that once the patients' scheduling decisions are made, patients are treated, and the disease progression occurs after the second period. The capacity assignment decisions are made without knowing the patient's health status realization, whereas patients' scheduling decisions are made based on the realized information. The stochastic characterizations in our framework are the changes in the patients' health status and the changes in the disease progression after the treatment. The stochastic parameters in our model are exogenous. We use a discrete distribution of these random parameters and denote each possible realization $\psi \in \Psi$ as a scenario with a corresponding probability z^{ψ}. Using this framework, we develop a two-stage stochastic programming model for capacity planning and patients' scheduling problem. Representation of the described general decision process is depicted in Figure 5.1.

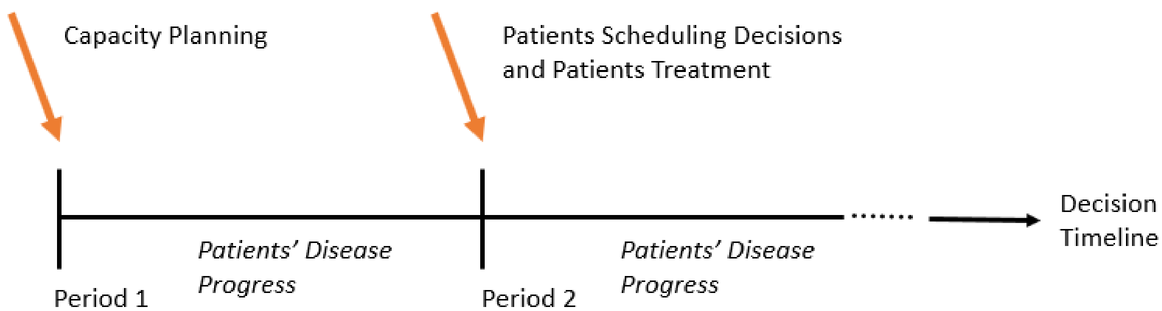

Figure 5.1: General decision process for capacity planning and patients' schedule

For a general mathematical representation of this problem, let $\mathcal{L}$ represent the set of patients that need chronic care and $i \in \mathcal{L}$ represents patient i in the set of patients $\mathcal{L}$. We consider two health status $j \in \{0,1\}$ where 0 represents the controlled health state and 1 represents the uncontrolled health state in chronic disease. The physician has a prior belief about patients' health status, which is described using the probability of patient i being in the health status at period t in scenario ψ (π_{it}^{ψ}) before appointments are scheduled. Then, we denote the health information vector, $\boldsymbol{\pi}_{it}^{\psi} = (\pi_{it}^{\psi}, 1 - \pi_{it}^{\psi})$, to reflect the health status of patient i at period t in scenario ψ. Patients' health status is stochastic and is realized after the first period. A patient can be in the controlled health state with probability π_{it}^{ψ} and in the uncontrolled health state with probability $(1-\pi_{it}^{\psi})$. To reflect the realization of a patient's true health state $j \in \{0,1\}$, we define a vector $\boldsymbol{e}_{it}^{\psi}$. $\boldsymbol{e}_{it}^{\psi} = (1,0)$ if patient i is in controlled heath state at period t in scenario ψ and $\boldsymbol{e}_{it}^{\psi} = (0,1)$ if patient i is in uncontrolled health state at period t in scenario ψ.

To differentiate office and virtual appointments, we define $k \in \mathcal{K} = \{o, v\}$, where "$o$" corresponds to office appointments and "v" corresponds to virtual appointments. To reflect the treatment process, we define matrices, $\boldsymbol{S}_o$ and $\boldsymbol{S}_v$, to represent the effect of treatment for office and virtual appointments, respectively, as follows:

$$\boldsymbol{S}_o = \begin{bmatrix} 1 & 0 \\ s_o & 1-s_o \end{bmatrix} \qquad \boldsymbol{S}_v = \begin{bmatrix} 1 & 0 \\ s_v & 1-s_v \end{bmatrix}$$

For both types of appointments, we assume that if a patient is diagnosed in the controlled heath state at period t, she/he will remain in the controlled heath state after the treatment with probability 1. If a patient is diagnosed in the uncontrolled heath state at period t, she/he will be in the controlled heath state with probability s_o after the treatment in the office appointments. This probability is s_v if the patient is scheduled for a virtual appointment. Since office appointments are expected to be more effective than virtual appointments in treatment, we assume $s_o \geq s_v$.

In our model, we further incorporate the disease progression. Due to the disease progression, we assume that a patient in the controlled health state at period t will remain in the controlled heath state at the beginning of period $t+1$ with probability w. We assume that there is no natural improvement in a patient's status if s/he is in the uncontrolled health state, thus, a patient in the uncontrolled health state at period t will remain in the uncontrolled health state at the beginning of period $t+1$ with probability 1. Hence, the disease progression matrix $\boldsymbol{W}$ can be defined as follow:

$$\boldsymbol{W} = \begin{bmatrix} w & 1-w \\ 0 & 1 \end{bmatrix}$$

After the definition of the treatment process and the disease progression, we can update the patients' health information vector by linking period t and period $t+1$. If a patient is not scheduled for office or virtual appointments in period t, the patients' health information vector at the beginning of period $t+1$ is updated by multiplying the health information vector at period t with the disease progression matrix (i.e., $\boldsymbol{\pi}_{i(t+1)}^{\psi} = \boldsymbol{\pi}_{it}^{\psi}\boldsymbol{W}$). If a patient is scheduled for a virtual appointment at period t, the patient's health status will be realized, s/he will receive the treatment, and a disease progression will occur until the next appointment. Hence, the patients' health information vector at the beginning of period $t+1$ will be $\boldsymbol{\pi}_{i(t+1)}^{\psi} = \boldsymbol{e}_{it}^{\psi}\boldsymbol{S}_v\boldsymbol{W}$ after virtual appointments. This health information vector is $\boldsymbol{\pi}_{i(t+1)}^{\psi} = \boldsymbol{e}_{it}^{\psi}\boldsymbol{S}_o\boldsymbol{W}$ for office appointments case. To reflect the cost that clinic assign the appointments, we denote c_o and c_v as the fixed cost of unit time of office and virtual appointments, respectively.

To model the capacity planning and the patients' schedule, the number of appointments planned for an appointment type $k \in \mathcal{K}$ is denoted by M_k, which is the decision variable in the first stage. We define $x_{it}^{k\psi} = 1$ if patient $i \in \mathcal{L}$ is scheduled for an appointment type $k \in \mathcal{K}$ at period $t \in T$ in scenario $\psi \in \Psi$, and $x_{it}^{k\psi} = 0$ otherwise. We note that $x_{it}^{k\psi}$ denotes patient scheduling decisions, and it is made in the second stage. We also define $\boldsymbol{b} = [1,0]$ as the QALY score vector corresponding to the health information vector. Then, the two-stage stochastic programming model can be written as follows:

$$Objective = max\, \mathbb{E}_{\psi} \sum_{t\in T} \sum_{i\in\mathcal{L}} \boldsymbol{\pi}_{it}^{\psi}\boldsymbol{b} \tag{5.1}$$

Subject to:

$$\sum_{k\in\mathcal{K}} x_{it}^{k\psi} \leq 1 \qquad \forall i \in \mathcal{L}, \forall t \in \mathcal{T}, \forall \psi \in \Psi \tag{5.2}$$

$$\sum_{i\in\mathcal{L}} x_{it}^{k\psi} \leq M_k \qquad \forall k \in \mathcal{K}, \forall t \in \mathcal{T}, \forall \psi \in \Psi \tag{5.3}$$

$$\boldsymbol{\pi}_{i,t+1}^{\psi} = \sum_{k\in\mathcal{K}} x_{it}^{k\psi}\boldsymbol{e}_{it}^{\psi}\boldsymbol{S}_k\boldsymbol{W} + \left(1 - x_{it}^{o\psi} - x_{it}^{v\psi}\right)\boldsymbol{\pi}_{it}^{\psi}W \qquad \forall i \in \mathcal{L}, \forall t \in \mathcal{T}, \forall \psi \in \Psi \tag{5.4}$$

$$c_oM_o + c_vM_v \leq B \tag{5.5}$$

$$M_o + M_v \leq |\mathcal{L}| \tag{5.6}$$

$$x_{it}^{v\psi}, x_{it}^{o\psi} \in \{0,1\} \qquad \forall i \in \mathcal{L}, \forall t \in \mathcal{T}, \forall \psi \in \Psi \tag{5.7}$$

$$M_o, M_v \geq 0 \tag{5.8}$$

Function (5.1) is the objective function, which is to maximize the expected total QALY score of all patients. Constraint (5.2) ensures that a patient can be scheduled for at most one type of appointment in any period. Constraint (5.3) states that the sum of patients scheduled for an appointment type $k \in \mathcal{K}$ is less than the total capacity planned for that type of appointment. Constraint (5.4) is the health information vector transition equation that is used to update the health information vector between periods. Constraint (5.5) ensures that the total cost is limited by B. B is a constant parameter representing the budget and resources for capacity allocation decisions. Constraint (5.6) ensures that the total number of appointments allocated in one period is less than the number of patients who need the care. Constraint (5.7) ensures the decision variables are binary. Constraint (5.8) defines that the total capacity allocated is non-negative for both types of appointments.

