1-1-2017

Essays On Stochastic Programming In Service Operations Management

Sina Faridimehr
Wayne State University,

ESSAYS ON STOCHASTIC PROGRAMMING IN SERVICE OPERATIONS MANAGEMENT

by

SINA FARIDIMEHR

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2017

MAJOR: INDUSTRIAL ENGINEERING

Approved By:

____

Advisor Date

____

____

DEDICATION

To my wife, my parents and my sister.

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisors Prof. Chinnam and Prof. Venkatachalam for continuous support during my Ph.D. studies and related research, for their patience, motivation, and immense knowledge. Their guidance helped me in all facets of the research and writing of this thesis. I could not have imagined having better advisors and mentors for my Ph.D. studies.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Yang, Dr. Dalkiran, and Dr. Fisher, for their insightful comments and encouragement, but also for the hard questions that encouraged me to broaden my research from various perspectives.

Last but not the least, I would like to thank my family: my wife, my parents and my sister for supporting me spiritually throughout writing this thesis and my life in general.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Within the domain of operations research (OR), *deterministic* mathematical modeling is a branch of optimization that has been extensively employed for decision making in real-world problems. In general, optimization involves finding the best solution for an objective function often limiting the search to a specific domain. While deterministic models assume that the data and parameters that form the objective function and any relevant constraints are known, real-world applications almost invariably include some 'unknown' parameters. The presence of this uncertainty can make the optimal solution of a deterministic model infeasible or sub-optimal to the decision making problem.

When the parameters are known within certain bounds, an approach to tackling such problems is *robust* optimization. On the other hand, *stochastic programming* allows the parameters and coefficients to be unknown and assumes that probability distributions governing the data are known or can be estimated. Stochastic programming can involve uncertainties in the objective function and/or constraints. Often, the goal is to find some policy or solution that is reasonably feasible for possible data instances and maximizes the expectation of some function of the decisions and the random variables [66].

More generally, such models are formulated, solved analytically or numerically, and analyzed in order to provide useful information to a decision-maker. Stochastic programming is seeing a growing range of applications in manufacturing production planning, machine scheduling, dairy farm expansion planning, asset liability management, traffic management, and automobile dealership inventory management that involve uncertainty in decision making. For a good overview, see Wallace & Ziemba [73], Uryasev & Pardalos [71].

In spite of the recognition that real-world problems often entail uncertainty, the OR literature is predominantly focused on deterministic optimization. There are several reasons for this, including: 1) Simplicity and better computational tractability of deterministic models, 2) Readily available commercial and open-source software, and 3) Inconvenience in terms of effort involved in characterizing uncertainties for stochastic programming. As noted by OR pioneer R.L. Ackoff [5], OR has devolved from "its original focus as a market-oriented profession ... to its current status as largely input-oriented" profession, and that "this descent has taken the profession away from the most important needs of the organizations it could serve." Meaning, OR has lost its way and shifted its focus from producing solutions to real and important problems to methodology development [27].

In an attempt to improve the practicality of mathematical programming models and contribute to their adoption in the real-world, this dissertation contributes two original essays highlighting the contribution that stochastic programming can offer in solving important practical problems of interest to operations research. The first essay studies a strategic problem in transportation industry and the target customer is planning agencies, and the second essay investigates a tactical problem in healthcare industry and the target customer is clinic administrators. We do this with the recognition that stochastic programming, in considering uncertainty in mathematical modeling, often leads to large-scale programming problems. However, advances made in recent decades by algorithms in optimization software combined with advances in computing hardware allow us to tackle problems of greater complexity to provide meaningful solutions and decision support for the real-world.

The most widely used stochastic programming approach involves the 'two-stage' stochastic programming model. In this model, first-stage decision variables are determined before observing the realization of uncertainties and second-stage decision variables are selected after exposing first-stage variables to the uncertainties. The goal is to determine the value of first-stage decisions in a way to maximize (minimize) the expected value of second-stage objective function. We employ the two-stage stochastic programming approach for tackling two important problems of OR. The first problem entails the task of designing a community-aware charging station network for electric vehicles and is more of a strategic problem. The second problem involves the development of effective scheduling policies for managing appointments to primary care clinics.

1.1 Motivation for Designing Community-Aware Charging Network for Electric Vehicles

Electric vehicles (EVs) are attracting more and more attention these days due to increase concern about global warming and future shortage of fossil fuels. These vehicles have potential to reduce greenhouse gas emissions, improve public health condition by reducing air pollution and improving sustainability, and address diversification of transportation energy feedstock.

Governments and policy makers have proposed two types of policy incentives in order to encourage consumers to buy an EV: direct incentives and indirect incentives. Direct incentives are those that have direct monetary value to consumers and include purchase subsidies, license tax/fee reductions, Electric Vehicle Supply Equipment (EVSE) financing, free electricity, free parking and emission test exemptions. On the other hand, indirect incentives are the ones that do not have direct monetary value and consist of high-occupancy vehicle access, emissions testing exemption time savings, and public charger availability. Lack of access to public charging network is a major barrier in adoption of EVs [39]. Access to public charging infrastructure will provide confidence for EV owners to drive longer distances without going out of charge and encourage EV ownership in the community.

The current challenge for policy makers and city planners in installing public charging infrastructure is determining the location of these charging service stations, number of required stations and level of charging since the technology is still in its infancy and the installation cost is high. Since recharging of EV battery takes more time than refueling conventional vehicles, parking lots and garages are considered as potential locations for installing charging stations. The aim of this research is to develop a mathematical programming model to find the optimal locations with potentially high utilization rate for installing community-aware public EV charging infrastructure in order to improve accessibility to charging service and community livability metrics. In designing this charging network, uncertainties such as EV market share, state of battery charge at the time of arrival, driver's willingness to charge EV away from home, arrival time to to the community, driver's activity duration (parking duration), and driver's walking distance preference play major role. By incorporating these uncertainties in the model, we propose a two-stage stochastic programming approach to determine the location and capacity of public EV charging network in a community.

1.2 Motivation for Managing Access to Primary Care Clinics

Patient access to care along with healthcare efficiency and quality of service are dimensions of health system performance measurement [1]. Improving access to primary care is a major step of having a high-performing healthcare system. However, many patients are struggling to get an in-time appointment with their own primary care provider (PCP). Even two years after health insurance coverage was expanded, new patients have to wait 82% longer to get an internal-medicine appointment. A national survey shows that percentage of patients that need urgent care and could not get an appointment increased from 53% to 57% between 2006 and 2011 [31]. This delay may negatively impact patient health status and may even lead to death. Patients that cannot get an appointment with their PCP may seek care with other providers or in emergency departments which will decrease continuity of care and increase total cost of health system.

The main issue behind access problem is the imbalance between provider capacity and patient demand. While provider panel size is already large, the shortage in primary care providers and the increase in number of patients mean that providers have to increase their panel size and serve more patients which will potentially lead to lower access to primary care. The ratio of adult primary care providers to population is expected to drop by 9% between 2005 and 2020 [12].

Moreover, patient flow analysis can increase efficiency of healthcare system and quality of health service by increasing patient and provider satisfaction through better resource allocation and utilization [40]. Effective resource allocation will smooth patient flow and reduce waste which will in turn results in better access to care.

One way to control patient flow in clinic is managing provider capacity through appointment scheduling system. A well-designed appointment scheduling system can decrease appointment delay and waiting time in clinic for patients and idle time and/or overtime for provider simultaneously and increase their satisfaction. Appointment scheduling requires to make a balance between patient needs and facility resources [13].

The purpose of this study is to develop appointment scheduling models using two-stage stochastic programming in order to improve access to care while maintaining high levels of provider capacity utilization and improving patient flow in clinic by leveraging uncertainties in patient demand volume, patient no-show, nurse service time and provider service time.

CHAPTER 2: DESIGNING COMMUNITY-AWARE CHARGING NETWORKS FOR ELECTRIC VEHICLES

2.1 Introduction

Electric vehicles (EVs) hold much promise including diversification of the transportation energy feedstock, reduction of greenhouse gas and other emissions, and improved public health by improving local air quality. In general, widespread adoption of EVs is in alignment with sustainable transportation objectives due to its social, economic, and environmental perspectives. It is estimated that an EV that draws its power from the U.S. electrical grid emits at least 30% less CO_2 than comparable gasoline or diesel-fueled vehicles [9]. As EV usage for daily commute increases, the consideration for the ability to recharge these vehicles away from home will become even more important. Ever-growing need to recharge EVs away from home necessitates designing effective networks of charging stations. Using multiple linear regression, Sierzchula et al. [67] examined the effect of consumer financial incentives and several socio-economic factors on national EV market shares of 30 countries for the year 2012. The analysis shows that installing one charging station (per 100,000 residents) could have twice the impact on EV adoption rate compared to a $1,000 financial incentive.

Many studies have been done on locating charging stations for EVs. However, majority of them concentrated on large-scale state-wide networks and only a few articles have investigated design of public charging station network in an urban area. Existing papers on charging station location problem often assume that demands for charging service are deterministic and known to the decision makers, while in reality, the traffic flows are stochastic in nature (varying by hour of day, weekday, weekend, commute purpose, destination etc) and carry significant uncertainty. The optimal solution of a deterministic model might become infeasible and/or significantly sub-optimal in the presence of these uncertainties. This paper adds to the growing field of designing EV charging station network by proposing a two-stage stochastic programming model to determine location and size of charging stations for a community. Considering uncertainties in charging pattern, demand, and drivers' behavior, the proposed stochastic model provides more robust charging network design decisions and thus access to charging service can be improved. However, a two-stage stochastic programming model often needs a large number of scenarios for good representation of uncertainties. We use sample average approximation (SAA) method as this will asymptotically converge to an optimal solution for a two-stage stochastic problem. SAA is a Monte Carlo simulation-based sampling technique in which we approximate the expected value of the objective function using a finite sample of scenarios. Since SAA can only solve small size problems within reasonable amount of time in general, an effective heuristic is also proposed for large-scale instances. The two-stage model and solution approach are evaluated by a case study constructed using the data representing Detroit midtown area in Michigan, U.S. In summary, the major contributions of this paper include: (1) formulation of a two-stage stochastic programming model to determine the location and capacity of public EV charging stations in an urban area to maximize access; (2) incorporation of uncertainties in EV demand flows, EV drivers' charging patterns, arrival and departure time, purpose of arrival to a community, and preferred walking distance; (3) adoption of SAA to solve the two-stage model; (4) an effective heuristic that provides near optimal solutions for large-scale instances; and (5) a case study representing public charging network planning in Detroit midtown area and a post-analysis framework to analyze the outputs of the two-stage model on accessibility and utilization of charging service. The remainder of this paper is organized as follows: A review of related literature is presented in Section 2.2. Section 2.3 provides problem description and the uncertainties considered in our model. Model formulation and the solution methodology are presented in Section 2.4. Section 2.5 presents the case study, scenario construction, computational experiments and evaluations of results. Finally, conclusion and directions for future studies are provided in Section 2.6.

2.2 Literature review

During the last decade, many researchers have focused on optimally locating alternative-fuel-vehicle refueling stations. However, most of them are focused on EV charging network in large networks to cover demand between cities and metropolitan areas, and only a few articles examined the design of charging network in a community or an urban area. We review the existing literature related to design of an EV charging network and categorize it into two major groups: (A) deterministic approach which assumes that all parameters and demand are known for charging station network problem, and (B) stochastic approach that considers uncertainties regarding available budget for constructing charging network, type of charging stations, total short-term and long-term charging demand, and charging behavior of EV drivers.

2.2.1 Deterministic approach

Upchurch et al. [70] introduced capacitated flow refueling location model that considers a limit on the traffic flow that any location can refuel to maximize vehicle miles traveled by alternative-fuel vehicles. Frade et al. [29] proposed a maximal covering model to find the optimal location of EV charging stations in an urban area by maximizing covered demand within a given distance. To deal with the computational burden of generating combinations of locations capable of serving the round trip on each route, a mixed-binary-integer optimization model is developed [15]. Capar et al. [16] presented a more computationally efficient model for flow-refueling location model to answer some strategic questions such as what is the minimum number of charging stations required for refueling a certain percentage of traffic flow; and what are the impacts of refueling demand forecast on the location of fuel stations. A mixed-integer programming method to model capacitated multiple-recharging-station-location problem considering budget constraint and vehicle routing behavior, and using the concepts of set coverage and maximum coverage is proposed [75]. The model in [7] finds the optimal locations of charging stations for EVs in an urban area while minimizing total costs, consisting the travel cost from demand zones to charging locations and investment cost. Cavadas et al. [18] proposed a mixed-integer programming model to locate slow-charging stations for EVs in an urban environment considering the possibility that there might be several stops by each driver during the day and the driver can only charge the vehicle at one of these locations. Since tour-based network equilibrium model can precisely track the state-of-charge (SOC) of the battery and also consider the dwell time at each destination, model is proposed to optimally locate public charging stations for EVs considering recharging behavior of drivers [35]. Huang and Zhou [37] developed an integer programming formulation to minimize the lifetime cost of equipment, installations, and operations of charging stations for plug-in EVs at workplaces by considering different charging levels and demographics of employees. In order to maximize the amount of vehicle-miles-traveled for an EV, a model is presented to select the optimal locations for public charging stations considering vehicle travel patterns [65]. The authors applied their model on vehicle trajectory data of taxi fleet over a three week period in Beijing, China. A major limitation with all these studies is that they assume a deterministic problem setting. As we confirm through our experiments, employing a stochastic formulation can lead to a significant improvement in the objective of the planners.

2.2.2 Stochastic approach

While planning under uncertainty has been addressed in many settings such as transportation, energy, disaster planning, supply chain management and production planning, the literature considering uncertainty in planning for EV charging network is limited. By considering both the transportation system and the power grid, Pan et al. [59] developed a two-stage stochastic programming model to find the optimal locations for battery exchange stations for plug-in hybrid electric vehicles (PHEV) accounting for uncertainty in demand for battery, loads, and generation capacity of renewable power sources. Tan and Lin [68] formulated the EV charging problem as a flow capturing location-allocation problem. They compared a deterministic case where charging demand is fixed over time to a stochastic one where consumer demand for charging service is random, and concluded that stochastic programming provides more realistic results. Hosseini and MirHassani [36] proposed a two-stage stochastic program to locate permanent and portable charging stations with and without considering capacities to maximize the served traffic flows. A stochastic flow-capturing location model is also developed to locate a predetermined number of fast EV-charging stations within a given region considering uncertainties in EV flows [77].

To efficiently assist city planners and policy makers in planning for public EV charging network within a community, we need to adequately capture uncertainties that exists in demand for public charging service. To the best of our knowledge, this is the first study to address the problem of locating public EV charging stations for a community using a two-stage stochastic programming approach while accounting for uncertainties in total customer demand for public charging service, arrival and dwell time, battery SOC at the time of arrival, preference for charging away from home and willingness to walk patterns of EV drivers.

2.3 Problem description and uncertainties

Unlike a conventional vehicle, an EV must often be parked for several hours to be recharged. Hence, public parking facilities are considered as potential locations for installing charging stations, which can in turn improve access to EVs as well as their adoption. Maximum number of installable charging stations depends on the total capacity of a parking lot. Without loss of generality, we assume that all charging point terminal types are semi-rapid charging ones (level 2 type charging stations) that are typically recommended for public and private parking lots, and provide 10 to 20 miles range per hour of charging. Also, a driver's walking distance to final destination is considered as the decisive contributing factor in choosing a parking lot [8]. Based on a driver's walking distance preference, we determine a possible set of parking lots that a driver can park the EV and then driver is randomly assigned to one of them. If charging stations are installed in any of parking lots that are within a driver's walking distance preference, driver will be attracted to one of those parking lots depending on the availability of a charging station at the time of arrival. If there is no parking lot within the maximum distance that a driver is willing to walk, we assume that driver will park the car on street, and since it is difficult to track the walking distance to final destination in this case, this demand is not considered in our analysis. It is also assumed that once a driver starts using a charging station, vehicle would not be unplugged until driver's activity is finished.

Designing a public EV charging network entails estimation of demand for charging service. Like facility location models, we assume that demand occurs at fixed points on a network. Demand will be attracted to different parking lots based on drivers' willingness to walk to use charging stations. Scenarios representing demand uncertainty in the two-stage model will represent time and purpose of arrival to the community, EV's battery SOC at the time of arrival, duration of activity, drivers' preference for charging away from home and willingness to walk based on demographics, community size and seasonality factors. The following uncertainties are considered to affect demand for public EV charging stations:

2.3.1 State of charge

A recent study analyzing two years of data from January 2011 to December 2013 of charging events that occurred away from home concluded that Nissan Leaf (pure battery electric vehicle, BEV) drivers prefer to charge their vehicles before their battery SOC drops to lower levels while Chevrolet Volt (a plug-in hybrid electric vehicle, PHEV) drivers tend to start recharging when there is a little charge in the battery since they rely on both electric motor and internal combustion engine [14]. Fig. 1 compares the probability of recharging for different values of battery SOC at the time of arrival for Nissan Leaf and Chevrolet Volt.

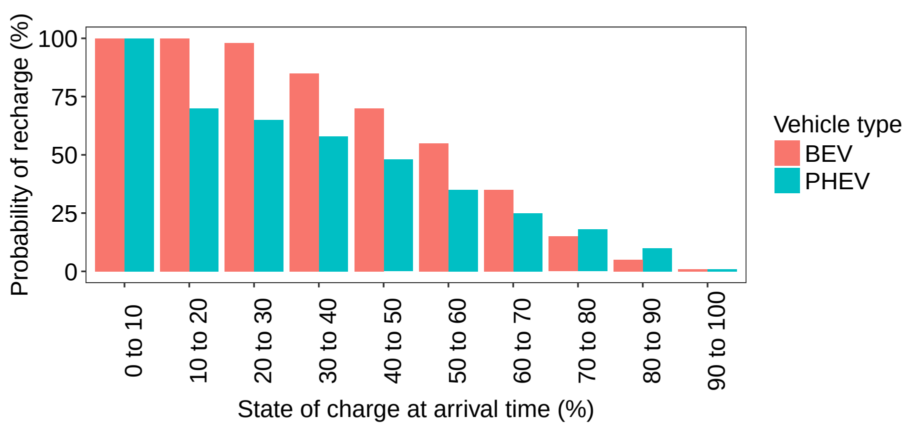

Figure 1: Probability of recharging as a function of the battery SOC at arrival time; Source: [14].

2.3.2 Dwell time

We define six different destination categories based on NHTS (National Household Travel Survey) data: Work, Social, Family, Meal, Study, and Shopping. Fig. 2 shows average time that people tend to park their vehicles based on their activity type [43]. Zhong et al. [82] concluded that Weibull, log-normal and log-logistic distributions are the best distributions for modeling duration of weekday and weekend activities. While their analysis shows that model type and parameters or both might be different for an activity in weekday versus weekend, they found Weibull distribution the most applicable one. In addition, they found that certain activities such as social and shopping tend to last longer during weekends. Weibull distribution is used in our analysis to estimate parking duration of EV drivers considering average staying time, and we have also differentiated the durations of all weekday and weekend activities except meal activity.

