HEALTHCARE FACILITY LOCATION AND CAPACITY CONFIGURATION UNDER STOCHASTIC DEMAND

A Dissertation

by

XUE HAN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Wilbert E. Wilhelm
Committee Members,	Sergiy Butenko
	Lewis Ntaimo
	Donald K. Friesen
Head of Department,	César O. Malavé

December 2014

Major Subject: Industrial Engineering

ABSTRACT

This dissertation addresses two topics. The first topic is strategic dynamic supply chain reconfiguration (DSCR) problem, in which the proposed capacity configuration network is employed in the second topic: healthcare facility location and capacity configuration under stochastic demand. The second topic investigates two problems: the stochastic, single healthcare facility location and capacity configuration problem (SSHFCP) in a competitive environment and the stochastic, multiple healthcare facility location and capacity configuration problem (SMHFCP) based on a location-allocation model.

The DSCR problem is to prescribe the location and capacity of each facility, select links used for transportation, and plan material flows through the supply chain, including production, inventory, backorder, and outsourcing levels. The objective is to minimize total cost. The network must be dynamically *reconfigured* (i.e., by opening facilities, expanding and/or contracting their capacities, and closing facilities) over time to accommodate changing trends in demand and/or costs. This research proposes a network-based model of DSCR and compares it with a traditional mixed integer programming (MIP) formulation via extensive, large-scale computational tests and sensitivity analyses, showing that the network-based model offers superior solvability.

The SSHFCP is to prescribe the location and multi-service, multi-period capacity configuration of facility facing competition from existing facilities under *uncertain* patient demand, so that the expected *excess revenue* (i.e., the amount by which revenue exceeds cost) of the new facility is maximized. This dissertation describes a solution methodology that relates practical features relative to healthcare, including a multiplicative competitive interaction (MCI) model to reflect competition among providers and a method to model the stochastic problem as a deterministic resource constrained shortest path problem (RCSPP) on a specially constructed

network, which can be solved in pseudo-polynomial time. This dissertation proposes two solution methods to SMHFCP. The dissertation shows that first method, a column-generation heuristic, solves test instances to near optimality; and the second one, an approximation method, provides a fast runtime with a bounding procedure to assess the quality of a solution. The application of SSHFCP and SMHFCP in locating and configuring new primary care centers in mid-Texas rural area validates the real business decision of industrial collaborators.

To my beloved grandparents, Shuxiang Liu and Baorong Zhang

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Wilbert E. Wilhelm, in the Department of Industrial and Systems Engineering, Texas A&M University, for his continuous guidance, support and inspiration throughout my Ph.D. study. I am grateful to be his student and learn from him the rigorous attitude in conducting research and teaching.

I want to thank Dr. Wilhelm for his significant contributions of time, idea and funding opportunities for my Ph.D. study. Throughout the editing process, much of the wording in this dissertation was edited and composed by Professor Wilhelm. Various chapters of this dissertation have been or will be published in referred journals, which will hold the associated copyrights.

I would like to thank Dr. Sergiy Butenko, Dr. Lewis Ntaimo and Dr. Donald K. Friesen for serving on the committee and for giving valuable suggestions. Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. Especially, I want to thank Dr. Chaehwa Lee for collaborating in our research. Our DSCR research is based on the previous research of Dr. Chaehwa Lee and Dr. Wilhelm. The original model uses random generated numbers to represent the amount of expansions and contractions. I refined the model to use decision variables to define such amounts and modified the cost definition in objective function. I also performed a new set of computational experiments.