5.3 Numerical Studies

In this section, we numerically analyze a clinic that provides chronic care through office and virtual appointments. We conduct numerical experiments to investigate the optimal capacity planning and patients scheduling decisions with respect to a change on different parameter values. Since patient scheduling decisions are the second stage decisions that vary for each scenario, we present our result for the first stage variables (i.e., the capacity allocation decisions) which are the same for all scenarios.

We consider 20 patients within the clinic network that $|\mathcal{L}| = 20$ and 4096 scenarios in the two-stage stochastic programming model. We randomly assign their health information vector at the beginning of the first period, $\boldsymbol{\pi}_{i1}^{\psi}$, as shown in Table 5.1, by ensuring that the probability of being in the controlled health state is uniformly distributed between 0 and 1 for each patient. The fixed capacity cost of office appointments is estimated as $c_o = \$84.6/\text{day}$, and that of virtual appointments is estimated as $c_v = \$57.53/\text{day}$ (Lee 2009, A. H. Association 2016). In terms of the treatment matrix, the effectiveness of treatment of office and virtual appointments are initially assigned as $s_o = 0.95$ and $s_v = 0.7$, respectively. We note that we perform sensitivity analysis by changing the values of these parameters to see the impact of treatment effectiveness on capacity allocation decisions. Similarly, the probability that is used to describe the disease progression is initially assumed as $w = 0.8$. Since there is no reference for these parameters, we perform the sensitivity analysis in this section to investigate the effect of these parameters.

5.3.1 Impact of Budget on Capacity Allocation

We study the impact of the budget on the optimal capacity planning of office and virtual appointments. Capacity allocation decision among office and virtual appointments is challenging since office and virtual appointments have different treatment effectiveness and unit costs. More specifically, office appointments have higher treatment effectiveness than virtual appointments,

Table 5.1: Initial health information vector of patients $i \in \mathcal{L}$

Patient $i \in \mathcal{L}$	$\boldsymbol{\pi_{i1}}$
1	[0.82, 0.18]
2	[0.74, 0.26]
3	[0.56, 0.44]
4	[0.35, 0.65]
5	[0.24, 0.76]
6	[0.88, 0.12]
7	[0.78, 0.22]
8	[0.63, 0.37]
9	[0.45, 0.55]
10	[0.79, 0.21]
11	[0.62, 0.38]
12	[0.84, 0.16]
13	[0.26, 0.74]
14	[0.45, 0.55]
15	[0.74, 0.26]
16	[0.18, 0.82]
17	[0.35, 0.65]
18	[0.27, 0.73]
19	[0.69, 0.31]
20	[0.86, 0.14]

while virtual appointments are less costly than office appointments. Hence, the clinic should balance those points by ensuring that the overall health statuses of patients are maximized with limited budget and resources. To make the budget variation easy to present, we first define the maximum budget that the clinic needs for the appointment assignment, B^{max}. Since there are $|\mathcal{L}|$ patients in total, there can be at most $|\mathcal{L}|$ patients who can be scheduled for an appointment at one period. Since the cost of office appointments is greater than that of virtual appointments, we define the maximum budget as follows: $B^{max} = |\mathcal{L}| \times c_o = \$1692/\text{day}$. Then, we consider that $B = \alpha B^{max}$ and we vary the budget coefficient, α, from 0 to 1 to show the changes in budget and its impact on the capacity allocation decisions in Figure 5.2.

In Figure 5.2, the solid and dashed lines represent the changes in office and virtual appointment capacities, respectively. The dotted line represents the change in the value of the objective function. As observed, the objective is a non-decreasing function with respect to the budget, and the marginal impact of the unit budget increase on the objective function is decreasing

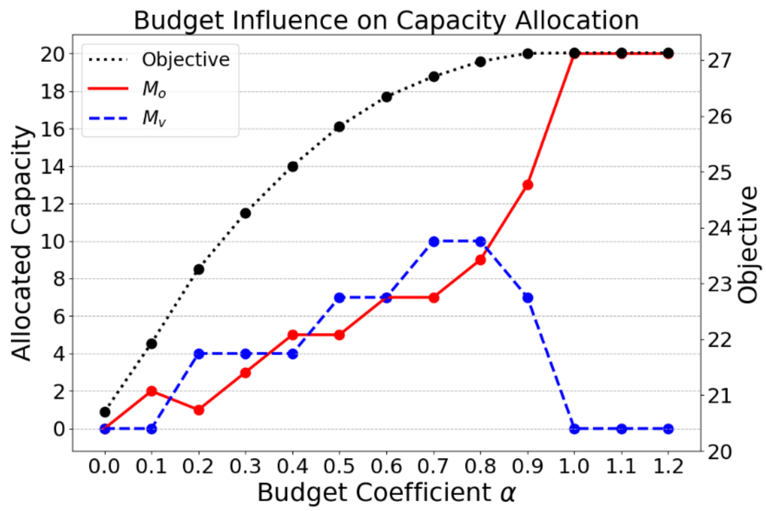

Figure 5.2: Budget influence on capacity allocation, $s_o = 0.95$, $s_v = 0.7$

until $\alpha \geq 1$.We allow α greater than 1 to investigate the effect of excessive budget. It is shown that the excessive budget does not improve the objective function and does not change the capacity allocation, which means that the health statuses of patients are not improved. Also, the number of virtual appointments increases from 0 to 10 when α increases from 0 to 0.8 and decreases from 10 to 0 when α increases from 0.8 to 1. we assume that the treatment effectiveness of office appointments is greater than that of virtual appointments. In this setting, $s_o = 0.95$ and $s_v = 0.7$. Due to the cost advantage of virtual appointments, when the budget is limited (i.e., $\alpha \leq 0.8$), more virtual appointments are assigned than the office appointments in most settings. However, as the budget increases (i.e. when the budget is not scarce), the cost advantage of virtual appointments is less significant in capacity allocation decisions. In this case, the impact treatment effectiveness is more significant in capacity allocation decisions, thus, office appointments start to replace virtual appointments as expected. As we increase s_v from 0.7 to 0.8 and 0.9, this phenomenon is clearer as shown in Figure 5.3 and Figure 5.4. However, no matter the value of s_v, $\alpha = 0.8$ is always a turning point that the clinic assigns 10 virtual appointments and 9 office appointments. When $\alpha > 0.8$, the number of office appointments starts to exceed that of virtual appointments.