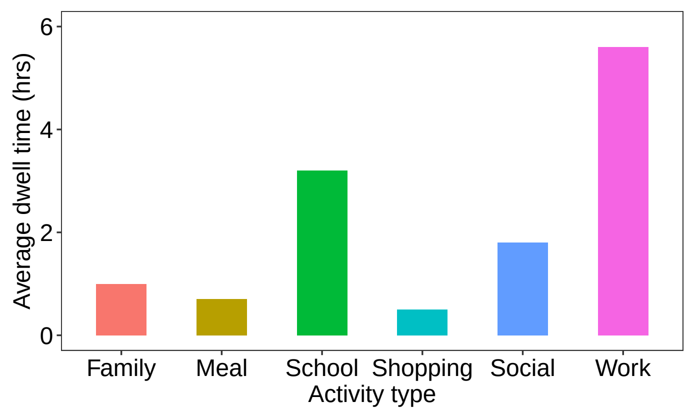

Figure 2: Average dwell time for activity types; Source: [14].

2.3.3 Weekday vs. weekend

Demand pattern for public EV charging service can vary from day to day since people tend to attend social events, visit their families and go to shopping centers more during weekends than weekdays, in which demand mostly consists of people traveling to work or school. Fig. 3 confirms that demand for charging stations depends on time and type of day. During weekdays, maximum load occurs in morning when people are arriving at work or school while maximum demand usually happens around noon during weekends when people are going to shopping malls and social places. According to [58], the best fitted distribution for arrival time to parking lot is a Weibull distribution. Hence, without loss of generality, we recommend the use of two Weibull distributions to estimate the arrival time of EVs to parking lots during weekdays and weekends.

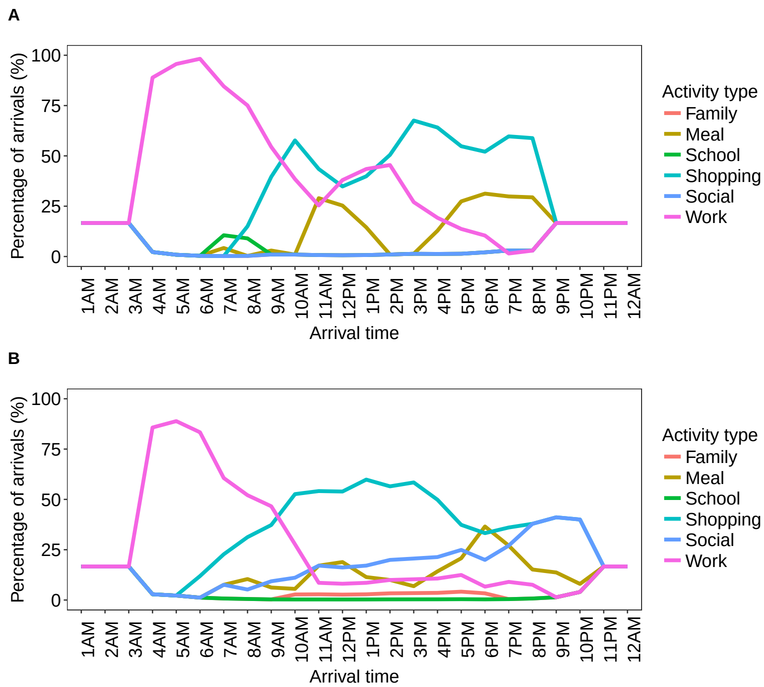

Figure 3: The expected breakdown of vehicle arrival percentages in A) weekdays and B) weekends; Sources: [14] and [43].

2.3.4 Preference for charging away from home

Analysis by Idaho National Laboratory on data from 2012 and 2013 over 4,000 Leafs and 1,800 Volts across the U.S. shows that 13% of Leaf drivers and 5% of Volt drivers only charge their vehicles at home. This indicates that vast majority of drivers intend to use publicly accessible charging stations. This analysis also shows that although many people that drive more daily miles tend to charge their vehicles in places other than their homes, the effect of daily miles traveled on the chance of charging away from home is small. Hence, without loss of generality, we do not consider the effect of driving distance to community as a factor that affects the chance of using EV charging stations.

2.3.5 EV market penetration

There are many social, environmental and economic factors that can significantly contribute to the increasing market share of different types of EVs. The survey in [17] about adult drivers in large U.S. cities in fall 2011 comprehended factors affecting the purchase of a plug-in EV. Besides demographic variables that can strongly predict intent of purchase, their results show that the presence of a charging station inside the community is the only awareness variable that has a significant effect on intent of purchase. Environmental Protection Agency estimated that 3.5% of the vehicle fleet will be BEV or PHEV in the 2022-2025 time frame [4].

2.3.6 Willingness to walk

The drivers' willingness to walk can be affected by their socio-demographic characteristics such as age, gender, education level and occupation. Many researchers have used distance decay function that shows the willingness to walk or bike as a distance towards different types of destinations. The parameter of this decay function depends on the activity type. Estimation results from [79] confirm that negative exponential distribution can better describe walking trips over short distances than other distributions such as Gaussian. They specify the distance decay function as

$$P(d) = e^{-\beta \times d} \quad (1)$$

which shows the percentage of people willing to walk d or longer distances than d. They used 2009 NHTS data to estimate the decay parameter β for different groups and trip purposes. Their analysis shows that people are more willing to walk for recreation, social events and work activities rather than for studying, shopping or eating meal. Table 1 shows the parameters of distance decay function influenced by variations in natural and built environment factors. The effects of season, region and community size on willingness to walk patterns are considered as well.

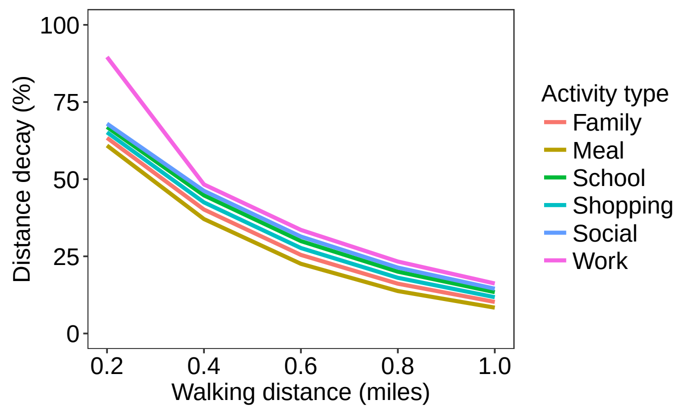

Figure 4: Distance decay function for walking trips to different destination types; Source: [79].

Table 1: Estimated parameter for distance decay function

Factor	Category	β
Season	Winter (Dec to Feb)	1.88
	Spring (Mar to May)	1.68
	Summer (Jun to Aug)	1.64
	Autumn (Sep to Nov)	1.7
Region	Northeast	1.85
	Midwest	1.65
	South	1.76
	West	1.65
Community	Town and country	1.68
	Suburban	1.63
	Urban and second city	1.78

Our research aims at maximizing coverage of demand for public EV charging network in a community by proposing a two-stage stochastic programming model considering uncertainties in EV total flow, arrival and departure time, battery SOC at arrival time, preference for charging EV away from home, and walking preference patterns in the community.

2.4 Model formulation and solution approach

2.4.1 Model formulation

Two-stage stochastic programming is a common approach for modeling problems that involve uncertainty in decision making. First-stage decision variables represent 'here-and-now' decisions which are determined before the realization of randomness, and the second-stage decisions are determined after scenarios representing uncertainties are presented. Through two-stage stochastic programming model, we aim at maximizing expected accessibility of EV owners to public charging service considering all uncertainties in demand for charging stations. In our model, binary variables in the first-stage determine the parking lots, and number of charging station installations for the selected parking lots. In the second-stage, a recourse decision is made on assigning EV drivers to one of their preferred parking lots based on their willingness to walk so that the expected access of EV drivers to public charging network is maximized.

We first define the following model sets, parameters and variables:

- Sets
  - S: Set of parking lots, indexed by $s \in S$.
  - L_s: Set of number of charging stations in a parking lot s, indexed by $l \in L_s$.
  - B: Set of buildings, indexed by $b \in B$.
  - T: Set of time slots, indexed by $t \in T$.

  - Γ: Set of arrival and departure times, indexed by $\gamma(t) \in \Gamma$ containing time slot $t \in T$.
  - Ω: Set of scenarios, indexed by $\omega \in \Omega$.

- Model Parameters
  - p: Number of parking lots to be considered for installing charging stations.
  - m_l: Number of charging stations, $l \in L_s$.
  - $d_{\gamma(t),b,s}(\omega)$: Demand with arrival and departure time set of $\gamma(t) \in \Gamma$ for a given $t \in T$ for building b that are willing to park their vehicle in parking lot $s \in S', S' \subset S$ in a scenario $\omega \in \Omega$.

- First Stage Decision Variables
  - x_s: 1, if parking lot $s \in S$ is considered for installing charging stations; 0, otherwise.
  - $z_{l,s}$: 1, if $l \in L_s$ charging stations are installed in parking lot $s \in S$.

- Second Stage Decision Variables
  - $y_{\gamma(t),b,s}(\omega)$: Proportion of demand with arrival and departure time set of $\gamma(t) \in \Gamma$ for building b willing to charge their vehicle in parking lot $s \in S', S' \subset S$ in a scenario $\omega \in \Omega$.

First-stage problem can be presented as follows:

$$\text{Max } E_{\Omega}[\varphi(x, z, \widetilde{\omega})] \tag{2}$$

s.t.

$$\sum_{s \in S} x_s = p \tag{3}$$

$$z_{l,s} \leq x_s \qquad \forall s \in S, l \in L_s \tag{4}$$

$$\sum_{l \in L_s} z_{l,s} \leq 1 \qquad \forall s \in S \tag{5}$$

$$x_s, z_{l,s} \in \{0, 1\} \qquad \forall s \in S, l \in L_s \tag{6}$$

where $\varphi(x, \widetilde{\omega})$ is the solution of the following second-stage problem:

$$\text{Max } \varphi(x, z, \omega) = \sum_{t \in T} \sum_{\gamma(t) \in \Gamma} \sum_{b \in B} \sum_{s \in S} y_{\gamma(t),b,s}(\omega) d_{\gamma(t),b,s}(\omega) \tag{7}$$

s.t.

$$\sum_{\gamma(t) \in \Gamma} \sum_{b \in B} y_{\gamma(t),b,s}(\omega) d_{\gamma(t),b,s}(\omega) \leq \sum_{l \in L_s} m_l z_{l,s} \qquad \forall s \in S, t \in T \tag{8}$$

$$\sum_{s \in S} y_{\gamma(t),b,s}(\omega) \leq 1 \qquad \forall t \in T, \gamma(t) \in \Gamma, b \in B \tag{9}$$

$$0 \leq y_{\gamma(t),b,s}(\omega) \leq 1 \quad \forall t \in T, \gamma(t) \in \Gamma, b \in B, s \in S \tag{10}$$

In this model, first-stage decisions are made regarding the locations of charging stations and charging capacity in each location. The first-stage objective function maximizes the expected access, and E_{Ω} is an expectation operator, and $E_{\Omega}[\varphi(x, z, \omega)]$ represents $\sum_{\omega \in \Omega} p_{\omega} \varphi(x, z, \omega)$, where p_{ω} is probability of occurrence for scenario ω, and $\sum_{\omega \in \Omega} p_{\omega} = 1$. Constraint (3) ensures that p parking lots are selected to install EV charging stations. Constraints (4) and (5) determine charging capacity in any parking lot that is selected for providing EV charging service. Constraints (6) define the feasible set for the binary first-stage variables. In the second-stage, recourse decisions are made to maximize the coverage of potential EV traffic flows based on the decisions chosen in the first-stage and a realization $\omega \in \Omega$. Constraints (8) describe the supply-demand balance restrictions. They ensure that demand that has arrival and departure time set of $\gamma(t)$ and are assigned to parking lot s for EV charging does not exceed the charging capacity in parking lot s. Constraints (9) state that demand with arrival and departure time set of $\gamma(t)$ can be assigned to at most one parking lot for EV charging. Constraint set (10) are the non-negativity constraints. Though we have not considered any budgetary restrictions, such constraints can be added to the first-stage model if appropriate.

2.4.2 Sample average approximation

According to [53], unless there are small number of scenarios that can represent uncertainties in a problem, it is usually impossible to solve a stochastic programming problem. They showed that optimal solution of stochastic programming can be approximated by a sample of scenarios much smaller than the actual size of scenarios and this approximation monotonically improves as we increase the number of scenarios. SAA is also an effective approach when sufficient number of scenarios to estimate optimal solution is unknown. SAA was proposed by [53] and for the sake of completeness, we provide the procedure for sample average approximation method as follows:

1. Estimating an upper bound for the optimal solution:

    - Generate M independent sample sets of scenarios each of size N, i.e., $(\omega_j^1, \omega_j^2, ..., \omega_j^N)$ for $j = 1, 2, ..., M$
    - For each sample set $j = 1, 2, ..., M$, find the optimal solution:

$$v_N^j = \frac{1}{N} \sum_{i=1}^{N} \varphi(x, z, \omega_j^i). \tag{11}$$

- Compute the followings:

$$\overline{v}_{N,M} = \frac{1}{M}\sum_{j=1}^{M} v_N^j \tag{12}$$

$$\sigma^2_{\overline{v}_{N,M}} = \frac{1}{M(M-1)}\sum_{j=1}^{M}(v_N^j - \overline{v}_{N,M})^2. \tag{13}$$

The expected value of v_N is greater than or equal to the optimal value v^*. Since the sample average $\overline{v}_{N,M}$ is an unbiased estimation of the expected value of v_N, $\overline{v}_{N,M}$ provides an upper statistical bound for the optimal solution.

2. Estimating a lower bound for the optimal solution:

- If $(\overline{x}, \overline{z})$ is a feasible solution for the first-stage problem, then $f(\overline{x}, \overline{z}) \leq v^*$. Hence, choosing any feasible solution of the first-stage problem will provide a lower statistical bound for the optimal value.

- Choose a sample of size N' of scenarios, much larger than N, i.e., $(\omega^1, \omega^2, ..., \omega^{N'})$ and independent of samples to find the upper limit and estimate the objective function:

$$f(\overline{x}, \overline{z}) = \frac{1}{N'}\sum_{i=1}^{N'} \varphi(x, z, \omega^i) \tag{14}$$

- Compute the variance for this estimation:

$$\sigma^2_{N'}(\overline{x}, \overline{z}) = \frac{1}{N'(N'-1)}\sum_{i=1}^{N'}(\varphi(x, z, \omega^i) - f(\overline{x}, \overline{z}))^2. \tag{15}$$

3. Estimating the optimality gap:

- Use the upper bound and the lower bound that are computed in previous steps to estimate the optimality gap:

$$gap_{M,N,N'}(\overline{x}, \overline{z}) = \overline{v}_{N,M} - f(\overline{x}, \overline{z}). \tag{16}$$

4. Checking the quality of the estimated optimality gap:

    - Variance of the estimated optimality gap can be found by

$$\sigma_{gap}^2 = \sigma_{\overline{v}_{N,M}}^2 + \sigma_{N'}^2(\overline{x}, \overline{z}) \tag{17}$$

2.4.3 Heuristic

SAA requires high computational resources, hence we developed a heuristic to solve large-scale problems efficiently. This heuristic is inspired by a score measure introduced by [69]. The score incorporates charging capacity of each parking lot as well as its distance to other parking lots. The heuristic consists of a construction phase during which we build an initial solution, and an improvement phase where we employ local search moves to find a better solution. The pseudo-code of the heuristic is presented as follows:

Algorithm 1 Pseudo-code of the heuristic

$bestsolution \leftarrow \emptyset$.
for $s \leftarrow 1$ to $NumberofParkingLots$ **do**:
    Compute score measure r_s.
end for
Construction phase:
$initialsolution \leftarrow \emptyset$
Compute attractiveness ratio ρ_s for all parking lots.
Add parking lots to the initial solution in decreasing order of the attractiveness ratio until p parking lots are selected.
Improvement phase:
$currentsolution \leftarrow initialsolution$
while $f(currentsolution)$ can be improved **do**
    remove-insert($currentsolution$)
end while
Store best solution found so far.

In the construction phase, a score measure for each parking lot as a potential location for installing charging stations is calculated as:

$$r_s = \sum_{s,s' \in S, s' \neq s} c_s e^{-\beta d_{s,s'}} \tag{18}$$

where β is a user parameter. The score is measured as an incentive for the charging capacity (c_s) of each parking lot, and distance ($d_{s,s'}$) to other parking lots as a cost. If a parking lot has more capacity for installing charging stations and is nearer to other parking lots, its score would be higher.

To consider randomness in constructing the initial solution, we use a set of sample scenarios to get the probability of parking lot s being chosen as one of the optimal locations for installing charging stations. This estimated probability for parking lot s, q_s, is computed based on the fraction of scenarios in which parking lot s is among the optimal locations. The attractiveness measure of parking lot s, ρ_s, is computed by multiplying this probability to the corresponding score measure:

$$\rho_s = r_s q_s \tag{19}$$

Parking lots will be added to the initial solution in a decreasing order of attractiveness measure until p parking lots are selected. In the improvement step, we use local search method of remove-insert procedure. For every parking lot that is already in the initial solution, we replace it with one of the parking lots that has not been selected based on a parking lot that has the highest attractiveness measure. This process is continued until there is no improvement in the objective function. We repeat this procedure for all parking lots that are selected in the initial solution and store the best value found for the objective function.

2.5 Case study and computational experiments

To demonstrate the efficacy of the proposed approach, our case study investigates the community area data of Detroit midtown area in Michigan, U.S. There is a wide range of employment types (type of final destinations) in this area, it attracts a lot of traffic, and is characterized by an urban university, commercial offices, hospitals, and museums. This area includes 135 buildings among which 67 are office buildings, 12 are social places, 5 are family related buildings, 4 are restaurants, 44 are schools buildings and 3 are shopping places. There are 32 parking lots that are considered as potential locations for installing EV charging stations. We assume that parking lots are open between 6am and 6pm, and have different capacities for installing charging stations. The center of each parking lot is considered as our candidate for installing a charging station, and Euclidean distance is used to measure distance between any two points in the community. Data from Southeast Michigan Council of Governments shows that average annual daily traffic of Detroit midtown area is approximately between 10,000 and 20,000 and like [16], we assume that total daily traffic of this community follows a uniform probability distribution.