I sincerely thank my parents, Peng Han and Yunhui Liu, for taking care of my son since his birth and playing the roles that my husband and I have not completely fulfilled. Thanks to my husband, Yuan Yao, for his continuous encouragement and support throughout my Ph.D. study and my future career path. I cannot finish this dissertation without all your support. Finally, thanks to my son, Logan Han Yao, for bringing enormous happiness to our extended family.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION[*]

This dissertation addresses two topics. The first topic is strategic dynamic supply chain reconfiguration (DSCR) problem, in which the proposed capacity configuration network is employed in the second topic: healthcare facility location and capacity configuration under stochastic demand. The second topic comprises two cases: the stochastic, single healthcare facility location and capacity configuration problem (SSHFCP) in a competitive environment and the stochastic and multiple healthcare facility location and capacity configuration problem (SMHFCP) based on a location-allocation model.

Section 1.1 introduces the DSCR problem. Section 1.2 introduces two cases of the healthcare facility location and capacity configuration problems, SSHFCP and SMHFCP.

This dissertation is organized in six chapters. Chapter II reviews literature relevant to DSCR and healthcare facility location and capacity configuration problems. Chapters III investigates the DSCR problem. Chapters IV and V investigate SSHFCP and SMHFCP, respectively. The last chapter presents conclusions of this dissertation research and outlines opportunities for future work.

1.1 Introduction to DSCR problem

A supply chain must be dynamically *reconfigured* (i.e., by opening facilities, expanding and/or contracting their capacities, and closing facilities) over time to cope with changes in demand

---

[*] Reprinted with permission from "Computational comparison of two formulations for dynamic supply chain reconfiguration with capacity expansion and contraction." by Wilhelm, W. E., Han, X. and Lee, C., 2013, *Computers & Operations Research*, 40, 2340-2356, Copyright 2013 by Elsevier.

and/or cost structures as the business environment evolves. Demand for products in each market and costs to produce them at each plant vary as economic factors change over time. For example, an economic downturn or a period of rapid growth may give rise to such changes and force an enterprise to reconfigure its supply chain to meet customer demands at the lowest possible cost (M. T. Melo, Nickel, & Saldanha da Gama, 2006). Another example of a phenomenon that gives rise to such changes is the product life cycle: demand increases after introduction, grows rapidly, plateaus, and then decreases as the end of the life cycle approaches.

The dynamic supply chain reconfiguration (DSCR) problem is to prescribe facility opening, capacity expansion and contraction, and facility closing at each potential location in a multi-period, multi-product, multi-echelon supply chain. This strategic problem involves a planning horizon of some 6-10 years. DSCR models are needed to provide decision support for management in dealing with changing business conditions in the competitive modern business environment. The objectives of this research are

- a traditional formulation and a network-based model of the DSCR problem,
- tests under different demand scenarios to promote an intuitive interpretation of our models,
- tests that identify the computational characteristics of our models to assess solvability, and
- tests to identify sensitivity of run time relative to primary parameters.

To achieve the first objective, this dissertation presents a traditional mixed integer program (MIP) and then proposes an alternative model that relates binary decision variables according to a network structure.

Even though the dynamic facility location problem with facility openings and closings has been studied for some time, there has not been adequate attention to cases that involve capacity expansion and contraction over the planning horizon. Furthermore, little research has been directed

to dynamic facility location within a multi-period, multi-product, multi-echelon supply chain network. No prior work has studied the solvability of different model forms.

Our DSCR models prescribe material flow through a four-echelon supply chain: suppliers, plants, distribution centers (DCs), and customer zones (CZs). Each echelon performs a unique function so that each product must be "processed" in each. Each viable transportation link allows shipment from a facility in one echelon to another in the next echelon; no links connect facilities within the same echelon. A viable transportation link is established between a pair of operating facilities and any product can be transported on it. Thus, the problem deals with a dynamic, multi-period, multi-product, multi-echelon supply chain network through which products are delivered to satisfy demands, which occur only in CZs.