5.3.2 Impact of Treatment Effectiveness on Capacity Allocation

Next, we study the impact of treatment effectiveness on the capacity allocation of office and virtual appointments. We define s_o/s_v as the relative treatment effectiveness. We set $s_o = 0.95$ and adjust the value of s_v based on the relative treatment effectiveness. The range of s_o/s_v is set

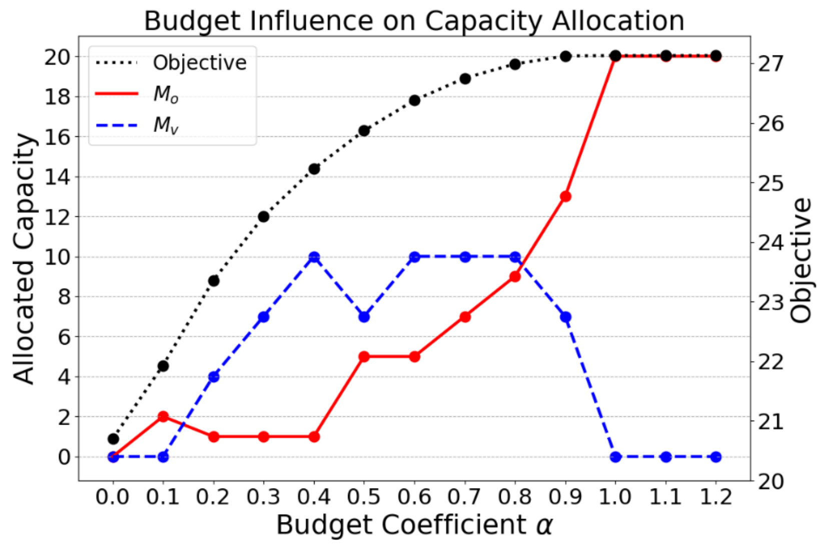

Figure 5.3: Budget influence on capacity allocation, $s_o = 0.95$, $s_v = 0.8$

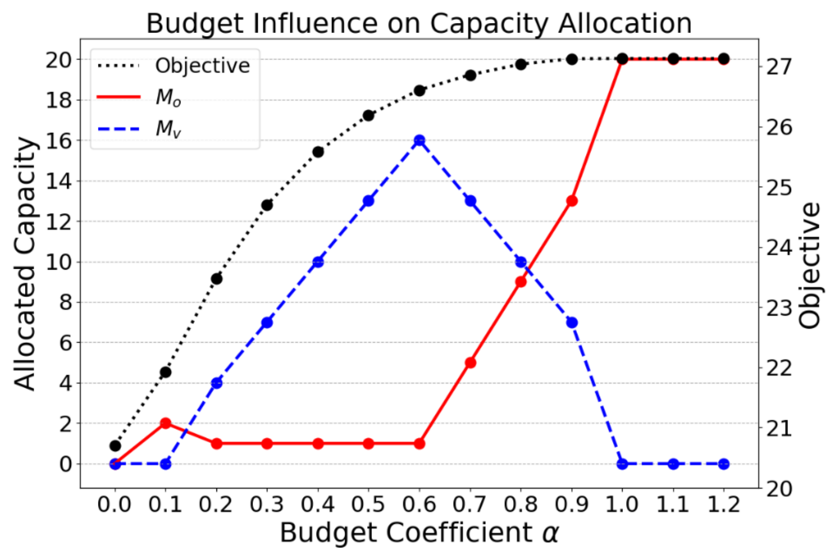

Figure 5.4: Budget influence on capacity allocation, $s_o = 0.95$, $s_v = 0.9$

from 1 to 1.5, which means that the treatment effectiveness of virtual appointments becomes worse as the ratio increases. As shown in Figure 5.5, we set the budget as $B = 0.4B^{max}$. The dashed and solid lines are the changes in the number of virtual and office appointments, respectively. When s_o/s_v equals to 1, the virtual appointments have the same treatment effectiveness as office appointments. Since virtual appointments are more cost-effective and they have the same treatment effectiveness with office appointments, all the budget is used to allocate the virtual appointments.

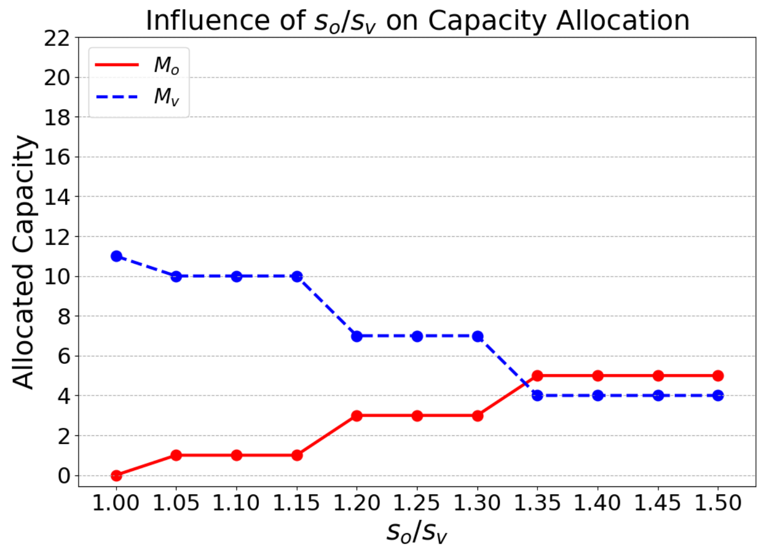

Figure 5.5: Influence of s_o/s_v on capacity allocation, budget coefficient $\alpha = 0.4$

However, as the ratio of s_o/s_v increases, the number of office appointments increases as well until $s_o/s_v = 1.35$.

We consider more settings by changing the budget where $\alpha = 0.6$ and 0.8 as shown in Figure 5.6 and Figure 5.7. When $\alpha = 0.4$, M_o and M_v become stable when s_o/s_v ratio reaches 1.35. On the other hand, M_o and M_v become stable when the s_o/s_v ratio is 1.25 and 1.05 for $\alpha = 0.6$ and 0.8, respectively. It shows that the impact of treatment effectiveness of virtual appointments is sensitive to the changes in budget. When the budget is scarce, even if the treatment effectiveness of virtual appointments is low, the cost advantage of virtual appointments is still a strength to the clinic. However, as the budget increases and as the clinic has sufficient budget the clinic replaces the number of virtual appointments with office appointments.

In Figure 5.7, it shows that even when s_o/s_v equals to 1.5, number of virtual appointments is still more than that of office appointments that $M_o = 9$ and $M_v = 10$. We find that under the budget constraint that $B = 0.8B^{max}$, if we add one more office appointment, we need to decreases the number of virtual appointments by 2 where $M_o = 10$ and $M_v = 8$. The total number of all appointments decreases by one in this case.

Table 5.2 and Figure 5.8 show the initial probability and the distribution of the number of patients in uncontrolled condition, which works for all the scenarios. We can see that there are on-average 9.27 patients in the uncontrolled health state and the number of uncontrolled patients is less and equal to 9 with the probability of 53.58%. Also, if a patient is in a controlled health state,

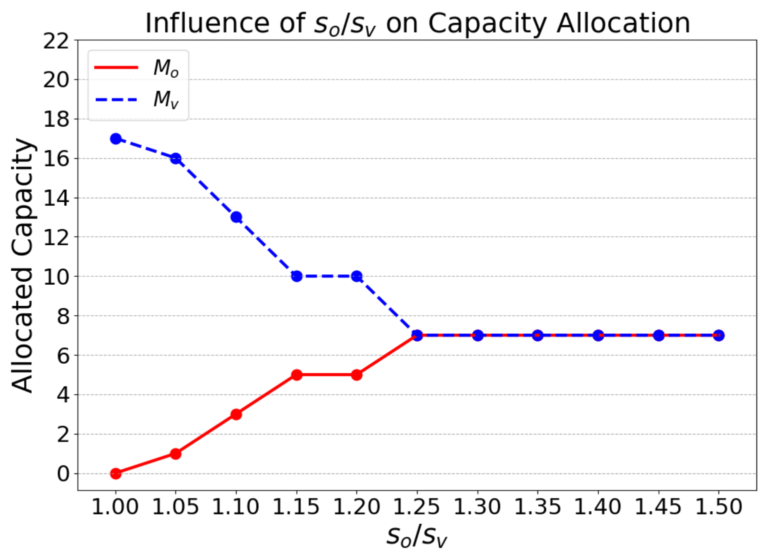

Figure 5.6: Influence of s_o/s_v on capacity allocation, budget coefficient $\alpha = 0.6$