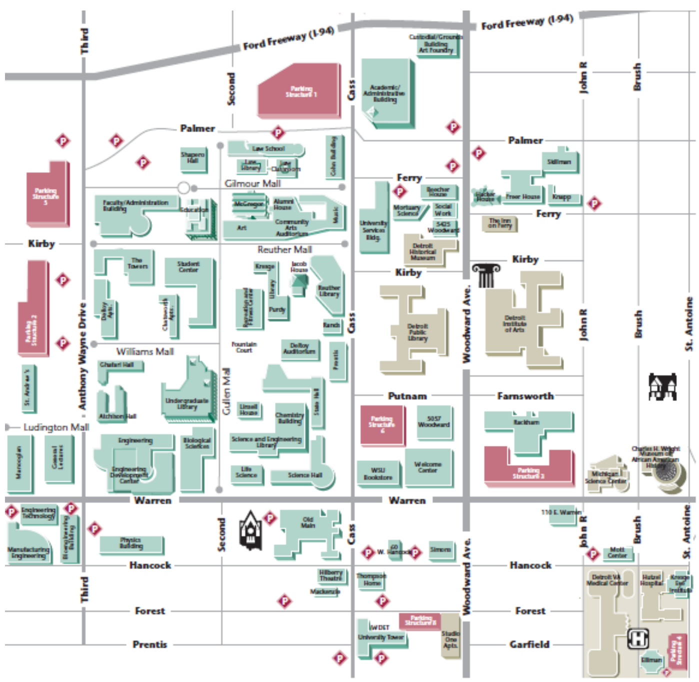

Figure 5: Part of Detroit midtown area used for our analysis.

Based on Environmental Protection Agency (EPA) analysis, we examine cases in which EVs constitute 3% and 5% of the light-duty vehicle fleet. According to [72], weather/climate is positively correlated with BEV market share. Since our case study is done in an area with low winter temperatures, BEV market share is considered lower than PHEV market share. Two cases are constructed for our computational experiments. In the first case, we assume that the market share is 1% and 2% for BEVs and PHEVs, respectively. In the second case, these market shares are assumed to be 2% and 3%, respectively.

In this study, negative exponential distribution functions estimated by [79] are used to describe willingness to walk patterns for various activity types which considers the effects of season and community size in U.S. Drivers are randomly assigned to a parking lot that is within their walking distance preference. In both cases of EV market share, 13% of total demand is not considered in our model since there is no parking lot within their walking distance preference, and also it is difficult to track their walking distance to their final destination if they use other EV charging sources placed in streets, etc. Four different values (2,4,6 and 8) for number of parking lots (p) to install charging stations are considered. The optimization models for SAA and heuristic were implemented in Python 2.7 using Gurobi 6.5.1 software for solving optimization problems. All the computations were performed using a system with Intel (R) Xeon(R) CPU 3.10 GHz and 24GB RAM.

2.5.1 Scenario construction

For the two-stage model, uncertainties are modeled by use case scenarios. A scenario represents a single day of public EV charging service and is influenced by short-term (weekday vs. weekend) and long-term (seasonal) variations, and total number of EVs arriving to the community. The probability of occurrence for a scenario is based on a uniform probability distribution. Without loss of generality, we assume that any given scenario day can belong to winter, spring, summer, and autumn seasons with equal probability.

In each scenario, a random number from $U(0, 1)$ determines type of each vehicle in a community, and if the random number is less than BEV market share, between BEV market share and sum of BEV and PHEV market shares, or greater than sum of BEV and PHEV market shares, then the vehicle is assumed to be a BEV, a PHEV or an ICE (internal

Table 2: Weibull distribution parameters for activity duration

Type of day	Work	Social	Family	Meal	School	Shopping
Weekday	(5.89,10)	(1.89,10)	(1.05,10)	(0.79,2)	(3.61,2)	(0.56,2)
Weekend	(6.04,6)	(2.03,2)	(1.13,2)	(0.79,2)	(3.36,10)	(0.25,0.5)

combustion engine), respectively. If it is an EV, Weibull distributions with parameters (8,3) and (13,4) are used to determine arrival time of EV drivers to the community in a weekday and weekend day, respectively. As explained in the earlier section, the purpose of arrival for a driver is determined based on arrival time and distributions. Furthermore, Weibull distribution is used to estimate duration of various types of weekday and weekend activities. Table 2 represents the parameters for this distribution based on type of activity. In this table, the first and second numbers represent the shape parameter and the scale parameter, respectively.

For a final destination, each EV driver is randomly assigned to a target destination/building using a uniform distribution based on a driver's purpose of arrival to the community. A random number is generated from exponential distribution as shown in Fig. 4 to determine each EV driver's willingness to walk distance based on his/her purpose of arrival, and also, community size and type of region are considered in willingness to walk distributions. If there is no parking lot within a driver's willingness to walk distance, then this demand is not considered in our model. In order to incorporate charging preference of EV drivers, uniform distribution $U(0, 1)$ is used. If the random number is greater than 13% for BEV or 5% for PHEV, a driver's willingness to charge away from home is decided. Consistent with recommendations from [23] and [51], without loss of generality, we assume that the initial battery SOC for vehicles arriving at the charging stations follows a normal distribution $N(0.3, 0.1)$ with a mean 0.3 and standard variation 0.1. Based on battery SOC at arrival time, uniform distribution $U(0, 1)$ is used to determine each EV driver's willing to charge EV at public charging stations. This is further compared with associated probability of recharge based on type of EV discussed in SOC section earlier. If the random number is less than or equal to the probability of recharge, that EV is considered as demand for EV charging network in the community. Similarly, multiple scenarios are constructed for the two-stage stochastic programming model to simulate the arrival pattern, battery SOC, dwell time, charging preference and willingness to walk in the community.

2.5.2 SAA settings

To estimate an upper bound for expected accessibility to public EV charging stations, N = 30, 50 and 100 scenarios are used and this is repeated M = 20 times. The average of these 20 runs is an estimate of upper bound on the accessibility. A sample of N' = 1,000 scenarios, which are separate from those that were used to get the upper bound, is used to estimate a lower bound for the optimal solution. Computation times for each test problem along with the heuristic performance are summarized in Tables 3 and 4. The computation times show that the optimization model using SAA method is able to solve problems with eight optimal locations in less than five hours. In these tables, UB (%) and LB (%) represent upper and lower bounds for expected accessibility to public EV charging service using SAA method. Gap (%) and gap SD indicate the differences between upper and lower bounds and standard deviation, respectively. Opt(s) is the running time of SAA. The best solution found by our heuristic for upper bound of the objective function and its running time are shown as Heuristic (%) and Heuristic (s).

2.5.3 Heuristic settings

In our heuristic, we consider $\beta = 0.0001$ since we found that this number gives better results in our setting. In order to consider the effect of uncertainties in our heuristic, we also use 10 pre-sampled scenarios to measure the attractiveness of each parking lot for installing charging stations. After computing attractiveness measures, we add parking lots to the initial solution in a decreasing order until we have p parking lots in the solu-

Table 3: SAA performance when (M, N') = (20,1,000) and (BEV,PHEV) = (1%,2%)

p	N	UB (%)	LB (%)	gap (%)	gap SD	Opt (s)	Heuristic (%)	Heuristic (s)
2	30	57.98	56.59	2.39	0.0064	397	57.98	68
	50	58.70	58.25	0.77	0.0062	1,226	58.70	74
	100	58.56	58.54	0.02	0.0055	4,564	58.56	93
4	30	73.89	73.42	0.63	0.0056	720	73.88	114
	50	74.61	73.85	1.02	0.0041	1,759	74.61	131
	100	74.59	73.74	1.14	0.0040	7,406	74.59	193
6	30	83.97	83.62	0.35	0.0039	1,071	83.21	160
	50	84.11	83.80	0.31	0.0034	2,173	83.17	186
	100	83.40	83.30	0.10	0.0031	9,572	82.86	303
8	30	91.16	90.61	0.61	0.0026	1,124	90.28	185
	50	91.13	90.78	0.38	0.0021	3,099	90.18	245
	100	90.87	90.86	0.02	0.0018	12,832	90.11	414

Table 4: SAA performance when (M, N') = (20,1,000) and (BEV,PHEV) = (2%,3%)

p	N	UB (%)	LB (%)	gap (%)	gap SD	Opt (s)	Heuristic (%)	Heuristic (s)
2	30	50.42	50.00	0.85	0.0056	462	50.42	82
	50	50.91	50.10	1.58	0.0054	1,141	50.91	87
	100	50.91	50.31	1.17	0.0048	4,761	50.91	106
4	30	63.35	63.16	0.30	0.0064	1,595	63.33	169
	50	63.19	63.11	0.13	0.0063	3,644	63.19	211
	100	63.46	63.42	0.07	0.0057	16,656	63.41	317
6	30	72.56	71.55	1.39	0.0071	1,663	72.34	208
	50	72.04	71.46	0.81	0.0059	3,246	71.84	273
	100	71.82	71.40	0.58	0.0050	12,165	71.73	474
8	30	78.91	78.49	0.52	0.0048	1,494	78.53	273
	50	79.44	78.92	0.66	0.0045	2,908	79.01	374
	100	79.12	78.69	0.54	0.0044	12,248	78.70	667

tion. Using local search moves, we try to improve this initial solution. To evaluate the performance of our heuristic, we compare 5 different runs of exact optimization and the heuristic results for 4 different number of scenarios. Tables 5 and 6 show the performance of the proposed heuristic in terms of running time and solution quality.

Table 5: Heuristic performance when (BEV,PHEV) = (1%,2%) when $p = 6$

Scenarios	Exact (%)	Exact (s)	Heuristic (%)	Heuristic (s)	gap (%)
50	73.09	203	72.48	12	0.84
	69.74	281	69.71	14	0.04
	71.72	96	71.72	16	0.00
	69.88	131	69.87	13	0.02
	71.13	86	71.13	15	0.00
100	72.00	409	72.00	23	0.00
	72.77	555	72.75	22	0.02
	71.31	747	71.31	25	0.00
	72.80	1,014	72.80	21	0.00
	70.23	588	70.23	24	0.00
150	71.72	2,312	71.71	36	0.01
	71.05	929	70.79	30	0.37
	70.31	813	70.31	33	0.00
	73.35	2,655	73.32	27	0.04
	71.57	1,294	71.56	35	0.01
200	73.09	5,280	73.04	45	0.06
	70.96	1,945	70.44	36	0.74
	71.35	3,904	71.35	40	0.00
	71.76	3,481	71.74	44	0.03
	71.44	1,680	71.44	40	0.00

As shown in Tables 5 and 6, the average running time for the heuristic algorithm was much less than the average running time of the exact algorithm (27 seconds versus 1420 seconds for $p = 6$ and 41 seconds versus 1336 seconds for $p = 8$). In addition, the results indicate that the heuristic is capable of producing good quality solution in all cases. The gap of feasible solution obtained from the proposed heuristic is on average 0.11% for $p = 6$ and 0.51% for $p = 8$.

Table 6: Heuristic performance when (BEV,PHEV) = (1%,2%) when $p = 8$

Scenarios	Exact (%)	Exact (s)	Heuristic (%)	Heuristic (s)	gap (%)
50	77.44	259	77.37	196	0.08
	78.30	133	78.09	20	0.27
	81.42	113	80.77	17	0.80
	79.04	201	78.65	18	0.49
	81.45	116	80.54	15	1.12
100	78.74	546	78.31	32	0.54
	76.99	810	76.99	38	0.00
	78.97	1,083	78.67	33	0.38
	81.73	413	80.69	33	1.27
	78.15	534	78.13	31	0.02
150	78.62	1,276	78.37	48	0.31
	79.14	909	78.80	44	0.42
	79.71	1,958	79.15	48	0.71
	78.50	1,194	78.34	53	0.20
	79.63	803	79.11	54	0.65
200	78.41	1,784	78.16	67	0.32
	79.14	2,710	78.75	67	0.50
	80.01	3,873	79.44	70	0.72
	79.43	4,128	78.93	67	0.63
	79.44	3,876	78.78	62	0.84

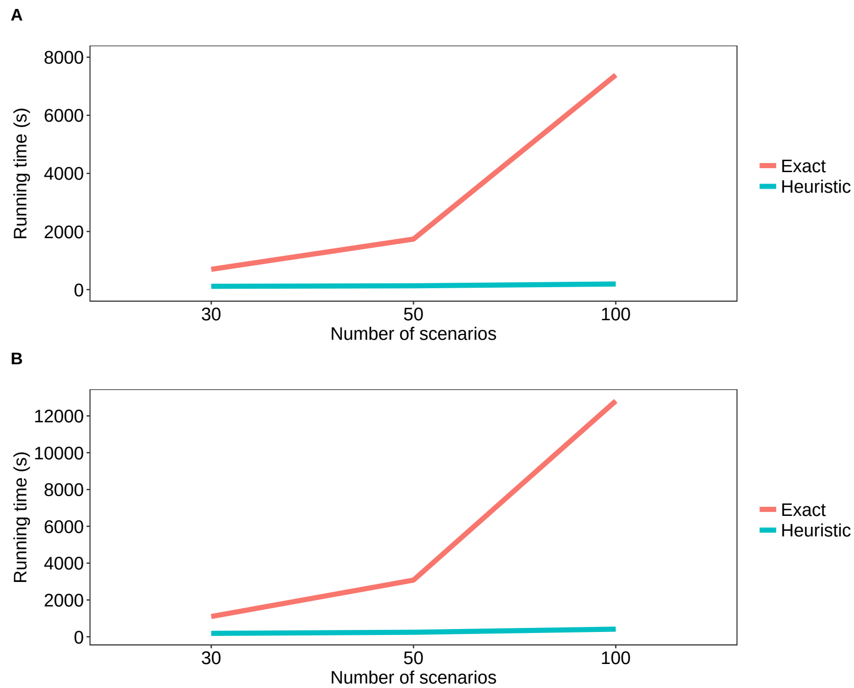

Figure 6: Comparison of exact running time vs. heuristic running time for a) $p = 4$ and b) $p = 8$ cases when (BEV,PHEV) = (1%,2%).

2.5.4 Performance measures

Number of public chargers per capita could have a significant effect on both BEV market share and PHEV market share. In terms of monetary benefits for EV consumers, the average of total benefits across 25 major metropolitan areas is around $2,800 per BEV and $1,600 per PHEV [52]. In order to deal with uncertainties in demand for public EV charging service and simulate the expected output measures with different number of chargers in the community, a set of 50 scenarios are generated and used for our analysis. We study two different cases for willingness to walk pattern in the community to generate optimistic and pessimistic bounds for level of walking in people that have access to public EV charging network. In the optimistic case, we assume that people are willing to walk long distances and will always choose the farthest available charging station to their final destination whereas in the pessimistic case people are willing to walk short distances and always choose the nearest available station to their building. Five different indicators are used to measure the performance of public EV charging placement: accessibility, lost demand, charging utilization, total walking distance, and walking distance per capita. Access is defined as the percentage of EV drivers that could charge their vehicles in public charging stations in the community, and lost demand is the percentage of EV drivers that are willing to use public EV network but there is not enough capacity to serve them. Charging utilization is the percentage of time that a charging station is being used by an EV. To assess walking patterns among people before network and after installing pubic EV charging stations, we use total walking distance and walking distance per capita measures.

As shown in Figs. 7 and 8, accessibility to public charging service increases in both cases of EV market share as more charging stations are installed in the community but utilization level of these stations reduces simultaneously. Increase in EV market share can reduce accessibility to public charging network up to 32% in both optimistic and pessimistic cases. However, this increase in demand will increase utilization level up to 41% and lost demand up to 68%.

Figs. 9 and 10 compare the average percentage of hourly utilization level of charging stations in weekdays versus weekends in an optimistic case of willingness to walk, and indicate a difference in utilization pattern from weekday to weekend. Utilization peaks around 8 in the morning during a weekday while it is around noon during weekend. These patterns match the expected arrival pattern of people to the community in weekdays and weekends. These plots also indicate that charging stations would not be fully utilized as more charging stations are available for EV drivers. This is important from revenue perspective since utilization level is among the major drivers of profitability of investment on public EV charging stations [22].

An important measure of livability analysis via transportation is increasing the travel options so that people can meet at least a part of their travel needs through walking and

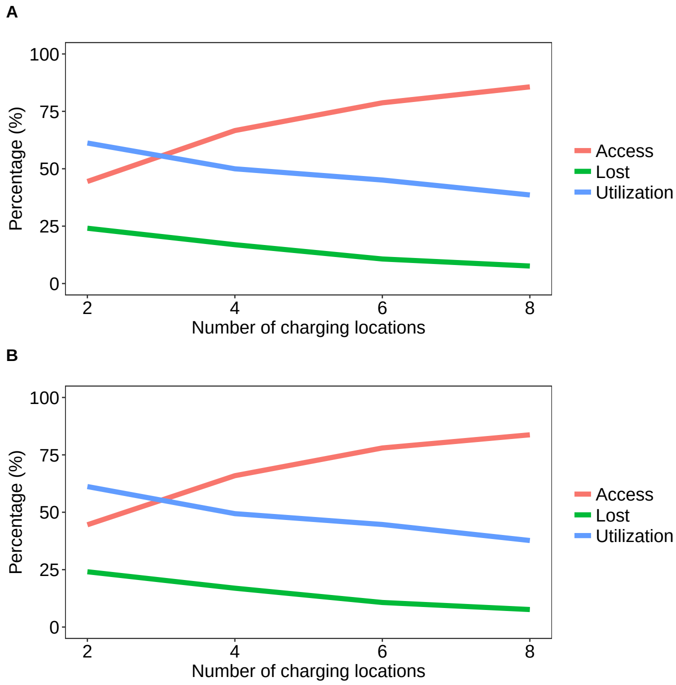

Figure 7: Percentage of accessibility, lost demand and charging utilization in A) optimistic and B) pessimistic cases when (BEV,PHEV) market shares are (1%,2%).

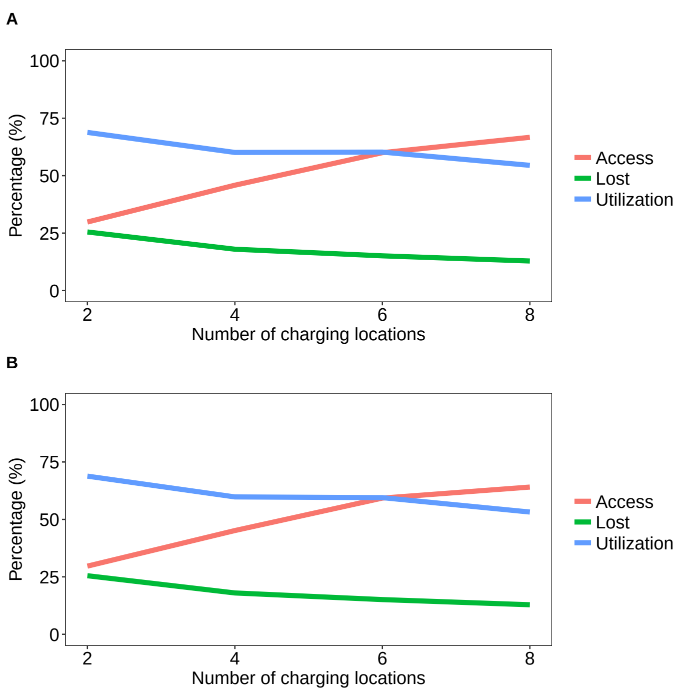

Figure 8: Percentage of accessibility, lost demand and charging utilization in A) optimistic and B) pessimistic cases when (BEV,PHEV) market shares are (2%,3%).