We model inventory, backorders, and outsourcing in each time period over the planning horizon because they allow peak demands to be satisfied and are thus essential to customer service. Previous studies have addressed these features in strategic planning (M. S. Daskin, Coullard, & Shen, 2002; M. S. Daskin, Snyder, & Berger, 2005; Yolanda Hinojosa, Kalcsics, Nickel, Puerto, & Velten, 2008; M. T. Melo et al., 2006; Shen, Coullard, & Daskin, 2003; W. Wilhelm et al., 2005). Of 60 papers reviewed by (M. T. Melo, Nickel, & Saldanha-da-Gama, 2009), 33 include inventory planning in facility location and supply chain management models.

Following (M. T. Melo et al., 2006), which studied dynamic facility location with inventory planning, we consider two cases for contraction and closure costs. In Case 1, it can be profitable for a contractor to close a facility to eliminate unused capacity because the associated cost is negative, indicating that a return can be obtained by selling infrastructure and equipment that is no longer needed. In Case 2, the costs to contract or close are positive, so that unused capacity can be economically eliminated only when doing so is less costly than maintaining it (e.g., heating, cooling, insurance, security and taxes).

Supply chains that are reconfigured abound in industry. After the North American Free Trade Agreement (NAFTA) was signed in 1993, many operations were moved from the U.S. to Mexico. Subsequently, China attracted many of these operations; and, now, because wages have risen in China, some are being moved to lower-cost countries in Asia and some are being on-shored to the U.S. (JR. Hagerty, 2012). For over a hundred years, the textile industry has moved from one country to another around the world, seeking low cost labor (Rivoli, 2009). The current paper deals with structures that underlie both domestic and international supply chains and need only be augmented with international financial issues (e.g., border crossing fees, tariffs, local content rules, and transfer prices) for application to global supply chains (W. Wilhelm et al., 2005). Domestically, General Motors has recently closed plants in the Midwest and open new ones in the south, while other manufacturers streamline their supply chains (JR Hagerty, 2012).

1.2 Introduction to SSHFCP and SMHFCP

This dissertation investigates two cases of healthcare facility location and capacity configuration, the single-facility case in a competitive environment and the multi-facility case based on a location-allocation model.

The first case, SSHFCP, is to prescribe the location and the multi-service, multi-period capacity configuration of a facility facing competition from existing facilities under *uncertain* patient needs. SSHFCP prescribes the location of a new facility from a set of potential sites as well as the capacity assigned to each specified service in each time period, allowing openings, expansions, contractions, and closures over the planning horizon, so that the expected *excess revenue* (i.e., the amount by which revenue exceeds cost) of the new facility is maximized.

The research objectives of SSHFCP are to synthesize a solution methodology that includes a model of SSHFCP; a model that reflects competition among healthcare providers; a model that

represents uncertain demand, allowing closed-form expressions for recourse (i.e., expected excess demand and capacity) to compensate for random outcomes that are realized over the planning horizon; as well as to demonstrate application of the methodology in a case study that configures a new primary care center and to reveal insights into the relative impacts of key competition factors and parameters through sensitivity analyses. In particular, we model the stochastic SSHFCP as a deterministic resource constrained shortest path problem (RCSPP) on a specially constructed network, which can be solved in pseudo-polynomial time.

We consider our models to be prototypical in the sense that they structure, for the first time, fundamental features that are relevant to a broad range of healthcare applications. These models can be adapted, as we exemplify, to particular applications and enhanced to incorporate additional features as they are identified in the future.

The second case, SMHFCP, is to prescribe the locations of new facilities, given a set of potential sites, as well as the capacity assigned to each specified service at each facility in each time period, allowing openings, expansions, contractions, and closures over the planning horizon, so that the total expected excess revenue of the system is maximized. We employ a location-allocation model which assigns demand from each population center to a particular facility.

This dissertation proposes two methods to solve SMHFCP, a column-generation heuristic and an approximation method. We also perform computational experiments on realistic case studies that locates primary care centers in mid-Texas rural area.