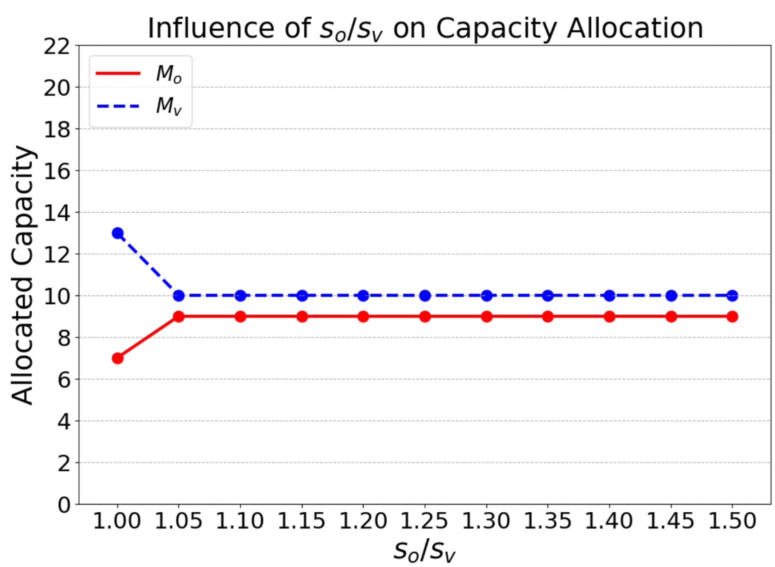

Figure 5.7: Influence of s_o/s_v on capacity allocation, budget coefficient $\alpha = 0.8$

the office and virtual appointments provide the same treatment service that ensures the patient remains in the controlled health state. Hence, in this situation, when the virtual appointments have relatively worse treatment effectiveness, its job is to serve the patients in the controlled patients and maintain their controlled condition while office appointments are responsible for serving the patients in uncontrolled condition, providing effective service, and improving their health condition. At this logic, the clinic can take greatly the advantages of both office and virtual appointments when the budget is limited for supporting enough office appointments.

Table 5.2: Distribution of number of patients in uncontrolled condition

# of patients in uncontrolled condition	Probability
0	0.000289
1	0.000928
2	0.003992
3	0.010306
4	0.022866
5	0.046139
6	0.070370
7	0.108204
8	0.126824
9	0.145863
10	0.137669
11	0.117590
12	0.089937
13	0.057291
14	0.034512
15	0.016475
16	0.007281
17	0.002568
18	0.000713
19	0.000167
20	0.000018
Avg	9.27

5.4 Conclusion

In this study, we use a two-stage stochastic programming model to formulate a capacity planning and patients scheduling problem. We consider the stochastic process of patients' disease progression and based on this fact, we develop the model to allocate the office and virtual appointments along with the patient scheduling decisions to maximize the patients' health conditions which are reflected by the probability that the physicians' belief that the patients stay in the controlled health condition. From the numerical studies, we find that even though the treatment effectiveness of virtual appointments is worse than that of office appointments, virtual appointments are still valuable due to its cost advantage, especially when the budget is limited for providing enough office appointments. However, the capacity allocation decision is more sensitive

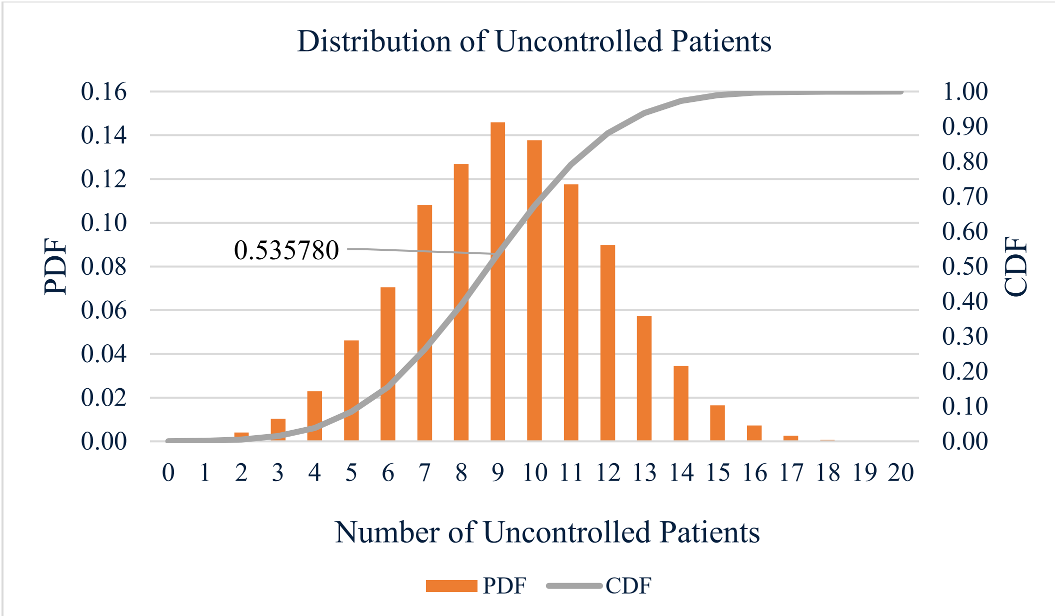

Figure 5.8: PDF and CDF function of the distribution of uncontrolled patients

to the relative treatment effectiveness of the interventions as the budget increases. We find that virtual appointments should be used more often to schedule controlled patients compared to treating uncontrolled patients. When virtual appointments and office appointments have the same treatment effectiveness, virtual appointments can be considered for all appointments since they are more cost-effective than the office appointments.

Chapter 6: Conclusion and Discussion

In the thesis, we study the integration of virtual appointments with traditional office appointments in a chronic care setting. Nowadays, virtual appointments play a significant role in the management of chronic conditions since virtual appointments are more convenient and cost-effective than office appointments. Deriving decision rules to combine virtual appointments with office appointments in an appropriate way will help healthcare providers and also provide patients efficient and affordable care in chronic care management.

To address the management issues in the integration of virtual appointments with the office appointments, we build several mathematical models in the thesis. In Chapter 3, we develop a migration network model to simulate the clinic system with both office and virtual appointments. We build a newsvendor-type optimization model to determine the capacity of the office and virtual appointments that maximize the average long-term profit of the clinic. We perform numerical experiments by using the data sources from the literature and open data sources. We show that as the follow-up rate increases, the required capacities of both office and virtual appointments increase which also results in an increase in the clinic's overall profit. We propose easy-to-implement policies for healthcare providers and we compare those policies for varying parameter values. We further find that the fixed-ratio capacity allocation policy performs worse than the dynamic optimal capacity allocation policy. Hence, our capacity allocation model is helpful for clinics when making capacity allocation decisions.

In Chapter 4, we use the migration network model that we develop in Chapter 3 and we investigate the optimal follow-up rate for office and virtual appointments when the capacity is fixed. We develop linear and nonlinear programming models to help clinics in the determination of the optimal follow-up rate for office and virtual appointments that maximizes the usage of resources and maximizes the clinics' profit. Our numerical experiments show that the number of patients that the clinic can serve is limited and as the new patients' arrival rate increases, the follow-up rate for existing patients needs to be decreased for a given capacity. The clinics should balance the number of existing patients and the new patients to keep the panel size in an appropriate range.

In Chapter 5, we consider capacity allocation and patient scheduling decisions simultaneously under uncertainty. More specifically, we develop a two-stage stochastic programming model to investigate capacity allocation decisions along with the patient scheduling decisions that maximize patients' overall health statuses. We consider that patients' health states are uncertain and this information is realized over time. Different from previous chapters, we also consider maximizing the overall health statuses of patients for a given budget. Our results show that due to the cost-effectiveness of virtual appointments, it is better to allocate more virtual appointments when the budget is limited. However, when the resources are not scarce, the difference in the treatment effectiveness of interventions becomes more important and the number of allocated virtual appointments decreases. Our results provide managerial insights for clinics in allocating capacity for varying parameter values. We find that virtual appointments can substitute for office appointments, which opens office appointment slots for other patients in need.

There are some limitations of this study that can be extended in several directions. First, the investigation of the new patients' arrival rate for virtual and office appointments can be another direction to study. Considering the arrival rate as a decision variable refers to the physician's panel size decision, where the physician can decide on the rate of new patients to accept in her/his panel (Green 2008, Ozen 2013, N. a. Liu 2014). Second, we assume the population of the patients is finite, which is a common but unrealistic assumption in the migration network literature. Our next step will be to consider an infinite number of patients that a clinic system can serve. Third, our data are based on the literature, which limits the generalization of the model. In the future, the model can be verified by using a specific clinic's or hospital's data. Finally, we consider that patients return to home from the office and virtual appointments and define one system state for this case. In a future study, the migration network model can be extended, and the system states can be divided to differentiate the patients who are back from office and virtual appointments.