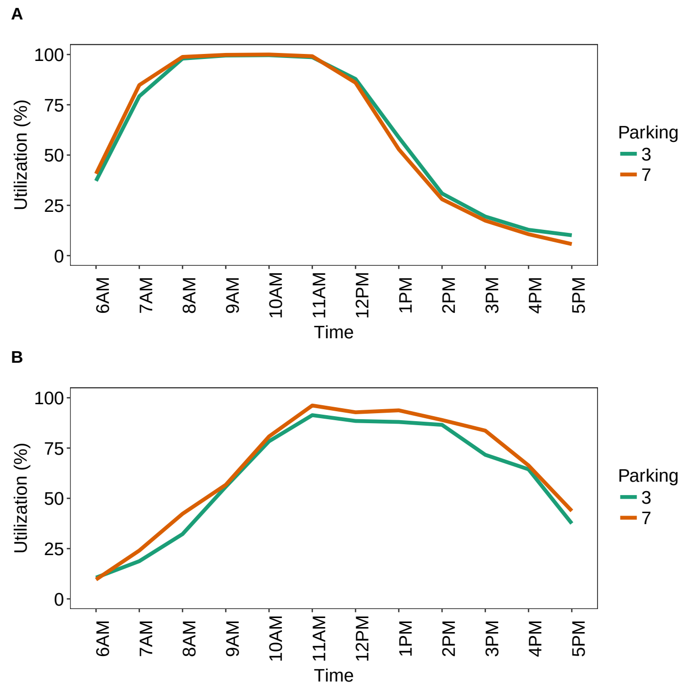

Figure 9: Average percentage of hourly utilization in A) weekdays and B) weekends in an optimistic case when $p = 2$ and (BEV,PHEV) market shares are (1%,2%).

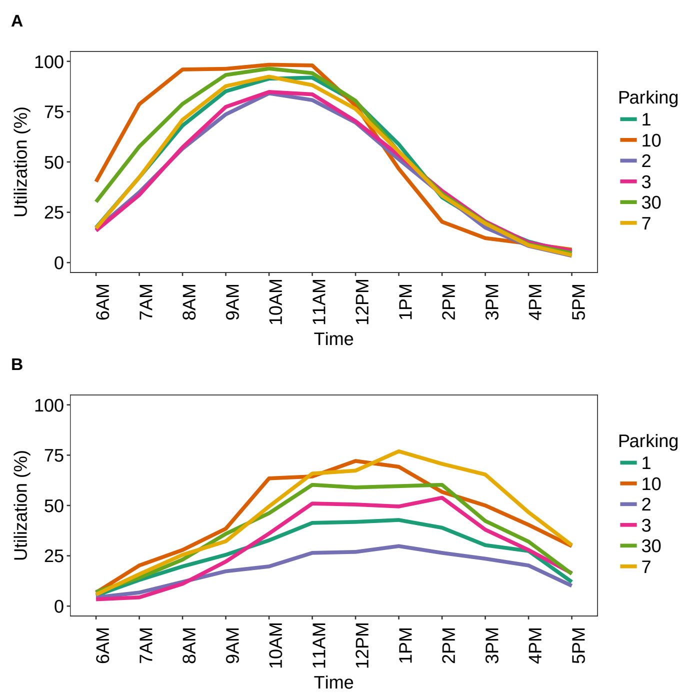

Figure 10: Average percentage of hourly utilization in A) weekdays and B) weekends in an optimistic case when $p = 6$ and (BEV,PHEV) market shares are (1%,2%).

biking, and improve their health condition [54]. It has been estimated that a shift from driving to walking can save the average approximately 25c per vehicle-mile traveled and 50c under urban-peak condition, when emission and parking costs are high, in external costs such as traffic congestion, noise and air pollution [46]. Design of an effective EV charging network can also provide opportunities for people in a community to increase their level of physical activity.

Figs. 11 and 12 compare total walking distance and walking distance per capita among people that have access to public EV charging service in the community before and after installing charging stations. As mentioned earlier, two cases are evaluated, an optimistic case where we assume that people will always choose the farthest available parking lot and a pessimistic case where people will always choose the nearest available parking lot for EV charging. These plots show that increasing number of charging stations in the community can raise total walking distance and walking distance per capita among people that have access to public EV charging stations up to 40% in an optimistic case. However, the rate of increase in total walking distance and walking distance per capita decreases as more charging stations are installed in the community. This happens as people get closer to the charging stations and their need to walk is reduced.

Another interesting aspect is the relationship between willingness to walk pattern and access to charging stations as young and old communities are expected to have a different level of willingness to walk. Young people tend to walk more while elderly people are not willing to walk long distances. Fig. 13 shows that if the average walking distance preference drops to half, accessibility to public EV charging stations will reduce by 4.23% and 1.32% when $p = 4$ and $p = 6$, respectively. However, if the average of willingness to walk distribution is doubled, accessibility increases by 2.86% and 2.43% when $p = 4$ and $p = 6$, respectively. This provides an additional perspective for policy makers, and also indicates the robustness of the model toward any change in willingness to walk pattern

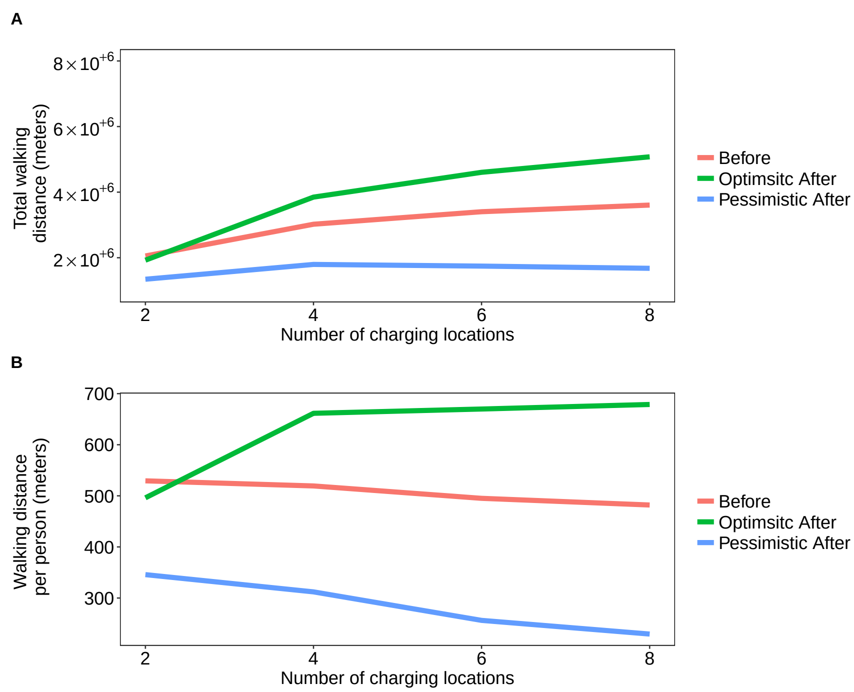

Figure 11: A) Total walking distance and B) walking distance per capita for people that have access to public EV charging service when (BEV,PHEV) market shares are (1%,2%).

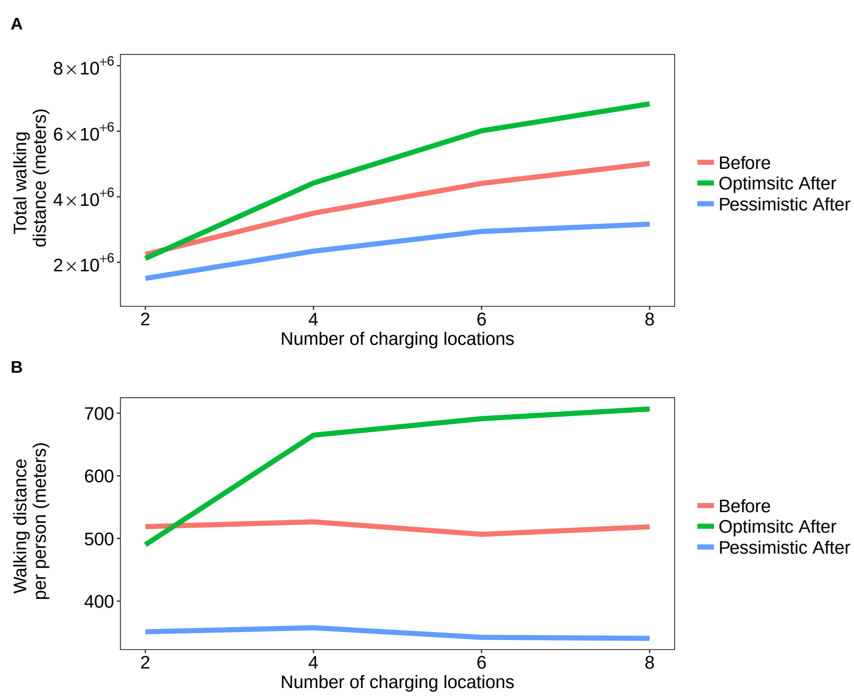

Figure 12: A) Total walking distance and B) walking distance per capita for people that have access to public EV charging service when (BEV,PHEV) market shares are (2%,3%).

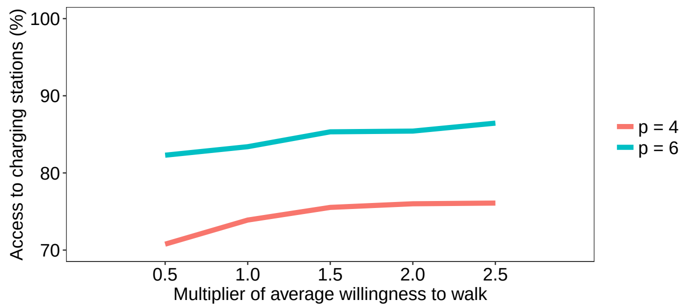

Figure 13: Accessibility for different average of willingness to walk distribution when (BEV,PHEV) market shares are (1%,2%).

in a community.

2.5.5 Value of stochastic solution

Value of stochastic solution was first introduced by [11], and is a standard means to quantify the usefulness of stochastic programming approach. Let the objective value of recourse problem be given as $RP = E_\Omega[\varphi(x,z,\omega)]$, and the *expected value problem* is obtained by replacing all random variables in scenarios with their expected values, $EV = \varphi(x,z,\bar{\omega})$, where $\bar{\omega}$ for the demand parameter will be $\sum_{\omega\in\Omega} p_\omega d_{\gamma(t),b,s}(\omega)$, p_ω representing probability of occurrence for a scenario ω, and $\sum_{\omega\in\Omega} p_\omega = 1$. Let $\bar{x}, \bar{z}$ represent solutions for EV problem, then the expected result of using expected value solution $(\bar{x},\bar{z})$, is given as $EEV = E_\Omega[\varphi(\bar{x},\bar{z},\omega)]$. Then, value of stochastic solution can be defined as $VSS = RP - EEV$. For obtaining VSS, we used the same number of scenarios as in SAA results. Based on five different runs, Fig. 14 shows that using stochastic programming brings up to 10.56% and 7.69% improvements in accessibility to public EV charging network when EV market share is 3% and 5%, respectively.

2.6 Conclusion

In this paper, we have presented a two-stage stochastic programming model for public EV charging station network design problem in a community. We considered several

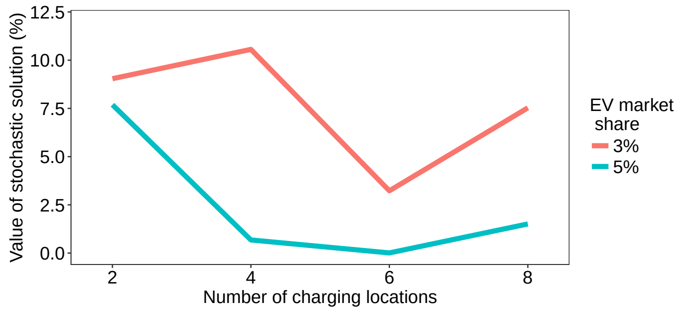

Figure 14: Median of value of stochastic solution for five different runs and different values of p and EV market share.

uncertainties such as total EV flows, arrival time, dwell time, battery SOC at the time of arrival, charging preference of EV drivers and willingness to walk patterns in estimating demand for public EV charging service. We used sample average approximation method, and for better computational performance, we proposed an effective heuristic that can solve large-scale problems and produce near optimal solutions. On a post analysis, our model presented a number of insights about the design of public EV charging network in an urban/community area. The results show that increasing number of charging stations in the community will improve accessibility to charging service for EV owners but will reduce the utilization level of these stations. Although all charging stations have similar demand patterns but increasing number of charging stations will increase the difference among stations in terms of utilization. While having more charging stations in the community can potentially increase total walking distance and walking distance per capita but the rate of increase in these measures decreases as we install more charging stations. Our model also shows robustness toward any change in willingness to walk pattern of community in the future. We suppose these analogies will provide better insights for a policy maker. Though we have used expected value function for the two-stage model, it will be interesting to see the use of risk-measures for these strategic decisions in the future.

2.7 Acknowledgment

This research was funded by the U.S. DoT through the Transportation Research Center for Livable Communities at Western Michigan University.

CHAPTER 3: MANAGING ACCESS TO PRIMARY CARE CLINICS

3.1 Introduction

Today, one of the most important challenges confronting healthcare facilities is the effective management of *access* for patients to outpatient care, especially primary care. American Academy of Family physicians defines primary care as "the care provided by physicians that are trained for comprehensive first contact and continuing care for patients with any undiagnosed sign, symptom, or health concern". There is strong evidence that patients in the U.S. often experience long waiting times to get appointments with their primary care providers. The average appointment delay to see a family physician ranged from a relatively healthy average of 5 days in Dallas to 66 days in Boston during 2014 [3]. Figure 15 reports aggregate access to care for returning patients from all primary care clinics affiliated with each VA medical facility across the U.S. for the calendar year 2013 as a function of realized appointment slot utilization. Slot utilization is defined here as the percentage of provider appointment slots actually used for providing care as a function of total number of available slots. Each data point in Figure 15 corresponds to a single facility and reports annual average access to patients served by all the primary care clinics affiliated with the facility. At the time, VA defined access as a binary measure based on whether or not a returning patient has been provided an appointment within 14 days of the patient's 'desired' appointment date (for new patients, the time window starts with call date). This plot clearly shows that there is a wide variation in access among facilities. This is in spite of the fact that the reported measure is an aggregate measure across all the primary care clinics of the facility, with some facilities carrying 20 to 25 primary care clinics. At the same time, the slot utilization is less than 60% for vast majority of the facilities. The plot clearly suggests that access is not poor due to high slot utilization. Further investigation of these facilities and clinics by our team revealed that the reasons for poor access can be attributed to poor and inconsistent appointment scheduling practices, high patient "no show" rates, appointment cancellations by patients/clinics etc.

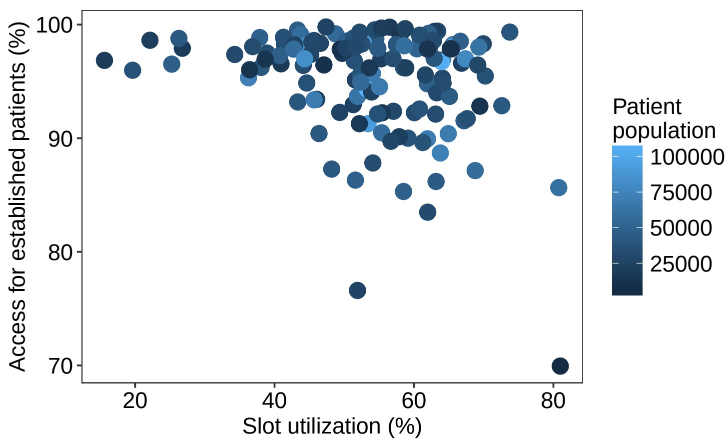

Figure 15: Access to primary care for returning patients vs. slot utilization at VA facilities nationwide in 2013. Color-coding refers to number of primary care patients cared for by each facility.

Access to primary care is expected to improve patient health outcomes, reduce overall healthcare costs, and increase health equality between population groups [2]. Analysis of facility level data from 89 VA medical centers merged with patient level data from geriatric outpatient clinics by Prentice et al. [61] revealed that long access delay has a significant impact on negative health outcomes such as mortality. Fahmy et al. [26] also found that experiencing delay in getting treatment is associated with adverse health outcomes in patients with bladder cancer. Appointment delay can also lead to patient attrition and lost opportunity for effective treatment in patients with mental health condition [60].

On the other hand, access to care, quality of care and health service efficiency are interrelated as dimensions of healthcare system performance [1]. One approach to improving the quality of health delivery process is to use patient flow analysis [24]. Patient flow analysis typically involves measuring patient waiting times within the clinic as well as utilization for staff and other resources. The amount of time a patient waits to get an appointment is generally referred to as 'indirect' waiting time and the amount of waiting time experienced by the patient on the day of the appointment within the clinic to receive care from the provider is referred to as 'direct' waiting time [33]. To some extent, one can improve indirect waiting time by compromising direct waiting time (and vice versa) by increasing slot utilization, employing over booking strategies (where multiple patients are assigned to same appointment slots), and reducing appointment slot lengths.

Study by Wellstood et al. [76] confirms that patient waiting time in primary care clinic is the most important barrier in access to care for different groups of patients. A recent survey by Software Advice shows that more than 40% of patients are willing to visit another physician to experience shorter wait times in the clinic. This study also shows that while 45% of the patients experience less than 15 minutes of direct waiting time for the provider, some 15% of the patients are experiencing more than 30 minutes of direct waiting time in the clinic [49]. Another study by Anderson et al. [6] shows that around 25% of patients are experiencing more than 30 minutes of direct waiting time for primary care. Their results demonstrate that longer direct waiting times were associated with lower patient satisfaction but it is moderated by service time with provider.

Institute of Medicine considers mismatched supply and demand as one of the causes of delay in access to healthcare [38]. While demand for healthcare is expected to increase by 29% from 2005 to 2025 due to population growth and aging, the number of adult primary care practitioners is estimated to grow only between 2% and 7% during the same period [12]. Balancing supply and demand in healthcare environment is usually done through appointment scheduling system. Appointment scheduling has been discussed extensively in the literature. However, prior research on outpatient appointment scheduling has mostly focused on managing patient flow inside the clinic through minimizing patient direct waiting time and provider idle time and/or overtime. Very few articles studied the indirect waiting time that patients experience in getting an appointment with their primary care provider [47]. Moreover, while most articles in the literature assume that patients call on the day that they want an appointment, data shows that many patients call well in advance to make an appointment. These patients are called 'routine' as opposed to patients that ask for 'urgent' or same day appointments. Figure 16 shows how soon patients call and ask for an appointment within three different VA primary care clinics in the U.S. Midwest. Data from VA primary care clinics also shows that number of patients that call during any particular day is stochastic in nature. Factors such as stochastic nurse and provider service time by patient type, patient preference toward day and time of the appointment, patient (un)punctuality, patient no-shows, and appointment cancellations (by patient or clinic) bring about more uncertainties into appointment scheduling problem. Ignoring these uncertainties will result in scheduling policies that are sub-optimal or infeasible in real clinical settings.