This work is motivated by the fact that the U.S. is seeking to expand access to healthcare services for individuals in underserved (e.g., rural) areas and for others through the Patient Protection and the Affordable Care Act (ACA). Healthcare is provided at primary, secondary, tertiary and quaternary levels. Primary care is the initial point of consultation with an internist, a family physician or a pediatrician. This level deals with all types of issues (e.g., high blood pressure, diabetes, arthritis)

but refers patients with more serious issues to specialists (e.g., cardiologists, urologists) for secondary or tertiary care. For example, a secondary-care specialist may perform a "routine" heart bypass surgery, but tertiary care may be required for a patient who has complications due to age or multiple morbidities. The quaternary level specializes in advanced procedures (e.g., transplants) that have relatively low demand.

Primary care is the entry into the healthcare system and leads to referrals that require hospital treatment; for this reason, it may be advantageous for a hospital to employ primary care physicians (PCPs). A patient may consider distance important in selecting a PCP because s\he may make repeated visits each year. In addition, factors like the skill, reputation, and treatment success of a physician may also be important in this selection process. Distance may be a lesser consideration at the tertiary and quaternary levels. For example, people are drawn from around the world to the Houston Medical Complex because of reputation, physician skills and success rate. Providers may compete on the basis of physician ratings as well as offering services that would otherwise be under-supplied in a particular locale. Furthermore, demographics change dynamically as the population ages and emigration (e.g., Detroit) and immigration (e.g., Texas) occur, affecting healthcare needs and selection decisions.

Our models can be expected to improve healthcare effectiveness, enhancing access to avoid underutilization of costly professionals and equipment or, worse, rationing of overloaded resources. Healthcare administrators can use our models to serve patient needs effectively in a variety of applications, including locating and prescribing the capacity configuration of a new hospital or primary care, urgent care, or treatment (e.g., dialysis or trauma) center; or emergency room. Some 36 states require a provider to obtain a Certificate of Need (CON) before constructing a new facility, expanding an existing facility, or enhancing capacity. We note that a provider can use our model to justify a CON application and that a regulatory body can use it to configure the need.

CHAPTER II

LITERATURE REVIEW*

This chapter reviews literature relevant to DSCR and healthcare facility location and capacity configuration problems in sections 2.1 and 2.2, respectively.

2.1 Literature review: DSCR

The DSCR problem is related to four classical OR problems: facility location, dynamic facility location, supply chain design, and production-distribution network design. The facility location problem involves siting a set of facilities to serve a set of customer demands with the objective of minimizing total distance (or cost) incurred by all transports (Daskin et al., 2005; Melo et al., 2009). An extension, the dynamic (multi-period) location problem, has been proposed to meet demands and costs as they change over time (Melo et al., 2009) and as a basis for building comprehensive supply chain network models (Melo et al., 2009).

A supply chain network comprises a number of facility types that perform operations ranging from acquiring raw materials, transforming materials into intermediate and finished products, and distributing finished products to customers (Yolanda Hinojosa et al., 2008; Melo et al., 2009). A specialization of the supply chain design problem is called the production-distribution network design problem (Andreas Klose & Drexl, 2005), which is also a special case of the network design problem in which the network is acyclic.

Due to the wide range of applications and its challenges to solution methods, the dynamic

---

* Reprinted with permission from "Computational comparison of two formulations for dynamic supply chain reconfiguration with capacity expansion and contraction." by Wilhelm,W. E., Han, X. and Lee, C., 2013, *Computers & Operations Research*, 40, 2340-2356, Copyright 2013 by Elsevier.

facility location problem with opening and closing has been studied widely since the first work of (Ballou, 1968), including both uncapacitated (Canel & Khumawala, 1997; Chardaire, Sutter, & Costa, 1996; Van Roy & Erlenkotter, 1982) and capacitated (Fong & Srinivasan, 1981a, 1981b; S.-B. Lee & H. Luss, 1987; Alexander Shulman, 1991) cases. The dynamic supply chain network problem, which includes locating facilities, has been studied by (Canel & Khumawala, 1997; Gue, 2003; Yolanda Hinojosa et al., 2008; Melachrinoudis & Min, 2000; Melo et al., 2006).