Despite these limitations and directions of extension and improvement, we believe the mathematical models in the thesis will bring value to the research topic focusing on the integration of office and virtual appointments and will ultimately help more policymakers to apply the virtual appointments in practice. Our models will further provide insights and easy-to-implement policy rules in capacity allocation, follow-up rate determination and patient scheduling decisions for the office and virtual appointments in a chronic care setting.

References

Ackerman, Michael J and Filart, Rosemarie and Burgess, Lawrence P and Lee, Insup and Poropatich, Ronald K. "Developing next-generation telehealth tools and technologies: patients, systems, and data perspectives." Telemedicine and e-Health (Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA) 16 (2010): 93-95.

Hospital American Association. "Telehealth: Helping Hospitals Deliver Cost-Effective Care." Available at: https://www.aha.org/system/files/content/16/16telehealthissuebrief.pdf, 2016.

Telemedicine American Association. What is telemedicine & Telehealth? 2020. https://s3.amazonaws.com/rdcms-himss/files/production/public/HIMSSorg/Content/files/Line \%2016\%20-\%20What\%20Is\%20Telemedicine.pdf.

Bashshur, Rashid L. "On the definition and evaluation of telemedicine." Telemedicine Journal 1 (1995): 19-30.

Batun, Sakine and Begen, Mehmet A. "Optimization in healthcare delivery modeling: Methods and applications." Handbook of Healthcare Operations Management, 2013: 75-119.

Bavafa, Hessam and Savin, Sergei and Terwiesch, Christian. "The impact of e-Visits on visit frequencies and patient health: Evidence from primary care." Management Science (INFORMS) 64 (2018): 5461-5480.

Bavafa, Hessam and Hitt, Lorin M and Terwiesch, Christian. "Redesigning primary care delivery: Customized office revisit intervals and e-visits." Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2363685, 2019.

Bayram, Armagan and Deo, Sarang and Iravani, Seyed and Smilowitz, Karen. "Managing virtual appointments in chronic care." IISE Transactions on Healthcare Systems Engineering, 2019: 1-17.

Bedi, Baljit Singh. "Telemedicine standards: need and Indian initiatives." Telemedicine and e-Health (Mary Ann Liebert, Inc. 140 Huguenot Street 3rd Floor New Rochelle, NY 10801 USA) 15 (2009): 597-599.

Block, Henry W and Langberg, Naftali A and Savits, Thomas H. "Repair replacement policies." Journal of Applied probability (Cambridge University Press) 30 (1993): 194-206.

Bodenheimer, Thomas and Wagner, Edward H and Grumbach, Kevin. "Improving primary care for patients with chronic illness." Jama 288 (2002): 1775-1779.

Bretthauer, Kurt M and Cote, Murray J. "A model for planning resource requirements in health care organizations." Decision Sciences (Wiley Online Library) 29 (1998): 243-270.

The United States Census Bureau. "Population and Housing Unit Estimates." Available at: https://www.census.gov/programs-surveys/popest.html?intcmp=serp, 2018.

Caceres, Cesar and Gomez, Enrique J and Garcia, Felipe and Gatell, Jose M and del Pozo, Francisco. "An integral care telemedicine system for HIV/AIDS patients." International Journal of Medical Informatics 75 (2006): 638-642.

Pinnacle Care. "The Human Cost and Financial Impact of Misdiagnosis." Available at: https://www.pinnaclecare.com/forms/download/Human-Cost-Financial-Impact-Whitepaper.pdf, 2016.

Clinics.com. "Clinics in Michigan. 2019" available at: https://www.freeclinics.com/sta/michigan.

Craig, John and Chua, Richard and Russell, Clive and Patterson, Victor and Wootton, Richard. "The cost-effectiveness of teleneurology consultations for patients admitted to hospitals without neurologists on site. 1: A retrospective comparison of the case-mix and management at two rural hospitals." Journal of telemedicine and telecare (SAGE Publications Sage UK: London, England) 6 (2000): 46-49.

Dall, Timothy M and Roary, Mary and Yang, Wenya and Zhang, Shiping and Zhang, Yiduo and Arday, David R and Gantt, Cynthia J and Chen, Yaozhu J. "Peer Reviewed: Health Care Use and Costs for Participants in a Diabetes Disease Management Program, United States, 2007-2008." Preventing Chronic Disease (Centers for Disease Control and Prevention) 8 (2011).

Deo, Sarang and Iravani, Seyed and Jiang, Tingting and Smilowitz, Karen and Samuelson, Stephen. "Improving health outcomes through better capacity allocation in a community-based chronic care model." Operations Research (INFORMS) 61 (2013): 1277-1294.

Partnership to Fight Chronic Disease. "The Growing Crisis of Chronic Disease in the United States." Available at: https://www.fightchronicdisease.org/sites/default/files/docs/GrowingCrisisofChronicDiseaseintheUSfactsheet81009.pdf, 2017.

The Physicians Foundation. "2018 Survey of American Physicians: Practice Patterns and Perspectives." Available at: https://physiciansfoundation.org/research-insights/the-physicians-foundation-2018-physician-survey/, 2018.

Fox, Bennett. "Discrete optimization via marginal analysis." Management Science (INFORMS) 13 (1966): 210-216.

Goodarzi, Mandana and Ebrahimzadeh, Issa and Rabi, Alireza and Saedipoor, Bahman and Jafarabadi, Mohammad Asghari. "Impact of distance education via mobile phone text messaging on knowledge, attitude, practice and self efficacy of patients with type 2 diabetes mellitus in Iran." Journal of Diabetes & Metabolic Disorders (BioMed Central) 11 (2012): 10.

Green, Linda V and Savin, Sergei. "Reducing delays for medical appointments: A queueing approach." Operations Research (INFORMS) 56 (2008): 1526-1538.

Gupta, Diwakar and Denton, Brian. "Appointment scheduling in health care: Challenges and opportunities." IIE Transactions 40 (2008): 800-819.

Heffler, Stephen and Smith, Sheila and Won, Greg and Clemens, M Kent and Keehan, Sean and Zezza, Mark. "Health spending projections for 2001--2011: the latest outlook." Health Affairs (Project HOPE-The People-to-People Health Foundation, Inc.) 21 (2002): 207-218.

Hulshof, Peter JH and Kortbeek, Nikky and Boucherie, Richard J and Hans, Erwin W and Bakker, Piet JM. "Taxonomic classification of planning decisions in health care: a structured review of the state of the art in {OR}/{MS}." Health Systems 1 (2012): 129-175.

Kelly, Frank P. Reversibility and stochastic networks. Cambridge University Press, 1979.

Kochanek, Kenneth D and Miniño, Arialdi M and Murphy, Sherry L and Xu, Jiaquan and Kung, Hsiang-Ching. "Deaths: final data for 2009." National Vital Statistics Reports, 2011.

Kucukyazici, Beste and Verter, Vedat and Mayo, Nancy E. "An analytical framework for designing community-based care for chronic diseases." Production and Operations Management 20 (2011): 474-488.

Kucukyazici, Beste and Verter, Vedat. "Managing community-based care for chronic diseases: The quantitative approach." Operations Research and Health Care Policy, 2013: 71-90.

LaGanga, L. R. and Lawrence, S. R. "Appointment overbooking in health care clinics to improve patient service and clinic performance." Production and Operations Management 21 (2012): 874-888.

Lee, Donald KK and Zenios, Stefanos A. "Optimal capacity overbooking for the regular treatment of chronic conditions." Operations research (INFORMS) 57 (2009): 852-865.

Leggett, Paul and Graham, Lorraine and Steele, Keith and Gilliland, Andrew and Stevenson, Michael and O'Reilly, Dermot and Wootton, Richard and Taggart, Alister. "Telerheumatology--diagnostic accuracy and acceptability to patient, specialist, and general practitioner." Br J Gen Pract (British Journal of General Practice) 51 (2001): 746-748.

Li, Yan and Zhang, Yi and Kong, Nan and Lawley, Mark. "Capacity planning for long-term care networks." IIE Transactions (Taylor & Francis) 48 (2016): 1098-1111.

Liu, Xiang and Hu, Michael and Helm, Jonathan E and Lavieri, Mariel S and Skolarus, Ted A."Missed Opportunities in Preventing Hospital Readmissions: Redesigning Post-Discharge Checkup Policies." Production and Operations Management (Wiley Online Library) 27 (2018): 2226-2250.