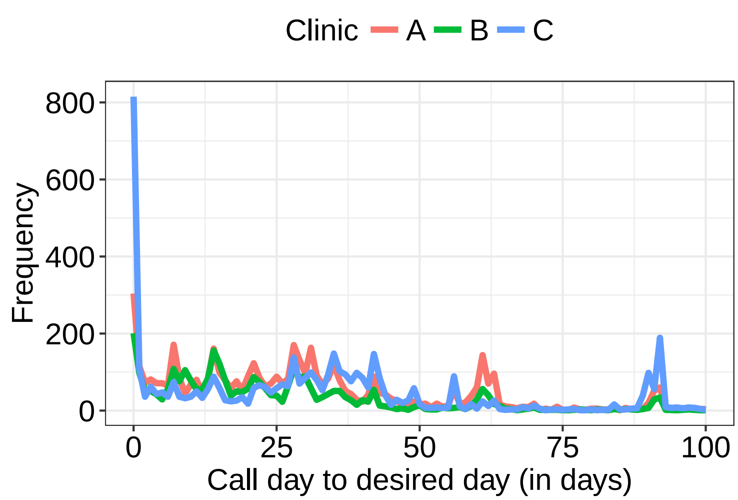

Figure 16: Time (in days) between call date and desired date for patients in three different VA primary care clinics. Clinic names are coded.

The main goal of this research is to address the gap between appointment scheduling and patient flow in primary care clinics to jointly improve both direct and indirect waiting times for patients. We propose a two-stage stochastic programming model that incorporates uncertainties in demand volume, patient call date, patient desired date, and patient no-show in appointment scheduling to minimize indirect waiting time of patients while accounting for clinic patient flow objectives. In particular, the goal is to produce optimal appointment scheduling 'templates' to be shared with the "call center" for scheduling of patients. In most healthcare facilities, it is the call center that is responsible for taking calls from patients seeking appointments and actual appointment scheduling. The templates can be for a day or a week or a month and so on. The call center is expected to follow the guidance provided by the template in providing actual appointments. For example, the template might suggest patients seeking annual 'physical' appointments be given appointments between 8-10AM on any day of the week.

The overall process is iterative. First, the two-stage stochastic programming model employs a rolling planning/booking horizon and yields a scheduling template using the supplied input parameters regarding supply and demand. As noted earlier, this template determines the allocation of arising demand into different days and appointment slots based on "patient or appointment types" and resource availability during the booking horizon. The performance of the template in terms of clinic patient flow is evaluated through simulation (termed "short-term simulated feedback"; simulations here assume that the scheduling template is fully populated with appointments). If the patient flow metrics are not satisfactory, additional constraints are added to the optimization model to avoid certain (sub-)sequences within the optimal template and we re-optimize the model. The process is repeated until patient flow metrics are satisfied and the resulting template is passed on to the call center for patient scheduling.

Over time, given the non-stationary nature of primary care (due to seasonal factors and others), demand and supply processes can change. If so, these input parameters (e.g., appointment request call volumes, mix of appointment types etc) are updated at regular

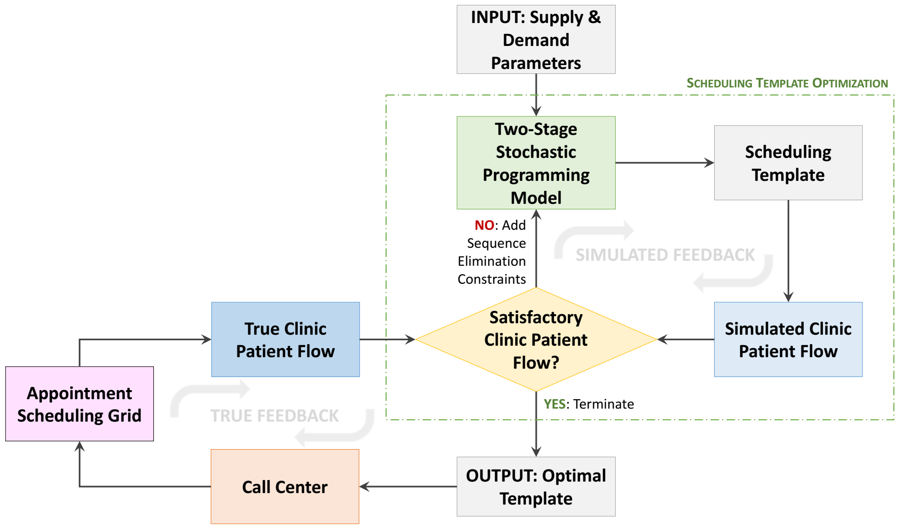

Figure 17: Appointment scheduling template optimization process

intervals or as necessary and the scheduling template is re-optimized. In addition, the feedback from the simulated clinic patient flow may not be fully representative of the true patient flow performance being achieved within the clinic. For this reason, the true performance feedback from the clinic is also fed back (termed "true feedback") to the optimization model for improved performance over time. Figure 17 gives an overview of this proposed approach to optimize the appointment scheduling template.

Main contributions of this study are as follows: 1) We propose a two-stage stochastic programming approach to develop scheduling policies/templates for patients in primary care by minimizing their indirect waiting time; 2) we employ simulation modeling to manage patient flow in clinic by introducing sequencing rules that control patient waiting time and provider idle time and/or overtime in clinic; 3) An index policy is proposed for appointment scheduling in call center considering several factors such as patient preference for date and time of the appointment, and patient and clinic appointment cancellation. The proposed method is validated using data informed by the literature, VA clinics, as well as synthetic experiments.

3.2 Literature review

There is a rich body of health care operations management literature on outpatient appointment scheduling and several methodologies such as operations research, queuing theory and simulation have been used to solve these problems. However, most of prior research has focused on proposing appointment scheduling systems to manage patient flow in clinic, and very few researchers addressed the problem of indirect waiting time of patients.

3.2.1 Clinic patient flow

Vast majority of articles in the literature have focused on minimizing patient direct waiting time, provider idle time and/or overtime work in clinic.

Patient flow measures: Muthuraman et al. [56] proposed a stochastic overbooking model to optimize appointment scheduling in an outpatient clinic where patients have different probability of not showing up for their appointments based on their history. Their objective function captures patient waiting time, provider overtime and idle time. Zeng et al. [81] formulated appointment scheduling to maximize clinic's expected profit, consisting of revenue from patients as well as cost of patient waiting and physician overtime and idle time, when patients have different no-show probabilities. They observed that performance of scheduling practices using homogeneous overbooking models by the mean value of show-up probabilities is not good enough. Chakraborty et al. [21] developed a sequential scheduling algorithm to minimize total expected cost resulting from patient waiting time and physician out-of-regular hour work. In their study, service time of each patient is randomly distributed and every scheduled patient has a chance of not showing up for the appointment. They showed that their model leads to higher profits and less overtime than the policies considering service periods pre-divided into slots.

No-show and overbooking: A major challenge that outpatient facilities are dealing with is the patient no-show problem. To mitigate the negative impact of no-shows on scheduling practice, Laganga et al. [44] developed an appointment scheduling approach considering overbooking to make a balance between patient waiting time and clinic overtime. They concluded that it is impossible to draw general conclusions about constructing overbooking schedules. Zacharias et al. [80] proposed an overbooking model to mitigate the negative impact of patient no-show on clinic performance when patients have different no-show probabilities. They studied static and dynamic scheduling problems and showed that no-show rate and patients' heterogeneity in terms of not showing up have a great negative impact on the scheduling process.

Patient choice: Among papers that studied appointment scheduling in outpatient clinics, some have considered patient preferences in their modeling. Qu et al. [63] proposed a Markov chain model to optimize appointment scheduling while considering patient choice. They assume that patients independently request appointments and when their desired clinic is fully booked, they independently decide to make an appointment later or seek care in other clinics. They also examined the effect of patient choice on physician productivity. Yan et al. [78] proposed sequential appointment scheduling policies to balance clinic efficiency and patient satisfaction considering patient choice and service fairness. Their objective is to maximize clinic profit that consists revenues from patients and costs of patient waiting time and provider idle time and overtime.

What arises from all these studies is that they only considered patient flow in outpatient clinic. However in real applications, provider has to make a decision on the percentage of appointments in every day that is assigned to each patient type considering the fact that many patients call and ask for an appointment well in advance while some other ask for same-day appointments.

3.2.2 Indirect waiting time

On the other hand, very few researchers have studied the patient indirect waiting time problem in outpatient clinics.

Advanced access scheduling: Today, an increasing number of clinics are using advanced access or open-access system to reduce the indirect waiting time in primary care. The idea behind this policy is doing today's work today, which means that patients are given appointments in or near the day in which they want an appointment. Taking into account that patients may cancel their appointment or do not show up for their visit, Liu et al. [48] proposed a dynamic scheduling policy in an outpatient clinic. Their finding shows that open access scheduling policy is performing better when demand rate is relatively low. Dobson et al. [25] examined the effect of keeping some slots open for same-day demand on two quality measures in primary care clinics: average number of same-day demand that is not served during normal working hours and average number of non-urgent patients in the queue. They demonstrated that encouraging non-urgent patients to call for same-day appointment is an important factor in implementing advanced-access scheduling system in primary care facilities. Qu et al. [62] demonstrated the percentage selection for open appointments in an open access appointment system by using a mean-variance approach. Their results indicate that when both demand rate and no-show rate are high for appointments reserved for routine patients, there are one or more Pareto optimal percentages of open appointments that decrease the variability in the number of patients seen. Lee et al. [45] compared open-access scheduling policy with overbooking methods in primary care facilities in terms of staff overtime, patient waiting time, proportion of unmet demand and capacity utilization. They concluded that although it has been reported that open-access is performing well when supply and demand are in balance, overbooking performs even better. In addition, if same-day demand is high, over 80%, overbooking outperforms open-access approach.

Patient choice: Patient choice is also important from indirect waiting time point of view. Rubin et al. [64] investigated patient preference when making an appointment with a general practitioner to describe the relationship between getting an appointment sooner and choice of time. They found that speed of access is of limited importance compared to patient choice of appointment time and this is much true for patients who are employed. Gupta et al. [34] developed a Markov decision process to manage access to appointment slots when patients have different choices between accepting a same-day appointment and a future one. They provided optimal solutions in two cases in which clinic is either a single-physician or multiple-physician one. Wang et al. [74] studied appointment scheduling to optimize clinic revenue by finding the optimal balance between number of slots that should be remained for same-day demand and number of slots that can be filled by routine patients considering patient preferences toward physician of choice and time of the appointment. Their model is limited because they did the scheduling only for one day and did not consider the interactions between multiple scheduling days.

Our work is more close to the research of Luo et al. [50] in which they developed a Tandem Queue model to study the relationship between appointment queue and service queue. They obtained various system performance measures such as server utilization and long-run average appointment delay and service delay. The main research question that we are trying to address in this study is that how can primary care practices schedule patients so that patients experience minimum delay in getting an appointment while patient flow in clinic is as smooth as possible? Our work is different from above studies in several important ways. First, we take into account indirect waiting time of patients that may call in advance to book an appointment or ask for same-day appointment over planning horizon T. Second, some patient flow measures such as patient direct waiting time, provider overtime and amount of lunch time that provider spends with patients resulted from the optimal sequence are being analyzed and considered in the optimization model. The other distinguishing feature of our study is that we explicitly consider patient no-show rate and use overbooking to mitigate its negative impact on indirect waiting time of patients while accounting for workload of primary care provider in terms of patient complexity in every appointment slot and every scheduling session. We also include patient choice toward appointment date and time in our call center simulation.

3.3 Problem description

A well-established appointment scheduling system will have direct positive effect on patient satisfaction. In this research, without loss of generality, we study a primary care outpatient clinic that is being managed by a single provider and capacity over the planning horizon is predetermined. In this clinic, there are 8 appointment slots with equal length of 60 minutes every day and 40 different sessions of length 4 hours in each month, with 10 sessions per week (Monday through Friday, morning and afternoon). We focus on single-day templates and the method can be readily extended to weekly and other template settings. Patients may call in advance to book an appointment or ask for a same-day appointment. Patients are only scheduled with their own primary care provider (PCP) so that continuity of care, rate of patients within a panel that visit their own PCP, is ensured. We assume that the available provider time could be divided into different sets for different types of patients. Providing appropriate ratios of appointments to different patient types will ensure fairness among patients in terms of access to care.

Clinic might cancel appointments due to the lab result delay or absence of the provider. However, these cancellations should be managed since it will increase patient dissatisfaction, and staff workload in the future. Patients may also cancel their appointment or simply not show up for their visit. Overbooking is a means to mitigate the negative effect of patient no-show on provider's slot utilization. In this way, provider becomes certain that there is always a patient to visit. However, scheduling system should consider the available resources, from exam rooms to number of staff, in putting several patients in one appointment since overbooking several patients can increase patient direct waiting time.

In this study, we assume that patient's request cannot be denied. Otherwise, patient will seek care in specialty care or emergency department that are more costly than primary care. Since the excess capacity that providers can add to the available capacity through overtime is limited, our model needs to moderate the effect of capacity shortage due to competing responsibilities of the provider, holidays and emergency closures on patient indirect waiting time. Complexity of our model also stems from the fact that patients have different preferences toward day of the appointment. There is always uncertainty regarding number of patients that call every day in the planning horizon and their desired date.

Our goal is to minimize indirect waiting time of patients by considering all available days in booking horizon of length T while accounting for patient flow in clinic and provider workload in every appointment slot and every scheduling session in terms of patient complexity. There is an infinite time horizon, but a finite rolling scheduling horizon. We represent this problem as a two-stage stochastic programming where first-stage decision variables are determined before realization of uncertainties and second-stage decision variables are determined after presenting scenarios that are representing uncertainties.

3.4 Model formulation

To formulate the access to care problem, we introduce a two-stage stochastic programming model. We will consider uncertainties in demand volume, patient call date, patient desired date and patient no-show rate in our model. First-stage decisions are made on the number of different patient types that can be scheduled in each appointment slot based on maximum tolerable patient complexity by provider in each appointment slot and each scheduling session, and these decisions are exposed to uncertainties in the second-stage. In the second-stage, patient allocation decisions are made in order to minimize total indirect waiting time of patients to visit the primary care provider. The output of two-stage stochastic programming is a daily scheduling template for the booking horizon. Table 7 shows the notations that are used for the two-stage stochastic programming model.

Table 7: Model Notation

Symbol	Description
Sets:	
R	Set of patient types, indexed by $r \in R$
A	Set of appointment slots, indexed by $a \in A$
S	Set of sessions, indexed by $s \in S$
G	Set of sequences in which patient flow requirements are not met, indexed by $g \in G$
D	Set of days in the booking horizon, indexed by $d \in D$
L_r	Set of number of patient type r that can be scheduled in every appointment slot, indexed by $l \in L_r$
Ω	Set of scenarios, indexed by $\omega \in \Omega$
Model Parameters:	
c_r	Average complexity of patient type r
κ	Maximum acceptable patient complexity for each appointment slot
η	Maximum acceptable patient complexity for each scheduling session
p_r	Average no-show probability of patient type r
$m_{r,l}$:	Number of patients of type r, $l \in L_r$
$\xi_{r,a}$	Number of scheduled patients of type r in appointment slot a
$f_{r,d}(\omega)$	Number of patients of type r that asked for an appointment in day d in scenario ω
First-stage Variables:	
$x_{r,a}$	Number of patients of type r that can be scheduled in appointment slot a
$z_{r,a,l}$	1 if l patients of type r can be scheduled in appointment slot a; 0 otherwise;
Second-stage Variables:	
$y_{r,d,d'}(\omega)$	Percentage of patients of type r that asked for an appointment in day d and are scheduled in day d' in scenario ω

First-stage problem can be presented as follows:

$$\text{Min } f(x,z) = E[\varphi(x,z,\widetilde{\omega})] \tag{20}$$

$$\text{s.t. } \sum_{r\in R} c_r x_{r,a} \leq \kappa \ \ \forall a \in A \tag{21}$$

$$\sum_{r\in R}\sum_{a\in s} c_r x_{r,a} \leq \eta \ \forall s \in S \tag{22}$$

$$x_{r,a} \geq \xi_{r,a} \ \forall r \in R, a \in A \tag{23}$$

$$x_{r,a} = \sum_{l\in L_r} m_{r,l} z_{r,a,l} \ \forall r \in R, a \in A \tag{24}$$

$$\sum_{l\in L_r} z_{r,a,l} = 1 \ \forall r \in R, a \in A \tag{25}$$

$$\sum_{r,a,l\in g} z_{r,a,l} \leq |g| - 1 \ \forall g \in G \tag{26}$$

$$x_{r,a} \in \mathbb{Z}^+ \ \forall r \in R, a \in A, z_{r,a,l} \in \{0,1\} \ \forall r \in R, a \in A, l \in L_r \tag{27}$$

where $\varphi(x,z,\widetilde{\omega})$ is the solution of the following second-stage problem:

$$\text{Min } \varphi(x,z,w) = \sum_{r\in R}\sum_{d\in D}\sum_{d'\in D:d\leq d'} w_r y_{r,d,d'}(\omega) f_{r,d}(\omega)[(d'-d)^{(1+\epsilon)}] \tag{28}$$

$$\text{s.t. } \sum_{d\in D:d\leq d'} (1-p_r) y_{r,d,d'}(\omega) f_{r,d}(\omega) \leq \sum_{a\in d'} x_{r,a}(\omega) \ \ \forall r \in R, d' \in D \tag{29}$$

$$\sum_{d'\in D:d\leq d'} y_{r,d,d'}(\omega) = 1 \ \ \forall r \in R, d \in D \tag{30}$$

$$y_{r,d,d'}(\omega) \geq 0 \ \forall r \in R, d, d' \in D : d \leq d' \tag{31}$$

The objective function minimizes expected delay of patients to get an appointment with their own primary care provider. We define the difference between desired date and appointment date as a super-linear to consider fairness in assigning delays to different patients. First-stage decision variable $x_{r,a}$ determines the number of patients of type r that can be scheduled in appointment slot a. Second-stage decision variable $y_{r,d,d'}$ assigns patients of type r that ask for an appointment in day d to day d'.

Every primary care provider has a certain threshold in terms of patient complexity that can be handled in every appointment slot and every scheduling session. Constraints (21) and (22) make sure that total complexity of patients that are scheduled in every appointment slot and every session does not exceed the maximum patient complexity that provider can handle in an appointment slot and a session, respectively. Constraints (23) aims at keeping the template according to the number of each patient type that has already been scheduled in every appointment slot. Constraints (24), (25) and (26) are sequencing rule constraints that are added dynamically to the optimization model once clinic patient flow simulation finds a certain sequence that does not meet patient flow requirements. Constraints (27) are non-negativity constraints. Constraints (29) make sure that there is enough capacity in each day in the booking horizon for every patient type and constraints (30) confirm that no patient request is denied.

3.4.1 Clinic patient flow simulation

The proposed solution for appointment scheduling may have limited success in reality if we only emphasize indirect waiting time and not considering patient flow in clinic. Simulation helps to include more complexities into the model. Patient direct waiting time, amount of service spillover to provider lunch time, and provider overtime at the end of the day are measured by patient flow simulation inside the primary care clinic (as illustrated in Figure 17). The model assumes that nurse and provider are ready before start time of the day. Without loss of generality, we model the patient flow inside the clinic as consisting of two stages: time with the nurse and time with the provider. Once patient walks into the clinic, he/she will wait in the lobby for the nurse to become available. After being visited by the nurse, patient waits in the exam room to meet the provider once the provider is available. We assume that patients that are scheduled for a particular day have to be served by the end of that day even if the provider has to work overtime to serve all patients.