The possibility of expanding capacity was considered by (Aghezzaf, 2005). (Lowe & Preckel, 2004) modeled the capacity-contraction case. A few studies (Behmardi & Lee, 2008; Melachrinoudis & Min, 2000; Melo et al., 2006; Wang, Batta, Bhadury, & Rump, 2003) considered both capacity expansion and contraction. (Daskin et al., 2005; Andreas Klose & Drexl, 2005; Melo et al., 2009) provided surveys of the dynamic facility location problem.

In particular, a few papers are closely related to this research. (Yolanda Hinojosa, Puerto, & Fernández, 2000) dealt with the multi-period, multi-product, two-echelon, capacitated location problem in which new facilities can be opened and existing facilities closed but didn't consider practical features like inventory, capacity expansion and contraction, or a budget limitation. (Melo et al., 2006) considered the step-wise reallocation of capacities under the assumptions that all existing facilities are operating at the start of the planning horizon; if an existing facility is closed, it cannot be reopened; and when a new facility is opened, it will remain in operation.

(Behmardi & Lee, 1914) studied a dynamic, multi-product, capacitated facility location problem in which each facility can be opened and subsequently closed with no reopening allowed. Extending (Yolanda Hinojosa et al., 2000), (Yolanda Hinojosa et al., 2008) formulated a model for a dynamic, two-echelon, multi-product, capacitated facility location problem with inventory and outsourcing. (Thanh, Bostel, & Péton, 2008) proposed a MIP to design of a multi-product, multi-echelon, production–distribution network, considering the opening, expanding, and closing

of facilities as well as supplier selection. Inventories were held only in warehouses, not in plants. (Torres Soto, 2009) studied the dynamic, capacitated facility location problem, which determines the optimal locations and times for opening facilities when demand and cost parameters are time-varying. This model minimizes costs of transporting and the opening, operating, closing, and reopening of facilities. As in (Wesolowsky & Truscott, 1975), (Torres Soto, 2009)employed binary variables for (re)opening, closing, and operating a facility, but neither allowed for capacity expansion or contraction.

In most models that allow only facility opening and closing (Behmardi & Lee, 2008; Yolanda Hinojosa et al., 2008; Yolanda Hinojosa et al., 2000; Melo et al., 2006; Thanh et al., 2008; Van Roy & Erlenkotter, 1982), the capacity of a facility cannot be increased or decreased over time. Facilities that are open at the start of the planning horizon can only be contracted or closed and, after closing, must remain closed until the end of planning horizon. Facilities that are not operating at the start of the planning horizon can only be opened and subsequently expanded; but an open facility must remain opened until the end of the planning horizon - it cannot be closed and its capacity cannot be contracted. In particular, this approach does not allow for a facility with excessive capacity to be closed or contracted. Our model fills this gap, allowing capacity expansion and contraction as well as closures.

A number of solution approaches have been proposed: commercial mathematical programming software (Gue, 2003; Melachrinoudis & Min, 2000; Melo et al., 2006), branch and bound (B&B) (Canel & Khumawala, 1997; Cem Canel, Basheer M Khumawala, Japhett Law, & Anthony Loh, 2001; Van Roy & Erlenkotter, 1982), Benders decomposition (Geoffrion & Graves, 1974; Torres Soto, 2009), dynamic programming (Balakrishnan, 2004; S.-B. Lee & H. Luss, 1987; Alexander Shulman, 1991), Lagrangian relaxation (Yolanda Hinojosa et al., 2000; Torres Soto, 2009) and heuristics (Chardaire et al., 1996; Fong & Srinivasan, 1981a, 1981b; Wang et al., 2003).

As (Andreas, Klose & Drexl, 2005) indicated, the computational challenge presented by the dynamic facility location problem increases drastically with the size of the model, reducing the chances to solve large-scale, real-world instances.