Liu, Nana and Ziya, Serhan. "Panel size and overbooking decisions for appointment-based services under patient no-shows." Production and Operations Management (Wiley Online Library) 23 (2014): 2209-2223.

Marcin, James P. "Telemedicine in the pediatric intensive care unit." Pediatric Clinics (Elsevier) 60 (2013): 581-592.

McKinstry, Brian and Hammersley, V and Burton, Christopher and Pinnock, Hilary and Elton, Rob and Dowell, J and Sawdon, N and Heaney, D and Elwyn, Glyn and Sheikh, A. "The quality, safety and content of telephone and face-to-face consultations: a comparative study." BMJ Quality & Safety (BMJ Publishing Group Ltd) 19 (2010): 298-303.

Ozen, Asli and Balasubramanian, Hari. "The impact of case mix on timely access to appointments in a primary care group practice." Health care management science (Springer) 16 (2013): 101-118.

Perednia, Douglas A and Allen, Ace. "Telemedicine technology and clinical applications." Jama 273 (1995): 483-488.

Pham, Hoang and Wang, Hongzhou. "Optimal (tau, T) opportunistic maintenance of ak-out-of-n: G system with imperfect PM and partial failure." Naval Research Logistics (NRL) (Wiley Online Library) 47 (2000): 223-239.

Center for Disease Control and Prevention. "Behavioral risk factor surveillance system survey (BRFSS), Atlanta, Georgia: US Department of Health and Human Services, Centers for Disease Control and Prevention." 2010.

Center for Disease Control and Prevention. "The Power of Prevention." Available at: https://www.cdc.gov/chronicdisease/pdf/2009-Power-of-Prevention.pdf, 2009.

Russo, Jack E and McCool, Ryan R and Davies, Louise. "VA telemedicine: an analysis of cost and time savings." Telemedicine and e-Health (Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA) 22 (2016): 209-215.

Department of Health and Human Services. "Chronic Care Management Services." ICN 909188 December, (2016).

Valero, MA and Arredondo, MT and Del Nogal, F and Rodriguez, JM and Frias, E. "Patient satisfaction with a home televisiting service based on interactive television over a cable network." Journal of telemedicine and telecare (SAGE Publications Sage UK: London, England) 6 (2000): 99-101.

Wu, Shin-Yi and Green, Anthony. "Global Telemedicine Market Outlook 2022." Research and Markets, (2018).

Wu, Shin-Yi and Green, Anthony. "Projection of chronic illness prevalence and cost inflation." Santa Monica, CA: RAND Health 18 (2000).

Appendices

Appendix A: Summary of the Notations

Notation	Definition
Migration Network Model	
$\mu_i, i \in \{o, v\}$	The service rate of office and virtual appointments, respectively
$\lambda_i, i \in \{o, v\}$	The arrival rate of new patients of office and virtual appointments, respectively
$\sigma_i, i \in \{o, v\}$	The frequency that a patient re-visit the physician of office and virtual appointments, respectively
δ	The departure rate of the patients in the system.
$\frac{1}{\gamma}$	The average time for a controlled patient to progress to uncontrolled state.
P_0	The probability that the patients diagnosed in a controlled condition stay in controlled condition after a virtual treatment.
P_1	The probability that the patients diagnosed in an uncontrolled condition improve in controlled condition after a virtual treatment.
$P_{j\|j'}, j, j' \in \{0,1\}$	The probability that the virtual diagnosis treats the patient in condition j as patient in condition j', where $j, j' = 0$ represents the controlled condition; $j, j' = 1$ represents the uncontrolled condition.
P_h	The probability that the new arrival patient in virtual appointment is in controlled condition.
h_0	The state represents controlled condition at home.
h_1	The state represents uncontrolled condition at home.
o	The state represents scheduled patients at the office appointment.
v	The state represents scheduled patients at the virtual appointment.
v_0	The state represents scheduled controlled patients at the virtual appointment.
v_1	The state represents scheduled uncontrolled patients at the virtual appointment.
$\alpha_k, k \in \{h_0, h_1, o, v_0, v_1\}$	The expected number of patients at node k in the steady-state condition.
$x_k, k \in \{h_0, h_1, o, v_0, v_1\}$	The number of patients at node k in the steady-state condition.
$\pi_k(x_k), k \in \{h_0, h_1, o, v_0, v_1\}$	The probability that x_k patients at node k in the steady state.

Notation	Definition
Capacity Allocation Model	
$r_k, k \in \{o, v_0, v_1\}$	The marginal profit for each patient treatment at node k.
$c_k, k \in \{o, v_0, v_1\}$	The fixed cost of each unit of capacity at node k.
$f_k, k \in \{o, v_0, v_1\}$	The penalty cost for each overflow patient at node k.
$x_k(t), k \in \{o, v_0, v_1\}$	The current number of patients in the node k at time t.
e_{id}	The penalty cost on each imperfect diagnosis.
Capacity Allocation Model	
M_o	The capacity of office appointments.
M_v	The capacity of virtual appointments.
M_{v_0}	The capacity of virtual appointments for controlled patients.
M_{v_1}	The capacity of virtual appointments for uncontrolled patients.
$\boldsymbol{M} = (M_o, M_{v_0}, M_{v_1})$	The vector to save the decision variables (capacity of each node).
$A(M)$	The total long-run average earnings of the clinic system.
$A_k(M_k), k \in \{o, v_0, v_1\}$	The long-run average earnings from the node k.
TC	The limited total capacity added to the base model.
T_w	The limited total working time per day.
$M_k^{min}, k \in \{o, v_0, v_1\}$	The optimal capacity that maximize the average long-run earnings at the node k.
$M_k^t, k \in \{o, v_0, v_1\}$	The variable used to save the value of the capacity of node k in the t^{th} iteration.
$A'(M_k), k \in \{o, v_0, v_1\}$	The partial differential of objective function to M_k.
$Z_k(M_k), k \in \{o, v_0, v_1\}$	The marginal profit of node k under unit of time.
M^*	The vector to save the optimal value of the decision variables (optimal capacity of each node) for basic capacity allocation model.
M_{TC}^*	The vector to save the optimal value of the decision variables (optimal capacity of each node) under limited total capacity, TC.
$M_{T_w}^*$	The vector to save the optimal value of the decision variables (optimal capacity of each node) under limited total working time, T_w.
M_{T_w}	The vector to save the value of the decision variables (capacity of each node) under limited total working time, T_w from Algorithm 2.
W_{id}	The average long-run penalty cost on imperfect diagnosis.

Notation	Definition
Stochastic Programming	
$t \in \{1,2\}$	The state period t in the problem framework.
$\psi \in \Psi$	The specific scenario ψ in the problem.
$z^{\psi}, \psi \in \Psi$	The probability that scenario ψ happens.
$\mathcal{L}$	The set of patients that need chronic care.
$\pi_{it}, i \in \mathcal{L}, t \in \{1,2\}$	Probability of patient i being in the health status at period t.
$\boldsymbol{\pi}_{it} = (\pi_{it}, 1 - \pi_{it})$	The health information vector of patient i at period t.
$\boldsymbol{e}_{it}$	The diagnosis information vector of patient i at period t.
$s_k, k \in \{o, v\}$	The treatment effectiveness of appointment with appointments' type k for uncontrolled patients.
$\boldsymbol{S}_k, k \in \{o, v\}$	The treatment matrix of appointment with appointments' type k.
w	The probability that patients in controlled status stay controlled.
$\boldsymbol{W}$	The disease progress matrix.
$x_{it}^{k\psi}$	$x_{it}^{k\psi} = 1$ if patient $i \in \mathcal{L}$ is scheduled for an appointment type $k \in \{o, v\}$ at period $t \in T$ in scenario $\psi \in \Psi$, and $x_{it}^{k\psi} = 0$ otherwise.
$\boldsymbol{b}$	QALY score vector corresponding to the health information vector.
B	The budget that the clinic plans for the appointments.
B^{max}	Maximum budget that needs for the appointment's assignment.
α	The budget coefficient, $B = \alpha B^{max}$.