Patients are assumed not to leave the exam room until provider finishes all required tasks.

Patients may arrive late for their appointments. By arriving late to the clinic, patients may increase provider idle time and cause waiting time for succeeding patients to increase. We assume that patients are called in the order of their arrival time. If patient arrives on time, waiting time to visit the nurse is measured from actual scheduled time but if the patient arrives late, waiting time to see the nurse would be measured from patient's arrival time.

3.4.2 Call center simulation

In the absence of an opportunity to implement the proposed method at primary care clinics, we rely on simulated experiments to test the performance of the proposed method. We simulate the call center where patients call and ask for an appointment or cancel their current scheduled appointment. The scheduler might also cancel an appointment on behalf of the clinic.

We propose an index policy to simulate patient scheduling in call center. When a patient requests an appointment, scheduler finds patient's desired date and calculates the following index for each appointment slot from patient's desired date on:

$$I_j = s_j \tag{32}$$

where s_j is the remaining capacity in appointment slot j. Then, scheduler ranks appointment slots based on their index and starts offering patient an appointment until patient accepts one. We simulate this process by generating a random number from $U(0, 1)$ and comparing it to an acceptance threshold. If the random number is higher than this threshold, patient accepts the corresponding appointment slot. We repeat this process until patient accepts an appointment. Once the appointment is scheduled, it will not be changed in the future.

Appointment cancellations are also handled in call center simulation. When a patient is scheduled, a random number is generated from $U(0,1)$ and if the random number is less than the cancellation rate of the clinic, patient or clinic will cancel the appointment before the appointment date. To determine the cancellation date, we generate a random number between patient's call date and actual appointment date. Once the cancellation date arrives, appointment becomes canceled and patient is removed from the scheduling grid.

3.4.3 Practical considerations

Unlike the literature, we used a better patient categorization process that represents effectively the amount of time that provider needs to spend with each patient type. Since provider service time follows a random distribution, we used average patient complexity in order to schedule patients in an appointment slot based on maximum patient complexity that provider can handle in one appointment slot and a scheduling session. In this way, we make sure that complex visits are evenly distributed across the day (or by limiting number of patients assigned to a slot) and provider has a balanced workload in every appointment slot and every scheduling session. We also considered a reserved buffer in each scheduling session that could be used for administrative tasks or accommodating urgent or walk-in visits.

In two-stage stochastic programming, patients could be assigned to appointment slots or appointment days in the second-stage. Assigning patients into appointment slots has this deficiency that many demand parameters become one patient and having a continuous assignment variable in the second-stage does not make sense. Since second-stage of two-stage stochastic programming is more of a tactical planning, we combined demand parameters for days so that we do not have continuous assignment problem anymore.

In clinic patient flow simulation, we realized that if patients that are assigned one appointment are scheduled to come at the same time, patient waiting time will artificially become higher. Therefore, we staggered visits by breaking each one hour slot into three 20 minute intervals (assuming 3 patients will be assigned to the slot). In addition, there are two different policies according which patients can be called in clinic. One policy is calling patients based on their scheduled time and the other one is calling patients based on their arrival time. Since simulation did not show any significant difference between these two policies, we assume that patients are being called based on their arrival time for visit. Another important factor that has not been studied in the literature is that some providers may ask for scheduling patients with lower complexity before the break between two sessions so that they could have enough time for lunch. So in addition to common patient flow metrics, we considered the amount of time that provider spends with patients during lunch break as a patient flow indicator.

3.5 Performance measures

We study a wide range of performance measures in order to show the impact of our appointment scheduling approach on patient and provider experience.

3.5.1 Indirect waiting time measures

The main performance measures for patient indirect waiting time are indirect waiting time distribution for all patients and each patient type and standard deviation of indirect waiting time for all patients and each patient type. We report all quartiles of indirect waiting time distribution in order to get a better picture of how much delay patients experience to get an appointment with their provider. As suggested by Cayirli et al. [20], we use the standard deviation of indirect waiting time as the fairness measure among patients.

3.5.2 Patient flow measures

In order to represent the relationship between appointment scheduling and patient flow inside the clinic, we use three common performance measures. One performance measure for patient flow in clinic is patient direct waiting time. Patient direct waiting time includes waiting time in the lobby to see the nurse and waiting time in the exam room to visit the provider. Other performance measures are service spillover to provider lunch time and overtime work of the provider. Since we are using overbooking to both deal with patient no-show problem and to provide service to more patients in a day, provider may need to work overtime during lunch time and/or after regular hours to serve all scheduled patients.

3.6 Numerical study

We performed a computational study to evaluate the performance of our proposed two-stage stochastic programming model in patient appointment scheduling. Because many factors play major role in real appointment scheduling setting, we chose a reasonable-sized problem in order to provide insight for clinic operations and process improvement. The parameters for our base problem are represented in Table 8.

Table 8: Base problem parameters and settings

Maximum patient complexity that provider can handle in an appointment slot = 1
Maximum patient complexity that provider can handle in a session = 3.28
Threshold for patient acceptance of offered appointment slot = 0.2
Average no-show rate for all patient types = 18%
Average cancellation rate for all patient types = 16%
Threshold for patient direct waiting time = 30 minutes
Threshold for spillover amount to provider lunch time = 30 minutes
Threshold for provider overtime = 45 minutes
Percentile of patient flow metric distributions in clinic patient flow simulation = 80%
Number of weekly requests = Triangular distribution with parameters (70,85,100)
Patient arriving time distribution = $\mathcal{N}(-16.62, 27)$
Booking horizon = 40 days
Planning horizon = 220 days

We used real data from a VA primary care clinic to estimate the number of weekly requests, patient no-show probability, appointment cancellation probability, patient call day (Monday through Friday) distribution, patient desired day (Monday through Friday) distribution and time between call date and desired date distribution.

By analyzing the National Ambulatory Medical Care Survey (NAMCS) for 2003, Yarnall et al. [42] determined the distribution of time that family physicians spend to provide service for different types of patients: acute, chronic and preventive. Table 9 reports service time with nurse and provider for each patient type along with the percentage of each patient type in provider panel. Average complexity of each patient type is computed by dividing average provider service time over the length of an appointment slot. We used the empirical study by Oh et al. [57] to proportionate the amount of time that nurse spends with patients of each type. As in Cayirli et al. [19], we assume that service time with nurse and provider for each patient type follows a log-normal distribution.

Table 9: Service time with nurse and primary care provider

Visit Type	(%) of Total Visits	Time with Nurse (minutes)	Time with Provider (minutes)
Acute	49.3	11.3 (8.3)	17.3 (8.7)
Chronic	36.1	12.6 (8.8)	19.3 (9.2)
Preventive	14.6	13.9 (11.3)	21.4 (11.8)

Patient unpunctuality for appointment is prevalent in outpatient clinics. By collecting data from a primary health care clinic in New York metropolitan hospital and using Kolmogorov-Smirnov test, Cayirli et al. [19] showed that a Gaussian distribution with a mean of -16.62 minutes and standard deviation 27.07 minutes fits the empirical distribution of patient arriving time to the clinic. Negative average means that patients on average arrive earlier than the starting time of their appointment.

In our proposed appointment scheduling process, we consider three patient flow metrics including patient direct waiting time, provider overtime at the end of the day and amount of spillover into provider's lunch hour and study their relationship with patient indirect waiting time. Practices have different standards for these performance measures. As considered in Michael et al. [55], 30 minutes of direct waiting time is an acceptable threshold for patients in the clinic. The thresholds for overtime and spillover are assumed to be 45 and 30 minutes, respectively. In clinic patient flow simulation, we find the distribution of patient direct waiting time, provider overtime and amount of spillover to provider lunch time and if a certain quantile of any of these distributions become larger than the specified threshold, we add the corresponding constraints of the patient sequence to the two-stage stochastic programming and re-optimize the model.

The computational study is in four parts. In the first part, our aim is to show the trade-off between indirect waiting time and patient flow in outpatient primary care clinic. The proposed model is solved considering various thresholds for patient flow measures. In the second part, we show how indirect waiting time changes if patients become more sensitive to appointment delay and their show up probability reduces as they wait longer to visit the provider. In the third part, we show the value of overbooking patients and how this policy can impact on both indirect waiting time and patient flow measures. The fourth part compares the cases where patients have different thresholds to accept appointment days and times that are being offered to them by the scheduler.

In order to show the effectiveness of the optimal sequence proposed by our appointment scheduling approach, we use two major sequencing rules that are proposed in the literature. The first rule is called SPT in which patients are scheduled in increasing order of mean service times. LCVB is the second rule in which patients are scheduled in increasing order of service time variability (low CV, σ/μ, in the beginning of the day) [10].

The model runs for 220 days and since the scheduling template is empty at the beginning, all performance measures will be reported between day 100 and day 180 to make sure that the system has reached a sufficient steady state. All computational studies were implemented using Gurobi 6.5 as the solver on a computer running Windows 7 with 2.6 GHz of processing speed and 80 GB of RAM.

Since sufficient number of scenarios to represent uncertainties in appointment scheduling problem is unknown, we use Sample Average Approximation (SAA) method to find this number. Because this problem is solved on a rolling horizon basis, we use SAA on various days to find the most reliable number of scenarios. Figure 18 compares the gap percentage between estimated upper bound and lower bound for three different days in the planning horizon. Since gap between upper and lower bounds is less than 5% for all three days when the number of scenarios are 30, we select 30 scenarios as sufficient number of scenarios for our analysis.

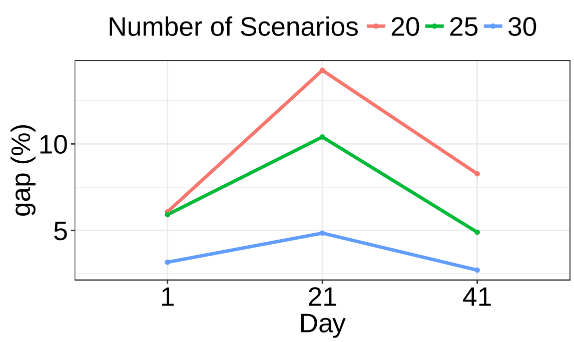

Figure 18: Gap percentage between estimated upper and lower bounds of optimal objective value in the three different days within planning horizon.

3.6.1 Trade-off between indirect waiting time and patient flow

By considering three different quantiles, α of 75, 80 and 85% for patient flow metric distributions, Figure 19 compares the indirect waiting time distribution of our scheduling approach with the same distribution of heuristic sequencing rules. The higher the quantile of patient flow performance measure distributions, the more the clinic manager is concerned about patient flow in clinic. We expect that as practitioner becomes more concerned about patient flow, patients experience longer delays to get an appointment. This hypothesis is confirmed by figure 19 in which indirect waiting time distribution skews to the left as α increases. This plot also shows the effectiveness of our scheduling approach compared to the heuristic policies in which patients have to wait longer to get an appointment.

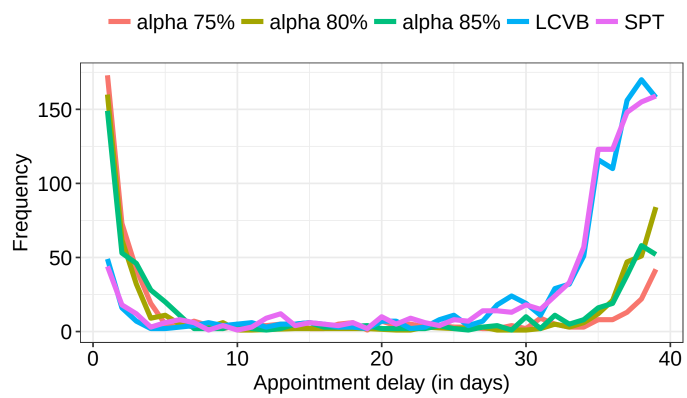

Figure 19: Indirect waiting time distribution for different quantiles of patient flow metrics. 62%, 60%, 58%, 21% and 20% of patients have no appointment delay with α of 75%, 80%, 85%, LCVB heuristic and SPT heuristic cases, respectively.

Table 10 represents the optimal daily templates under different values of α for patient flow metrics resulted from the two-stage stochastic programming and the heuristic scheduling templates. In this template, 'A', 'C' and 'P' stand for acute, chronic and preventive patient types, respectively.

Table 10: Optimal and heuristic daily templates

Slot	$\alpha = 75\%$	$\alpha = 80\%$	$\alpha = 85\%$	SPT	LCVB
1	A,C,P	A,C,P	A,C,P	A,A,A	C,C,C
2	A,C,P	A,C,P	C,P	A,A,A	C,C,C
3	A,C,P	A,C,P	A,A,C	A,A,A	A,C
4	A	P	A,C	C,A	A,A,A
5	A,C,P	A,A,C	A,C,P	C,C,C	A,A,A
6	A,C	A,A,C	A,C,P	C,C,C	A,A,A
7	A,A,C	A,C	A,C	P,P	P,P
8	A,C	A,C	A,A	P	P

Figures 20, 22 and 23 represent the patient flow metrics as level of concern about patient flow in clinic changes. Although difference between patient direct waiting time, amount of spillover to provider lunch time and provider overtime distributions is not significant for different values of α in Figures 20, 22 and 23, these distributions are very different from the ones resulted from heuristic scheduling approach. These plots show that patients will experience lower appointment delay and direct waiting time under our appointment scheduling approach. Amount of spillover to provider lunch time would be lower and provider needs less amount of time to work beyond regular hours to serve all patients in a day.

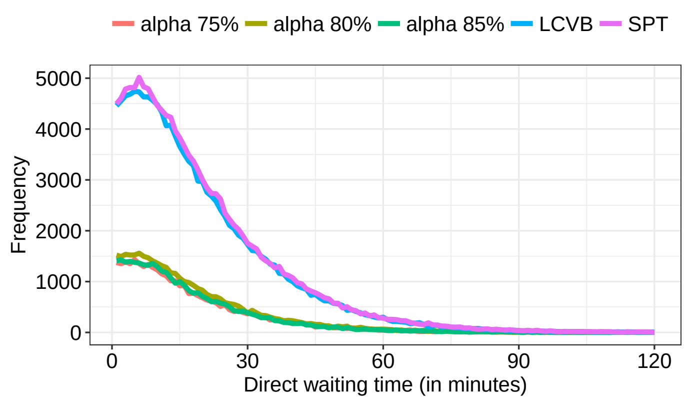

Figure 20: Direct waiting time distribution for different quantiles of patient flow metrics. 64%, 61%, 64%, 49% and 47% of patients do not wait in clinic in α of 75%, 80%, 85%, LCVB heuristic and SPT heuristic cases, respectively.

Another performance measure for appointment delay is the standard deviation of delay among patients. Standard deviation of appointment delay shows how fair the scheduling system treats patients in assigning appointment to them based on their desired date. Figure 21 shows that our appointment scheduling policy treats fairer than heuristics in assigning delay to patients.

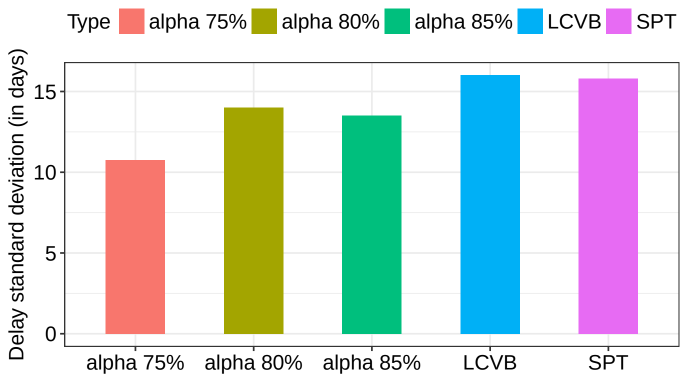

Figure 21: Standard deviation of indirect waiting time for different scheduling approaches.

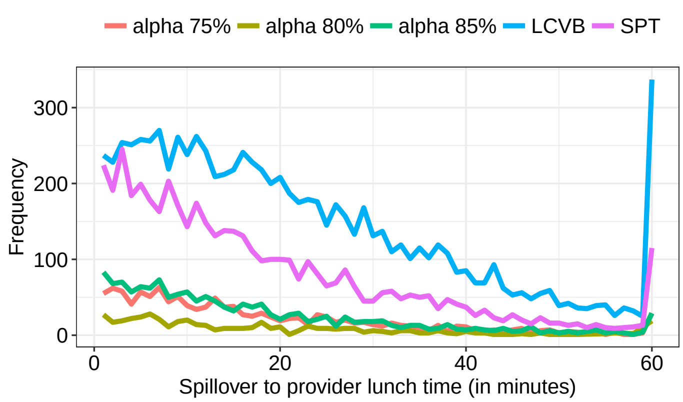

Figure 22: Amount of spillover to provider lunch hour distribution for different quantiles of patient flow metrics. In 84%, 92%, 82%, 46% and 70% of days, provider does not need to work during lunch time in α of 75%, 80%, 85%, LCVB heuristic and SPT heuristic cases, respectively.

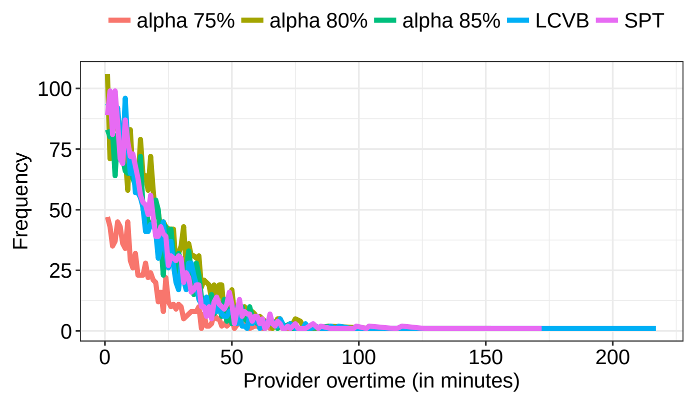

Figure 23: Provider overtime work distribution for different quantiles of patient flow metrics. In 89%, 70%, 74%, 87% and 87% of days, provider does not need to work overtime in α of 75%, 80%, 85%, LCVB heuristic and SPT heuristic cases, respectively.

3.6.2 No-show behavior

Researchers often assume that patient no-show probability remains constant and does not depend on patient appointment delay. However, evidence suggest that there is higher chance that patient does not show up if appointment delay becomes longer [28] [30].

Three different functions are proposed in the literature by Kopach et al. [41], Galluci et al. [30] and Green and Savin [32] to show the relationship between appointment delay and patient show-up probability:

$$p_j = \begin{cases} 1 - p * (1 - 0.5 * e^{-0.017j}) \\ 1 - (0.51 - 0.36 * e^{-j/9}) \\ 1 - (0.31 - 0.3 * e^{-j/50}) \end{cases}$$

where j represents appointment delay and p is the estimated patient no-show probability in the function proposed by Kopach et al. [41]. We assume that p is equal to the average no-show probability in our study. Figure 24 shows the sensitivity of patients to appointment delay under these functions. This plot clearly shows that patients are more sensitive to appointment delay under function proposed by Galluci et al. [30] and there is lower chance that they show up for their appointment. Figure 25 shows that as patients become more sensitive to appointment delay and their show-up probability reduces faster, the model tries to schedule patients near their desired date in order to lower their appointment delay.