2.2 Literature review: healthcare facility location and capacity configuration

The healthcare facility location and capacity configuration involves four topics. The first topic we address is healthcare, dynamic and stochastic facility location. Second, we review models of demand attraction and facility location in a competitive environment, including queueing models used to maximize the capture of demand. Third, we review clinic scheduling models, which assess cost penalties for patient waiting time and overtime, providing justification for maximizing expected excess revenue, as we do. Fourth, we review the location-allocation models in healthcare applications.

This section comprises four subsections; each involves a topic; 2.2.1, Healthcare, dynamic and stochastic facility location; 2.2.2, Demand attraction models under competition; 2.2.3, Optimizing expected excess revenue; and 2.2.4, Location-allocation model in healthcare applications.

2.2.1 Healthcare, dynamic and stochastic facility location

A variety of literature is related to this work, including research on healthcare and dynamic facility location. Several recent papers provided reviews of healthcare-facility location (Daskin & Dean, 2004; Rais & Viana, 2010), including emergency services (Adenso-Diaz & Rodriguez, 1997), hospitals (Sinuany-Stern, Mehrez, Tal, & Shemuel, 1995), preventive services (Verter & LaPierre, 2002; Zhang, Berman, Marcotte, & Verter, 2010), rural services (Griffin, Scherrer, & Swann, 2008), and hierarchical networks (Galvão, Espejo, Boffey, & Yates, 2006). In addition, studies

have addressed location (Mehrez, Sinuany-Stern, T., & Binyamin, 1996; Ndiaye & Alfares, 2008) and number (Schweikhart & Smith-Daniels, 1993) of facilities, assignment (Griffin et al., 2008) and reassignment (Gunes & Yaman, 2010) of services, capacity (Berman, Ganz, & Wagner, 1994; Jayaraman & Srivastava, 1995; Smith-Daniels, Schweikhart, & Smith-Daniels, 1988), and location-allocation involving travel and waiting times (Syam & Côté, 2010). A recent study located new traumatic brain-injury units to treat military personnel, who were injured by improvised explosive devices in Iraq and Afghanistan, within facilities operated by the Veterans Health Administration (Côté, S. Syam, Vogel, & Cowper, 2007). Even this problem of limited scope required a heuristic to address the *deterministic* case. Various objectives have been addressed, including minimizing travel (Wang et al., 2003), maximizing demand served (Gendreau, Laporte, & and Semet, 2006; Verter & LaPierre, 2002), and optimizing multiple criteria (Schweikhart & Smith-Daniels, 1993). Results of some studies have been used successfully in industry (Jacobs, Silan, & Clemson, 1996; Price & Turcotte, 1986). Solution methods include Lagrangian relaxation (Berman et al., 1994; Haghani, 1996) and heuristics (Adenso-Diaz & Rodriguez, 1997; M. Daskin & Dean, 2004). Several studies have addressed uncertainty, in particular for emergency services, using stochastic programming (Beraldi, Bruni, & Conforti, 2004; Berman et al., 1994), robust optimization (Baron & Milner, 2010), or simulation (De Angelis, Felici, & Impelluoso, 2003; Harper, Shahani, Gallagher, & Bowie, 2005).

Research on generic, dynamic facility location problems (Owen & Daskin, 1998) has invoked various rules for opening, expanding, contracting and closing, addressing both uncapacitated (Tcha & Lee, 1984) and capacitated (Fong & Srinivasan, 1981a, 1981b; Lee & Luss, 1987; A Shulman, 1991) cases as well as broader supply chain design (Canel, Khumawala, Law, & Loh, 2001; Klose & Drexel, 2005; Melo, Nickel, & Saldanha da Gama, 2009; Thanh, Bostel, & Peton, 2008), for which W. E. Wilhelm, Han, and Lee (2013) recently showed that the network