Appendix B: Proofs

B.1. Proof of Average Number of Patients at Each Node

Recall that the model with imperfect diagnosis and treatment meets the definition of an open migration network (Kelly 1979), number of patients at each node satisfy the following traffic equations:

$$\mu_v \alpha_{v_0} - \sigma_v P_{0|0} \alpha_{h_0} - \sigma_v P_{0|1} \alpha_{h_1} = P_h \lambda_v \qquad (A.1)$$

$$\mu_v \alpha_{v_1} - \sigma_v \left(1 - P_{0|0}\right) \alpha_{h_0} - \sigma_v \left(1 - P_{0|1}\right) \alpha_{h_1} = (1 - P_h) \lambda_v \qquad (A.2)$$

$$-\mu_v P_0 \alpha_{v_0} - \mu_v P_1 \alpha_{v_1} + (\sigma_v + \sigma_o + \delta + \gamma) \alpha_{h_0} - \mu_o \alpha_o = 0 \qquad (A.3)$$

$$-\mu_v (1 - P_0) \alpha_{v_0} - \mu_v (1 - P_1) \alpha_{v_1} - \gamma \alpha_{h_0} + (\sigma_v + \sigma_o + \delta) \alpha_{h_1} = 0 \qquad (A.4)$$

$$-\sigma_o \alpha_{h_0} - \sigma_o \alpha_{h_1} + \mu_o \alpha_o = \lambda_o \qquad (A.5)$$

These equations represent that the inflow to node k must to be equal to outflow from node k. Equations (A.1) - (A.5) are five equations with five unknowns, then we can solve the traffic equations and obtain the average number of patients in each node at steady-state and get:

$$\alpha_{v_0} = \frac{\Phi_1 \lambda_v + \left[\left(P_{0|1} + P_{0|0} P_1 - P_{0|1} P_1\right) \sigma_v^2 + P_{0|0} \delta \sigma_v + P_{0|0} \sigma_o \sigma_v + P_{0|1} \gamma \sigma_v\right] \lambda_o}{\left(\delta \mu_v \left(\delta + \gamma + \sigma_o + \sigma_v - P_0 P_{0|0} \sigma_v + P_0 P_{0|1} \sigma_v + P_{0|0} P_1 \sigma_v - P_{0|1} P_1 \sigma_v\right)\right)} \qquad (A.6)$$

$$\alpha_{v_1} = \frac{\Phi_2 \lambda_v + \left[\Phi_3 + \left(1 - P_{0|1}\right) \gamma \sigma_v\right] \lambda_o}{\left(\delta \mu_v \left(\delta + \gamma + \sigma_o + \sigma_v - P_0 P_{0|0} \sigma_v + P_0 P_{0|1} \sigma_v + P_{0|0} P_1 \sigma_v - P_{0|1} P_1 \sigma_v\right)\right)} \qquad (A.7)$$

$$\alpha_{h_0} = \frac{\Phi_4 \lambda_v + \left[\sigma_o + \left(P_1 + P_{0|1} (P_0 - P_1)\right) \sigma_v + \delta\right] \lambda_o}{\left(\delta \left(\delta + \gamma + \sigma_o + \sigma_v - P_0 P_{0|0} \sigma_v + P_0 P_{0|1} \sigma_v + P_{0|0} P_1 \sigma_v - P_{0|1} P_1 \sigma_v\right)\right)} \qquad (A.8)$$

$$\alpha_{h_1} = \frac{\Phi_5 \lambda_v + \left[\gamma + \left(1 - P_1 - P_{0|0} (P_0 - P_1)\right) \sigma_v\right] \lambda_o}{\left(\delta \left(\delta + \gamma + \sigma_o + \sigma_v - P_0 P_{0|0} \sigma_v + P_0 P_{0|1} \sigma_v + P_{0|0} P_1 \sigma_v - P_{0|1} P_1 \sigma_v\right)\right)} \qquad (A.9)$$

$$\alpha_o = \frac{\sigma_o \lambda_v + (\delta + \sigma_o) \lambda_o}{\delta \mu_o} \qquad (A.10)$$

$$\alpha_v = \alpha_{v_0} + \alpha_{v_1} = \frac{\sigma_v \lambda_o + (\delta + \sigma_v) \lambda_v}{\delta \mu_v} \qquad (A.11)$$

where

$$\Phi_1 = \left(P_{0|1} + P_{0|0}P_1 - P_{0|1}P_1\right)\sigma_v^2 + \left(P_{0|1} + P_h + P_{0|0}P_1 - P_{0|1}P_1\right)\delta\sigma_v \quad (A.12)$$
$$+P_{0|0}\sigma_o\sigma_v + P_{0|1}\gamma\sigma_v + P_h\delta^2 + P_h\delta\gamma + P_h\delta\sigma_o$$

$$\Phi_2 = \left(1 - P_{0|1} - P_0P_{0|0} + P_0P_{0|1}\right)\sigma_v^2 + \left(2 - P_{0|1} - P_h - P_0P_{0|0} + P_0P_{0|1}\right)\delta\sigma_v \quad (A.13)$$
$$+\left(1 - P_{0|0}\right)\sigma_o\sigma_v + \left(1 - P_{0|1}\right)\gamma\sigma_v + (1 - P_h)\delta^2 + (1 - P_h)\delta\gamma + (1 - P_h)\delta\sigma_o$$

$$\Phi_3 = \left[1 - P_{0|1} + P_0\left(P_{0|1} - P_{0|0}\right)\right]\sigma_v^2 + \left(1 - P_{0|0}\right)\delta\sigma_v + \left(1 - P_{0|0}\right)\sigma_o\sigma_v \quad (A.14)$$

$$\Phi_4 = \sigma_o + \left(P_1 + P_{0|1}(P_0 - P_1)\right)\sigma_v + \left(P_1 + P_h(P_0 - P_1)\right)\delta \quad (A.15)$$

$$\Phi_5 = \gamma + \left(1 - P_1 - P_{0|0}(P_0 - P_1)\right)\sigma_v + \left(1 - P_1 - P_h(P_0 - P_1)\right)\delta \quad (A.16)$$

B.2. Proof of Reformulation of the Objective Function

The original objective function of the base capacity allocation model is:

$$A(M) = \lim_{T\to\infty}\frac{1}{T}\left\{\sum_{k\in\{o,v_0,v_1\}}\int_0^T r_k min\,[x_k(t), M_k]dt - \int_0^T c_kM_kdt - \int_0^T f_k(x_k(t) - M_k)^+dt\right\} \quad (A.17)$$

To simplify the objective function (A.17), we make some changes:

$$min[x_k(t), M_k] = x_k(t) - [x_k(t) - M_k]^+ \quad (A.18)$$

$$M_k = x_k(t) - [x_k(t) - M_k]^+ + [M_k - x_k(t)]^+ \quad (A.19)$$

Substitute equation (A.18) and (A.19) into objective function (A.17) gives:

$$A(M) = \frac{1}{T}\left\{\int_0^T (r_k - c_k)x_k(t)dt - \int_0^T c_k[M_k - x_k(t)]^+dt - \int_0^T (f_k + r_k - c_k)[x_k(t) - M_k]^+dt\right\} (A.20)$$

Due to ergodicity, let $\mathbb{E}_{\pi_k}(x_k), k \in \{o, v_o, v_1\}$ be the expected number of patients at node i under the steady state distribution $\pi_k, k \in \{o, v_o, v_1\}$. We have:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x_k(t)dt = \mathbb{E}_{\pi_k}(x_k) = \alpha_k \quad (A.21)$$

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T [M_k - x_k(t)]^+dt = \mathbb{E}_{\pi_k}(M_k - x_k)^+ \quad (A.22)$$

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T [x_k(t) - M_k]^+dt = \mathbb{E}_{\pi_k}(x_k - M_k)^+ \quad (A.23)$$

Take equation (A.21) - (A.23) into objective function (A.20), we get:

$$A(M) = \sum_{k\in\{o,v_o,v_1\}}\left[(r_k - c_k)\alpha_k - c_k\mathbb{E}_{\pi_k}(M_k - x_k)^+ - (f_k + r_k - c_k)\mathbb{E}_{\pi_k}(x_k - M_k)^+\right] \quad (A.24)$$

In the Function (A.24), the first term is the marginal profit. The second term is the opportunity cost for unutilized capacity. The last term represents the cost due to patient overflow.