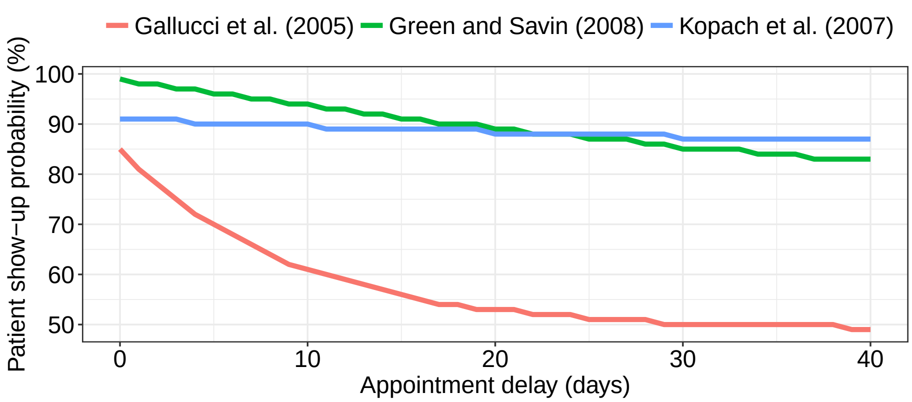

Figure 24: Patient show up probabilities

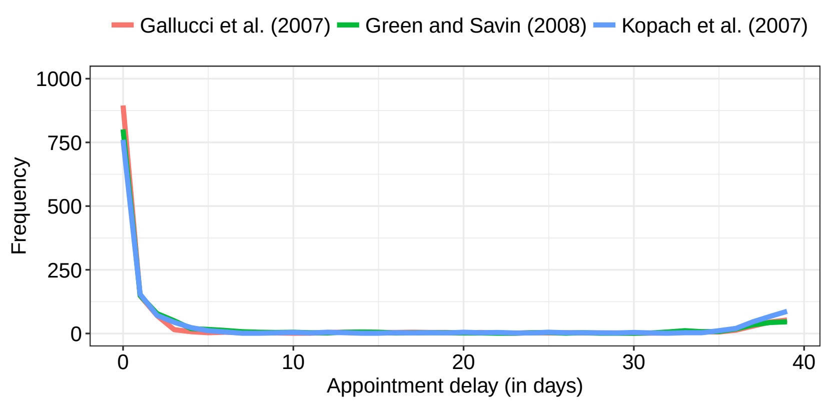

Figure 25: Indirect waiting time distribution under three different delay-dependent no-show function.

Table 11 compares the optimal scheduling templates under different behavior of patients to appointment delay.

Table 11: Optimal daily templates under different delay sensitivity of patient show-up probability

Slot	Kopach et al. [41]	Galluci et al. [30]	Green and Savin [32]
1	A,C,P	A,C,P	A,C,P
2	A,C	A,C,P	A,C,P
3	A,A,C	C,P	C,P
4	A,P	A,C	A,C
5	A,C,P	A,A,P	A,C
6	A,C	A,C	A,A,P
7	C,C,P	A,A,C	A,C
8	A,A	A,C	A,A,C

3.6.3 Value of overbooking

We have considered a maximum patient complexity that provider can handle in every appointment slot and every scheduling session. In this way, provider will have a balanced workload for the entire working day. Not using all provider capacity in every scheduling session will provide this opportunity to schedule some urgent walk-in patients between scheduled ones or do any administrative task during the idle time. On the other hand, scheduling more patients in a session will give this opportunity to patients to get an appointment sooner but it may result in higher patient direct waiting time, amount of spillover to provider lunch time and overtime work of the provider in clinic. Table 12 provides the optimal daily sequencing rules under different policies for overbooking patients in a session.

As Figure 26 shows, scheduling more complex patients in a session, even when the maximum slot complexity remains constant, can have significant effect on patient appointment delay. Figures 27 and 28 show that this increase in session maximum complexity will not change patient direct waiting time and spillover amount to provider lunch time but according to Figure 29, it will cost the provider to work more after regular hours to serve all scheduled patients.

Table 12: Optimal daily templates for different values of slot and session maximum complexity

Slot	(Slot = 1, Session = 2.7)	(Slot = 1, Session = 3)	(Slot = 1, Session = 3.28)
1	C,P	A,A,C	A,C,P
2	C,P	A,A,A	A,C,P
3	A,P	A,P	A,C,P
4	A,C	A,A	P
5	A,A,C	C,P	A,A,C
6	A,A	C,C,P	A,A,C
7	A,C	C,C	A,C
8	A,C	C,P	A,C

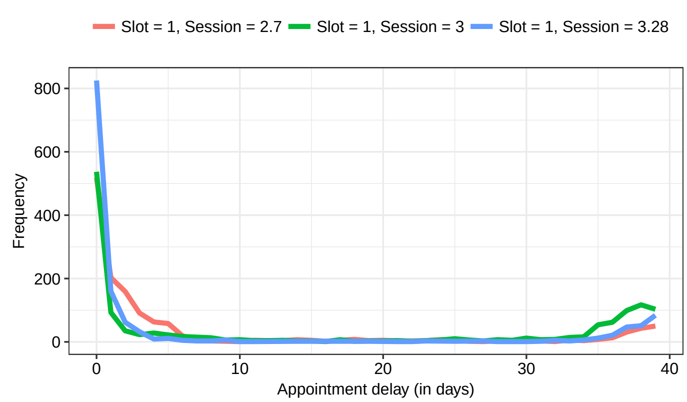

Figure 26: Indirect waiting time distribution for different values of maximum slot and session complexity.

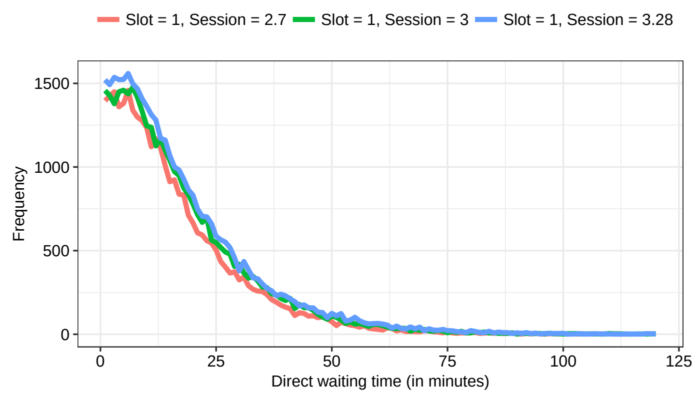

Figure 27: Direct waiting time distribution for different values of maximum slot and session complexity. 67%, 62% and 61% of patients do not wait in clinic in (1,2.7), (1,3) and (1,3.28) of appointment slot and scheduling session complexities, respectively.

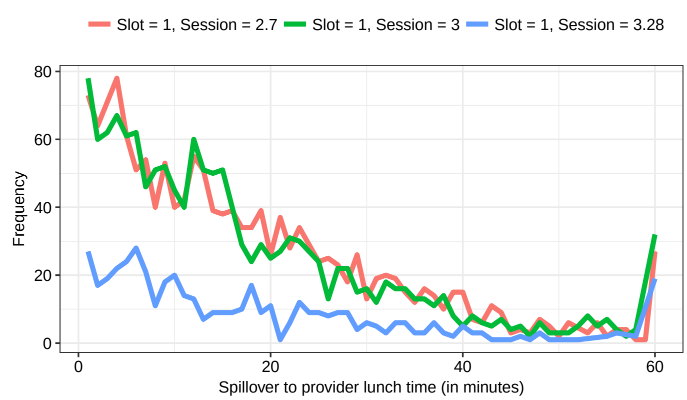

Figure 28: Amount of spillover to provider lunch hour distribution for different values of maximum slot and session complexity. In 81%, 82% and 94% days, provider does not need to work during lunch time in (1,2.7), (1,3) and (1,3.28) of appointment slot and scheduling session complexities, respectively.

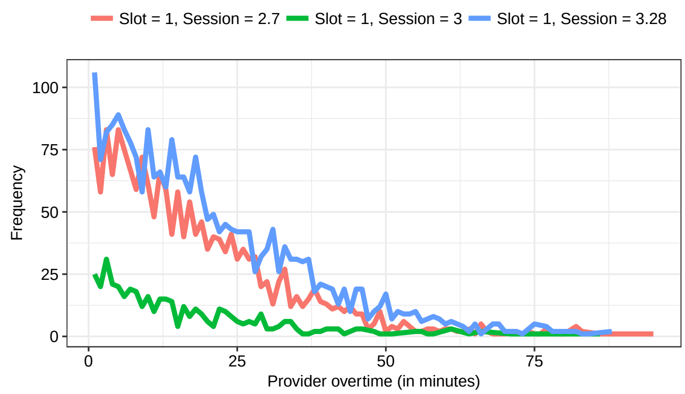

Figure 29: Provider overtime work distribution for different values of maximum slot and session complexity. In 77%, 95% and 70% of days, provider does not need to work overtime in (1,2.7), (1,3) and (1,3.28) of appointment slot and scheduling session complexities, respectively.

3.6.4 Patient 'toughness' with appointment slot timing

Patients might show different behavior when they are in contact with the scheduler to make an appointment. Once the scheduler realizes patient's desired date, she will start offering different appointment slots to patient. Some reasons such as employment or need for certain transportation might make patients more concerned about time and day of the appointment.

As expected, Figure 30 shows that if patients become tougher in accepting an appointment that is offered by the scheduler, appointment delay becomes longer for patients in the panel. However, our model, as well as heuristic sequences, shows a robust performance in terms of patient indirect waiting time distribution when patients show different behavior in appointment acceptance. Table 13 represents the optimal daily templates under different patient appointment acceptance thresholds.

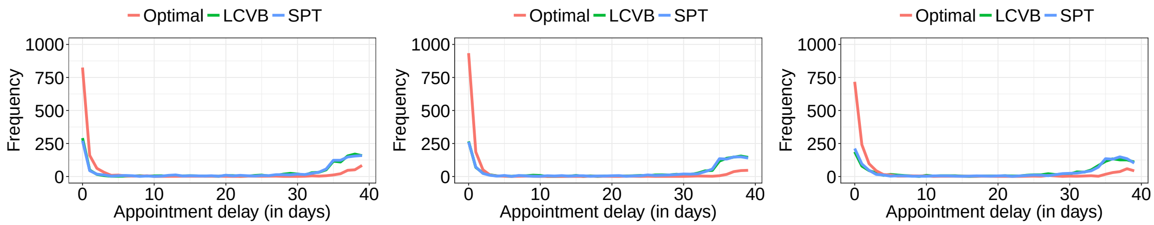

Figure 30: Indirect waiting time distribution for 0.2 (left), 0.4 (middle) and 0.6 (right) as appointment acceptance threshold.

Table 13: Optimal daily templates under different values of patient appointment acceptance threshold

Slot	Acceptance Threshold = 0.2	Acceptance Threshold = 0.4	Acceptance Threshold = 0.6
1	A,C,P	A,C,P	A,C
2	A,C,P	A,A,C	A,C,P
3	A,C,P	A,C	A,A,C
4	P	C,P	C,P
5	A,A,C	A,A,C	A,A
6	A,A,C	A,P	A,C,P
7	A,C	A,C,P	A,C
8	A,C	A,C	A,C,P

3.7 Conclusion

Although primary care is considered as the first contact point of patients with healthcare system, patients often have to wait long to visit their provider. A well established appointment scheduling system could help clinics to reduce patient indirect waiting time and improve patient flow in clinic.

Clinic managers have to handle multiple issues in patient appointment scheduling. While patients have different intentions to see the primary care provider and could have different complexities, they are also different in terms of call date and desired date and may not show up for their scheduled appointment. Some patients may call in advance to book an appointment whereas others ask for same-day one. This study examines the appointment scheduling problem by proposing a two-stage stochastic programming model in order to minimize patient indirect waiting time in getting an appointment while maintaining patient flow as smooth as possible inside clinic.

The numerical study shows the superiority of our proposed approach over heuristic approaches in patient appointment scheduling and lowering patient indirect waiting time as well as smoothing patient flow in clinic. Our model also shows a better performance in various cases when patient no-show is sensitive to appointment delay, patients become tougher in accepting appointment slots and provider is able to see more complex patients in every scheduling session.

While the two-stage stochastic programming model is offering a daily scheduling template, it could sometimes perform in favor of one patient type and result in longer appointment delay for other patient types. So an important direction for future research would be studying more flexible scheduling templates such as weekly or monthly template and comparing their performance with optimal daily scheduling template. Currently, optimal scheduling template only guides the scheduler on the number of patients of each patient type that could be scheduled in every appointment slot. An important question that could be answered in future research is that when should the scheduler start offering every open appointment slot to each patient type? We also studied the expected value of patient indirect waiting time in this research. However, some patients may still experience long appointment delay in this setting. It would be interesting to see how including risk-measures in the objective function will perform in terms of patient indirect waiting time and clinic patient flow.

CHAPTER 4: CONCLUSION AND FUTURE RESEARCH

As state in the introduction, in an attempt to improve the practicality of mathematical programming models and contribute to their adoption in the real-world, this dissertation contributes two original essays highlighting the contribution that stochastic programming can offer in solving important practical problems of interest to operations research. We do this with the recognition that stochastic programming, in considering uncertainty in mathematical modeling, often leads to large-scale programming problems. However, advances made in recent decades by algorithms in optimization software combined with advances in computing hardware allow us to tackle problems of greater complexity to provide meaningful solutions and decision support for the real-world.

We now present a summary of our research along with contributions and discuss directions of future research.

4.1 Summary

4.1.1 Designing Community-Aware Charging Network for Electric Vehicles

We studied the problem of designing a public charging infrastructure for electric vehicles in a community. Evidence shows that providing more charging service will increase EV owner's confidence in driving long distances with an EV and help promoting EV market share in the community. The presence of uncertainties in EV market share, daily arriving demand, arrival time to the community, state of battery charge at the time of arrival, dwell time at final destination, driver's preference to use public charging service and willingness to walk of driver bring complexities in designing this network. Incorporating these uncertainties, we proposed a two-stage stochastic programming model in order to find the location, number of chargers and level of charge for installing public EV charging stations. We adopted sample average approximation method to solve the two-stage stochastic model and developed an effective heuristic to solve large-scale problems.

Our model showed the interaction between access to EV charging service, utilization rate of charging stations and amount of change in walking distance of people in community for different number of charging stations in the community. We also showed that the model is robust towards any change in willingness to walk among people in the community.

4.1.2 Managing Access to Primary Care Clinics

We studied the problem of patient appointment scheduling in an outpatient primary care clinic setting in order to improve patient appointment delay while maintaining high utilization rate of staff and resources and considering patient flow in clinic. Uncertainties in demand volume, patient call date, patient desired date, patient no-show and patient service time with nurse and provide bring complexities to appointment scheduling problem. We proposed a two-stage stochastic programming model in order to provide a daily scheduling template that minimizes patient indirect waiting time in primary care clinics while maintain a smooth patient flow inside clinic and high utilization rate of provider capacity. We suggested an index policy to simulate patient appointment scheduling in call center by considering patient preference for day and time of the appointment and appointment cancellation by patient and by clinic.

Our numerical study showed the superior performance of our appointment scheduling approach over heuristic scheduling in terms of patient indirect waiting time and patient flow measures including patient direct waiting time, amount of time that provider spends with patients during lunch hour and amount of provider overtime work after regular hours. The study showed that if patients become more sensitive to appointment delay and their show up probability reduces as appointment delay increases, the model tries to schedule patients near to their desired date. Value of overbooking more patients in every scheduling session was investigated through assigning different values to maximum patient complexity that provider can handle in a session. Our model also illustrated a robust performance in appointment scheduling as patients show different behavior in accepting appointment slots.

4.2 Future research

In the case of the EV charging network design problem, the solution might not be implementable due to the assumption of centralized decision making. However, its solution serves as a reference point for evaluating existing and future network designs. More importantly, our model could be used to design incentive mechanism for charging station operators to make their location decision. In future research, we will develop an incentive allocation model which will optimize the allocation of incentive resources across multiple charging stations to influence their locations to come as close as possible to that of the centralized solution. Another avenue for future research is the inclusion of multi-modal transportation in the model and studying its impact on the optimal design of EV charging network.

While our appointment scheduling approach performs very well in reducing patient indirect waiting time and maintaining smooth patient flow, there are some important questions that can be answered in future research. Currently, we run the optimization model every day to find the optimal scheduling template but daily updating scheduling template might not be optimal. Finding the optimal time to re-optimizing the two-stage stochastic programming model is an important avenue of research. Moreover, while the two-stage stochastic programming model provides guidance on the number of patient for each patient type that can be scheduled in every appointment slot, it does not say anything about the optimal time of offering every open appointment slot to different patient types. Finding this optimal time could also result in better performance in patient indirect waiting time. Behavioral aspects of providing care by physician at different times of a day to different patient types and also the behavior of call center scheduler in offering open slots to patients seeking an appointment are other issues that can be incorporated in the appointment scheduling model.

In both studies, we only considered expected value of the objective function. Another avenue for future research is including risk-measures such CVaR and absolute semi-deviation in the objective function of two-stage stochastic programming.

REFERENCES

[1] Why not the best? results from the national scorecard on u.s. health system performance. 2011.

[2] Access is the answer: Community health centers, primary care & the future of american health care. 2014.

[3] Physician appointment wait times and medicaid and medicare acceptance rates. 2014.

[4] Midterm evaluation of light-duty vehicle greenhouse gas emission standards and corporate average fuel economy standards for model years 2022-2025. *Environmental Protection Agency*, 2016.

[5] Russell L Ackoff. Presidents' symposium: Or, a post mortem. *Operations Research*, 35(3):471–474, 1987.

[6] Roger T Anderson, Fabian T Camacho, and Rajesh Balkrishnan. Willing to wait?: the influence of patient wait time on satisfaction with primary care. *BMC health services research*, 7(1):31, 2007.

[7] Fouad Baouche, Romain Billot, Rochdi Trigui, and Nour-Eddin El Faouzi. Efficient allocation of electric vehicles charging stations: optimization model and application to a dense urban network. *IEEE Intelligent Transportation Systems Magazine*, 6(3):33–43, 2014.

[8] Angela S Bergantino, Claudia Capozza, Angela De Carlo, and Andrea Morone. What

drives parking choices: a laboratory experiment. 2014.

[9] Daniel J Berger and Andrew D Jorgensen. A comparison of carbon dioxide emissions from electric vehicles to emissions from internal combustion vehicles. *Journal of Chemical Education*, 92(7):1204–1208, 2015.

[10] Papiya Bhattacharjee and Pradip Kumar Ray. Simulation modelling and analysis of appointment system performance for multiple classes of patients in a hospital: a case study. *Operations Research for Health Care*, 8:71–84, 2016.