B.3. Proof of Proposition 1

Since the office and virtual processes are independent to each other, to maximize the objective $A(M)$, it suffices to maximize the sub-objective $A_k(M_k), k \in \{o, v_0, v_1\}$ separately.

$$A_k(M_k) = (r_k - c_k)\alpha_k - c_k E_{\pi_k}(M_k - x_k)^+ - (f_k + r_k - c_k)E_{\pi_k}(x_k - M_k)^+ \quad \text{(A.25)}$$

The differential and the second-order differential of function (A.25) are

$$\frac{\Delta A_k(M_k)}{\Delta(M_k)} = A_k(M_k + 1) - A_k(M_k) \quad \text{(A.26)}$$

$$= -c_k\left[\sum_{x_k=0}^{M_k+1}(M_k + 1 - x_k)\pi_k(x_k) - \sum_{x_k=0}^{M_k}(M_k - x_k)\pi_k(x_k)\right]$$

$$-(f_k + r_k - c_k)\left[\sum_{x_k=M_k+1}^{\infty}\big(x_k - (M_k + 1)\big)\pi_k(x_k) - \sum_{x_k=M_k}^{\infty}(x_k - M_k)\pi_k(x_k)\right]$$

$$= -c_k\sum_{x_k=0}^{M_k}\pi_k(x_k) + (f_k + r_k - c_k)\sum_{x_k=M_k+1}^{\infty}\pi_k(x_k)$$

$$= -c_k\pi_k(x_k \le M_k) + (f_k + r_k - c_k)\pi_k(x_k > M_k)$$

$$= f_k + r_k - c_k - (f_k + r_k)\pi_k(x_k \le M_k)$$

$$\frac{\Delta^2 A_k(M_k)}{\Delta^2(M_k)} = \Delta A_k(M_k + 1) - \Delta A_k(M_k) \quad (A.27)$$

$$= -(f_k + r_k)[\pi_k(x_k \le M_k + 1) - \pi_k(x_k \le M_k)] < 0$$

It is clear that the objective function $A_k(M_k)$ is a discrete concave function. Hence, to maximize $A_k(M_k)$, the optimal capacity of node k is the smallest positive integer $M_k = M_k^{min}$ that makes $\Delta A_k(M_k) \le 0$, and we have

$$\forall M_k \in \left[0, M_k^{min}\right),\ \Delta A_k M_k > 0,\ then,\ A_k(M_k) < A_k\left(M_k^{min}\right) \quad (A.28)$$

$$\forall M_k \in \left[M_k^{min}, \infty\right),\ \Delta A_k M_k \le 0,\ then,\ A_k(M_k) \le A_k\left(M_k^{min}\right) \quad \text{(A.29)}$$

In one case that if $\Delta A_k\left(M_k^{min}\right) = 0$, the optimal capacity can be M_k^{min} or $\left(M_k^{min} + 1\right)$, since $A_k\left(M_k^{min}\right) = A_k\left(M_k^{min} + 1\right)$. But it has a small probability that $\Delta A_k\left(M_k^{min}\right) = 0$ and even in that case, M_k^{min} is one of the optimal solutions. Hence, we conclude that M_k^{min} is the optimal capacity for node k that maximize $A_k(M_k)$.

$$M_k^{min} = min\left\{M_k \geq 0: \pi_k(x_k \leq M_k) \geq \frac{f_k + r_k - c_k}{f_k + r_k}\right\} \quad (A.30)$$

From Function (A.30), we obtain $M^* = \left(M_o^{min}, M_{v_0}^{min}, M_{v_1}^{min}\right)$, which is the optimal capacity for the $A(M)$.

B.4. Proof of Proposition 2

When physicians do not have enough working time, the solution is obtained through the Algorithm 2. Assume through t^{th} iteration, we obtain the solution from the algorithm, and the solution is $M_{T_w} = M_{T_w}^t = \left(M_o^t, M_{v_0}^t, M_{v_1}^t\right)$. Hence, M_{T_w} satisfies:

$$\sum_{k\in\{o,v_o,v_1\}} \frac{1}{\mu_k} M_k^t \leq T_w \quad (A.31)$$

$$\sum_{k\in\{o,v_o,v_1\}} \frac{1}{\mu_k} M_k^t + \frac{1}{\mu_k} > T_w \quad \forall\ k \in \{o, v_0, v_1\} \quad (A.32)$$

Now, if we relax the time constraint and let the algorithm runs one more iteration, we have $M_{T_w}^{t+1} = M_{T_w} + e_x$, where e_x is the x^{th} unit vector, and $x = argmax_j\ \mu_j A'\left(M_j^k\right)$. Refer to (Fox 1966), we have

$$A\left(M_{T_w}^{t+1}\right) > A\left(M_{T_w}^*\right) \geq A\left(M_{T_w}\right) \quad (A.33)$$

$$\sum_{k\in\{o,v_o,v_1\}} \frac{1}{\mu_k} M_k^{t+1} > T_w \geq \sum_{k\in\{o,v_o,v_1\}} \frac{1}{\mu_k} M_{k,T_w}^* \geq \sum_{k\in\{o,v_o,v_1\}} \frac{1}{\mu_k} M_k^t \quad (A.34)$$

Inequality (A.33) and (A.34) shows that the optimal average long-run earnings and the working time are between those under the sub-optimal solution M_{T_w} and the solution $M_{T_w}^{t+1}$ that we allow to run one more iteration. By considering inequality (A.33), we have

$$A\left(M_{T_w}^*\right) - A\left(M_{T_w}\right) < A\left(M_{T_w}^{t+1}\right) - A\left(M_{T_w}\right) = A'(M_x^t) \leq max\ A'(M_k^t) \quad (A.35)$$

where $x = argmax_j\ \mu_j A'\left(M_j^t\right)$.

With inequality (A.33) and (A.35), we are ready to prove that the relative error by using the solution from Algorithm 2, M_{T_w}, as an approximation of $M_{T_w}^*$ is no greater than $\frac{max\ A'(M_k^t)}{A\left(M_{T_w}\right)}$.

$$\frac{A\left(M_{T_w}^*\right) - A\left(M_{T_w}\right)}{A\left(M_{T_w}^*\right)} \leq \frac{A\left(M_{T_w}^*\right) - A\left(M_{T_w}\right)}{A\left(M_{T_w}\right)} < \frac{max\ A'(M_k^t)}{A\left(M_{T_w}\right)} \quad (A.36)$$

Appendix C: Graphs of Policies Comparison in Section 3.4.6

Table C.1: Comparison of Policy-4 and 5 with time/capacity constraint when the fluctuation rate is 5%

	Algorithm-1						% Profit Gap	
Model with Capacity Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-4**	**Policy-5**
Average	15.28	16.71	0.92	9.94	6.76	1.49	5.64%	1.20%
Max	18.00	20.00	1.14	12.00	9.00	2.20	16.46%	7.18%
Min	13.00	14.00	0.70	8.00	5.00	1.11	0.59%	0.00%

	Algorithm-2						% Profit Gap	
Model with Time Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-4**	**Policy-5**
Average	16.50	19.25	0.86	11.30	7.95	1.43	6.57%	3.69%
Max	20.00	23.00	1.11	14.00	10.00	2.00	24.62%	23.29%
Min	14.00	16.00	0.64	9.00	6.00	1.11	0.11%	0.00%

Table C.2: Comparison of Policy-4 and 5 with time/capacity constraint when the fluctuation rate is 10%

	Algorithm-1						% Profit Gap	
Model with Capacity Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-4**	**Policy-5**
Average	15.10	16.61	0.92	9.90	6.71	1.54	8.97%	4.59%
Max	20.00	21.00	1.50	14.00	10.00	3.33	58.96%	33.66%
Min	10.00	12.00	0.50	6.00	3.00	0.88	0.00%	0.00%

	Algorithm-2						% Profit Gap	
Model with Time Constraint	M_o	M_v	$\frac{M_o}{M_v}$	M_{v_0}	M_{v_1}	$\frac{M_{v_0}}{M_{v_1}}$	**Policy-4**	**Policy-5**
Average	16.15	19.05	0.86	11.24	7.80	1.49	11.62%	8.85%
Max	21.00	25.00	1.29	17.00	12.00	2.60	85.70%	90.13%
Min	13.00	14.00	0.54	7.00	5.00	0.88	-1.78%	0.00%