[11] John R Birge. The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical programming*, 24(1):314–325, 1982.

[12] Thomas Bodenheimer and Hoangmai H Pham. Primary care: current problems and proposed solutions. *Health Affairs*, 29(5):799–805, 2010.

[13] Lisa Brandenburg, Patricia Gabow, Glenn Steele, John Toussaint, and Bernard J Tyson. Innovation and best practices in health care scheduling. Technical report, Technical report, 2015.

[14] R Paul Brooker and Nan Qin. Identification of potential locations of electric vehicle supply equipment. *Journal of Power Sources*, 299:76–84, 2015.

[15] Ismail Capar and Michael Kuby. An efficient formulation of the flow refueling location model for alternative-fuel stations. *IIE Transactions*, 44(8):622–636, 2012.

[16] Ismail Capar, Michael Kuby, V Jorge Leon, and Yu-Jiun Tsai. An arc cover–path-cover formulation and strategic analysis of alternative-fuel station locations. *European Journal of Operational Research*, 227(1):142–151, 2013.

[17] Sanya Carley, Rachel M Krause, Bradley W Lane, and John D Graham. Intent to purchase a plug-in electric vehicle: A survey of early impressions in large us cites. *Transportation Research Part D: Transport and Environment*, 18:39–45, 2013.

[18] Joana Cavadas, Gonçalo Homem de Almeida Correia, and João Gouveia. A mip model for locating slow-charging stations for electric vehicles in urban areas accounting for driver tours. *Transportation Research Part E: Logistics and Transportation Review*, 75:188–201, 2015.

[19] Tugba Cayirli, Emre Veral, and Harry Rosen. Designing appointment scheduling systems for ambulatory care services. *Health care management science*, 9(1):47–58, 2006.

[20] Tugba Cayirli, Emre Veral, and Harry Rosen. Assessment of patient classification in appointment system design. *Production and Operations Management*, 17(3):338–353, 2008.

[21] Santanu Chakraborty, Kumar Muthuraman, and Mark Lawley. Sequential clinical scheduling with patient no-show: The impact of pre-defined slot structures. *Socio-Economic Planning Sciences*, 47(3):205–219, 2013.

[22] Daniel Chang, Daniel Erstad, Ellen Lin, A Falken Rice, C Tzun Goh, AA Tsao, and Jason Snyder. Financial viability of non-residential electric vehicle charging stations. *Luskin Center for Innovation: Los Angeles, CA, USA*, 2012.

[23] MengTing Dai, JingHong Zheng, Man Zhang, and WenZhuo Wang. Optimization of electric vehicle charging capacity in a parking lot for reducing peak and filling

valley in power grid. In *Advanced Power System Automation and Protection (APAP), 2011 International Conference on*, volume 2, pages 1501–1506. IEEE, 2011.

[24] Cinnamon A Dixon, Damien Punguyire, Melinda Mahabee-Gittens, Mona Ho, and Christopher J Lindsell. Patient flow analysis in resource-limited settings: A practical tutorial and case study. *Global Health: Science and Practice*, 3(1):126–134, 2015.

[25] Gregory Dobson, Sameer Hasija, and Edieal J Pinker. Reserving capacity for urgent patients in primary care. *Production and Operations Management*, 20(3):456–473, 2011.

[26] Nader Fahmy, Armen Aprikian, Mohammed Al-Otaibi, Simon Tanguay, Jordan Steinberg, Suganthiny Jeyaganth, Moamen Amin, and Wassim Kassouf. Impact of treatment delay in patients with bladder cancer managed with partial cystectomy in quebec: a population-based study. *Canadian Urological Association Journal*, 3(2):131, 2009.

[27] Timothy LJ Ferris. Churchman and measurement. In *Volume 1: Rescuing the Enlightenment from Itself*, pages 213–225. Springer, 2006.

[28] David S Festinger, RJ Lamb, Douglas B Marlowe, and Kimberly C Kirby. From telephone to office: Intake attendance as a function of appointment delay. *Addictive Behaviors*, 27(1):131–137, 2002.

[29] Inês Frade, Anabela Ribeiro, G Goncalves, and AP Antunes. An optimization model for locating electric vehicle charging stations in central urban areas. *Journal of Transport Research Record: Journal of the Transportation Research Board, DOI*, 10:2252–12,

2011.

[30] Gerard Gallucci, Wayne Swartz, and Florence Hackerman. Brief reports: Impact of the wait for an initial appointment on the rate of kept appointments at a mental health center. *Psychiatric Services*, 2005.

[31] Amireh Ghorob and Thomas Bodenheimer. Sharing the care to improve access to primary care. *New England Journal of Medicine*, 366(21):1955–1957, 2012.

[32] Linda V Green and Sergei Savin. Reducing delays for medical appointments: A queueing approach. *Operations Research*, 56(6):1526–1538, 2008.

[33] Diwakar Gupta and Brian Denton. Appointment scheduling in health care: Challenges and opportunities. *IIE transactions*, 40(9):800–819, 2008.

[34] Diwakar Gupta and Lei Wang. Revenue management for a primary-care clinic in the presence of patient choice. *Operations Research*, 56(3):576–592, 2008.

[35] Fang He, Yafeng Yin, and Jing Zhou. Deploying public charging stations for electric vehicles on urban road networks. *Transportation Research Part C: Emerging Technologies*, 60:227–240, 2015.

[36] Meysam Hosseini and SA MirHassani. Refueling-station location problem under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 84:101–116, 2015.

[37] Yongxi Huang and Yan Zhou. An optimization framework for workplace charging strategies. *Transportation Research Part C: Emerging Technologies*, 52:144–155, 2015.

[38] Garry Kaplan, Marianne Hamilton Lopez, and J Michael McGinnis. Transforming

health care scheduling and access: Getting to now. *Washington DC: Institute of Medicine*, 2015.

[39] Diego Klabjan and Timothy Sweda. The nascent industry of electric vehicles. *Wiley encyclopedia of operations research and management science*, 2011.

[40] Ellis Knight and Jeffrey Gorke. Managing patient flow by reducing variability. 2014.

[41] Renata Kopach, Po-Ching DeLaurentis, Mark Lawley, Kumar Muthuraman, Leyla Ozsen, Ron Rardin, Hong Wan, Paul Intrevado, Xiuli Qu, and Deanna Willis. Effects of clinical characteristics on successful open access scheduling. *Health care management science*, 10(2):111–124, 2007.

[42] KM Krause, KI Pollak, M Gradison, and JL Michener. Family physicians as team leaders: 'time' to share the care. *Preventing Chronic Disease*, 2009.

[43] John Krumm. How people use their vehicles: Statistics from the 2009 national household travel survey. Technical report, SAE Technical Paper, 2012.

[44] Linda R LaGanga and Stephen R Lawrence. Appointment overbooking in health care clinics to improve patient service and clinic performance. *Production and Operations Management*, 21(5):874–888, 2012.

[45] Sangbok Lee, Daiki Min, Jong-hyun Ryu, and Yuehwern Yih. A simulation study of appointment scheduling in outpatient clinics: Open access and overbooking. *Simulation*, 89(12):1459–1473, 2013.

[46] Todd Litman. Economic value of walkability. *World Transport Policy & Practice*, 10(1):5–14, 2004.

[47] Nan Liu and Serhan Ziya. Panel size and overbooking decisions for appointment-based services under patient no-shows. *Production and Operations Management*, 23(12):2209–2223, 2014.

[48] Nan Liu, Serhan Ziya, and Vidyadhar G Kulkarni. Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. *Manufacturing & Service Operations Management*, 12(2):347–364, 2010.

[49] Gaby Loria. Patient use of online reviews. 2013.

[50] Jianzhe Luo, Vidyadhar G Kulkarni, and Serhan Ziya. A tandem queueing model for an appointment-based service system. *Queueing Systems*, 79(1):53–85, 2015.

[51] Zhuowei Luo, Zechun Hu, Yonghua Song, Zhiwei Xu, and Haiyan Lu. Optimal coordination of plug-in electric vehicles in power grids with cost benefit analysis part 2: A case study in china. *IEEE Transactions on Power Systems*, 28(4):3556–3565, 2013.

[52] N Lutsey, S Searle, S Chambliss, and A Bandivadekar. Assessment of leading electric vehicle promotion activities in united states cities. *International Council on Clean Transportation*, 2015.

[53] Wai-Kei Mak, David P Morton, and R Kevin Wood. Monte carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, 24(1):47–56, 1999.

[54] Wesley E Marshall. An evaluation of livability in creating transit-enriched communities for improved regional benefits. *Research in Transportation Business & Management*, 7:54–68, 2013.

[55] Melanie Michael, Susan D Schaffer, Patricia L Egan, Barbara B Little, and Patrick Scott Pritchard. Improving wait times and patient satisfaction in primary care. *Journal for Healthcare Quality*, 35(2):50–60, 2013.

[56] Kumar Muthuraman and Mark Lawley. A stochastic overbooking model for outpatient clinical scheduling with no-shows. *Iie Transactions*, 40(9):820–837, 2008.

[57] Hyun-Jung Oh, Ana Muriel, Hari Balasubramanian, Katherine Atkinson, and Thomas Ptaszkiewicz. Guidelines for scheduling in primary care under different patient types and stochastic nurse and provider service times. *IIE Transactions on Healthcare Systems Engineering*, 3(4):263–279, 2013.

[58] Aydogan Ozdemir et al. Distributed storage capacity modelling of ev parking lots. In *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, pages 359–363. IEEE, 2015.

[59] Feng Pan, Russell Bent, Alan Berscheid, and David Izraelevitz. Locating phev exchange stations in v2g. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pages 173–178. IEEE, 2010.

[60] Andrew Pomerantz, Brady H Cole, Bradley V Watts, and William B Weeks. Improving efficiency and access to mental health care: combining integrated care and advanced access. *General hospital psychiatry*, 30(6):546–551, 2008.

[61] Julia C Prentice and Steven D Pizer. Delayed access to health care and mortality. *Health services research*, 42(2):644–662, 2007.

[62] Xiuli Qu, Ronald L Rardin, and Julie Ann S Williams. A mean–variance model to

optimize the fixed versus open appointment percentages in open access scheduling systems. *Decision Support Systems*, 53(3):554–564, 2012.

[63] Xiuli Qu and Jing Shi. Modeling the effect of patient choice on the performance of open access scheduling. *International Journal of Production Economics*, 129(2):314–327, 2011.

[64] Greg Rubin, Angela Bate, Ajay George, Phil Shackley, and Nicola Hall. Preferences for access to the gp: a discrete choice experiment. *Br J Gen Pract*, 56(531):743–748, 2006.

[65] Narges Shahraki, Hua Cai, Metin Turkay, and Ming Xu. Optimal locations of electric public charging stations using real world vehicle travel patterns. *Transportation Research Part D: Transport and Environment*, 41:165–176, 2015.

[66] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

[67] William Sierzchula, Sjoerd Bakker, Kees Maat, and Bert van Wee. The influence of financial incentives and other socio-economic factors on electric vehicle adoption. *Energy Policy*, 68:183–194, 2014.

[68] Jingzi Tan and Wei-Hua Lin. A stochastic flow capturing location and allocation model for siting electric vehicle charging stations. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 2811–2816. IEEE, 2014.

[69] Theodore Tsiligirides. Heuristic methods applied to orienteering. *Journal of the Operational Research Society*, 35(9):797–809, 1984.

[70] Christopher Upchurch, Michael Kuby, and Seow Lim. A model for location of capacitated alternative-fuel stations. *Geographical Analysis*, 41(1):85–106, 2009.

[71] Stanislav Uryasev and Panos M Pardalos. *Stochastic optimization: algorithms and applications*, volume 54. Springer Science & Business Media, 2013.

[72] Sydney Vergis and Belinda Chen. Comparison of plug-in electric vehicle adoption in the united states: A state by state approach. *Research in Transportation Economics*, 52:56–64, 2015.

[73] Stein W Wallace and William T Ziemba. *Applications of stochastic programming*. SIAM, 2005.

[74] Wen-Ya Wang and Diwakar Gupta. Adaptive appointment systems with patient preferences. *Manufacturing & Service Operations Management*, 13(3):373–389, 2011.

[75] Ying-Wei Wang and Chuah-Chih Lin. Locating multiple types of recharging stations for battery-powered electric vehicle transport. *Transportation Research Part E: Logistics and Transportation Review*, 58:76–87, 2013.

[76] Katie Wellstood, Kathi Wilson, and John Eyles. Reasonable access to primary care: assessing the role of individual and system characteristics. *Health & Place*, 12(2):121–130, 2006.

[77] Fei Wu and Ramteen Sioshansi. A stochastic flow-catching model to optimize the location of fast-charging stations with uncertain electric vehicle flows.

[78] Chongjun Yan, Jiafu Tang, Bowen Jiang, and Richard YK Fung. Sequential appointment scheduling considering patient choice and service fairness. *International Jour-*

nal of Production Research, 53(24):7376–7395, 2015.

[79] Yong Yang and Ana V Diez-Roux. Walking distance by trip purpose and population subgroups. *American journal of preventive medicine*, 43(1):11–19, 2012.

[80] Christos Zacharias and Michael Pinedo. Appointment scheduling with no-shows and overbooking. *Production and Operations Management*, 23(5):788–801, 2014.

[81] Bo Zeng, Ayten Turkcan, Ji Lin, and Mark Lawley. Clinic scheduling models with overbooking for patients with heterogeneous no-show probabilities. *Annals of Operations Research*, 178(1):121–144, 2010.

[82] Ming Zhong, John Hunt, and Xuewen Lu. Studying differences of household weekday and weekend activities: a duration perspective. *Transportation Research Record: Journal of the Transportation Research Board*, (2054):28–36, 2008.

ABSTRACT

ESSAYS ON STOCHASTIC PROGRAMMING IN SERVICE OPERATIONS MANAGEMENT

by

SINA FARIDIMEHR

August 2017

Advisors: Dr. Ratna Babu Chinnam & Dr. Saravanan Venkatachalam

Major: Industrial Engineering

Degree: Doctor of Philosophy

Deterministic mathematical modeling is a branch of optimization that deals with decision making in real-world problems. While deterministic models assume that data and parameters are known, these numbers are often unknown in the real world applications. The presence of uncertainty in decision making can make the optimal solution of a deterministic model infeasible or sub-optimal. On the other hand, stochastic programming approach assumes that parameters and coefficients are unknown and only their probability distribution can be estimated. Stochastic programming can include uncertainties in objective function of random variables and/or constraints. Stochastic programming has seen a growing range of applications in manufacturing production planning, machine scheduling, dairy farm expansion planning, asset liability management, traffic management, and automobile dealership inventory management that involve uncertainty in decision making. The most widely used stochastic programming approach is the two-stage stochastic

programming models. In this model, first-stage decision variables are determined before observing the realization of uncertainties and second-stage decision variables are selected after exposing first-stage variables into the uncertainties. The goal is to determine the value of first-stage decisions in a way to maximize (minimize) the expected value of second-stage objective function. In an attempt to improve the practicality of mathematical programming models and contribute to their adoption in the real-world, this dissertation contributes two original essays highlighting the contribution that stochastic programming can offer in solving important practical problems of interest to operations research.

In the first essay, We study the problem of designing a charging infrastructure for electric vehicles in an urban area. Electric vehicles (EVs) are attracting more and more attentions these days due to increase concern about global warming and future shortage of fossil fuels. These vehicles have potential to reduce greenhouse gas emissions, improve public health condition by reducing air pollution and improving sustainability, and address diversification of transportation energy feedstock.

Governments and policy makers have proposed two types of policy incentives in order to encourage consumers to buy an EV: direct incentives and indirect incentives. Direct incentives are those that have direct monetary value to consumers and include purchase subsidies, license tax/fee reductions, Electric Vehicle Supply Equipment (EVSE) financing, free electricity, free parking and emission test exemptions. On the other hand, indirect incentives are the ones that do not have direct monetary value and consist of high-occupancy vehicle access, emissions testing exemption time savings, and public charger availability. Lack of access to public charging network is a major barrier in adoption of EVs. Access to public charging infrastructure will provide confidence for EV owners to drive longer distances without going out of charge and encourage EV ownership in the community.

The current challenge for policy makers and city planners in installing public charging infrastructure is determining the location of these charging service stations, number of required stations and level of charging since the technology is still in its infancy and the installation cost is high. Since recharging of EV battery takes more time than refueling conventional vehicles, parking lots and garages are considered as potential locations for installing charging stations. The aim of this research is to develop a mathematical programming model to find the optimal locations with potentially high utilization rate for installing community-aware public EV charging infrastructure in order to improve accessibility to charging service and community livability metrics. In designing this charging network, uncertainties such as EV market share, state of battery charge at the time of arrival, driver's willingness to charge EV away from home, arrival time to final destination, driver's activity duration (parking duration), and driver's walking distance preference play major role. Incorporating these uncertainties in the model, we propose a two-stage stochastic programming approach to determine the location and capacity of public EV charging network in a community.

In the second essay, We study access to care problem in outpatient primary care clinics. Patient access to care along with healthcare efficiency and quality of service are dimensions of health system performance measurement. Improving access to primary care is a major step of having a high-performing healthcare system. However, many patients are struggling to get an in-time appointment with their own primary care provider (PCP). Even two years after health insurance coverage was expanded, new patients have to wait 82% longer to get an internal-medicine appointment. A national survey shows that percentage of patients that need urgent care and could not get an appointment increased from 53% to 57% between 2006 and 2011. This delay may negatively impact patient health status and may even lead to death. Patients that cannot get an appointment with their PCP may seek care with other providers or in emergency departments which will decrease continuity of care and increase total cost of health system.

The main issue behind access problem is the imbalance between provider capacity and patient demand. While provider panel size is already large, the shortage in primary care providers and increasing number of patients mean that providers have to increase their panel size and serve more patients which will potentially lead to lower access to primary care. The ratio of adult primary care providers to population is expected to drop by 9% between 2005 and 2020. Moreover, patient flow analysis can increase efficiency of healthcare system and quality of health service by increasing patient and provider satisfaction through better resource allocation and utilization. Effective resource allocation will smooth patient flow and reduce waste which will in turn result in better access to care.

One way to control patient flow in clinic is managing appointment supply through appointment scheduling system. A well-designed appointment scheduling system can decrease appointment delay and waiting time in clinic for patients and idle time and/or overtime for provider simultaneously and increase their satisfaction. Appointment scheduling requires to make a balance between patient needs and facility resources. The purpose of this study is to develop appointment scheduling model using two-stage stochastic programming to improve access to care while maintaining high levels of provider capacity utilization and improving patient flow in clinic by leveraging uncertainties in patient demand volume, patient no-show, nurse service time and provider service time.

AUTOBIOGRAPHICAL STATEMENT

Sina Faridimehr received his B.S. and M.S. degrees in industrial engineering from Sharif University of Technology, Tehran, Iran, in 2009 and 2011, respectively. He is currently working toward his Ph.D. degree in Industrial and Systems Engineering with Wayne State University, Detroit, Michigan. His research interests are in service operations management, and data-driven decision making. His papers have been published in top-tier journals such as OMEGA (The International Journal of Management Science), International Journal of Production Economics, and Scientia Iranica.