models of the type we propose outperform traditional integer programs in resolving dynamic configuration problems. The most-used solution approaches include Branch-and-Bound (B&B) (Canel et al., 2001; Tcha & Lee, 1984), dynamic programming (Lee & Luss, 1987; A Shulman, 1991), Lagrangian relaxation (Y Hinojosa, Kalcsics, Nickel, Puerto, & Velton, 2008), and heuristics (Fong & Srinivasan, 1981a, 1981b). The dynamic facility location problem presents computational challenges, giving rise to the need for a more effective, exact solution approach (A Klose & Drexel, 2005). Reviews are provided by (Klose & Drexel, 2005; Melo et al., 2009).

In contrast to the earlier work on facility location, the novel, key component we use to locating a facility and configuring its capacity involves formulating a RCSPP, which is known to be NP-hard, even if only one resource is limited (Handler & Zang, 1980), but can be solved in pseudo-polynomial time. Solution methods for RCSPP include dynamic programming (Dumitrescu & Boland, 2003; Joksch, 1966) and Lagrangian relaxation (Carlyle, Royset, & Wood, 2008; Handler & Zang, 1980; Santos, Coutinho-Rodrigues, & Current, 2007). A recent paper (Lozano & Medaglia, 2013) provided an exact solution method that solves certain large networks in less than 0.3 seconds. Our network formulation enables us to prescribe a globally optimal location and capacity configuration for multiple services over a multi-period planning horizon. We use a three-stage labeling method proposed by Zhu and Wilhelm (2012) to solve the RCSPP.

A few facility-location studies have addressed uncertainty explicitly (Louveaux, 1993; Snyder, 2007). Models that deal with broader issues in supply chains have been proposed (Santoso, Ahmed, Goetschalckx, & Shapiro, 2004), making some progress in dealing with uncertainty. However, stochastic programming capabilities are still evolving, especially in application to large-scale instances. By formulating a deterministic equivalent, (Albareda-Sambola, Fernández, & Saldanha-da-Gama, 2011) solved a facility location problem involving binomial demands. We show how to transform the more general, stochastic SSHFCP so that it can

be solved as a deterministic RCSPP, reflecting decisions that individuals make in selecting providers, dealing with all possible outcomes instead of representative scenarios as done by two-stage stochastic programming.

2.2.2 Demand attraction models under competition

We model competition among facilities to attract demand. The Huff model (Huff, 1963), originally used to reflect competition in locating a new shopping center, defines the attraction of a facility to an individual customer as the size of the facility divided by the distance to the facility. The MCI model of (Nakanishi & Cooper, 1974) is a general form of the Huff model. MCI uses the product of all attributes that affect a customer's choice to quantify the attraction of customers to a retail facility.

McFadden (1974) proposed a multinomial logit (MNL) model. The odds of visiting one location from a given set of locations defines the conditional probability of choosing it, given the measures associated with a set of attributes. MNL sets one location as a baseline and calculates the logarithm of the odds for each other location relative to the baseline (i.e., log-odds). An exponential function of the log-odds represents the attraction of an individual customer to a facility. In all such competition models, the probability a customer will seek service at a particular facility is expressed as the "gravity" of the attraction of this facility relative to the other facilities in the market. Our prototypical MCI model incorporates competition factors that are especially relevant to healthcare.

Research on competitive facility location has dealt with deterministic demand (Aboolian et al., 2007; Küçükaydin et al., 2012) and stochastic demand (e.g., Uno et al., 2010). Several types of models have been proposed to quantify cost penalties that result as uncertain demand is realized in a competitive environment. Marianov, Rios, and Icaza (2008) maximized the capture of demand

by a firm that enters new facilities in a market involving competition from existing facilities. They employed an MNL model that uses travel distance and waiting time associated with an $M/M/m/K$ queue as competition factors. Y Zhang, Berman, and Verter (2012) maximized the capture of demand for preventive care centers. They developed two attraction models, an MNL model and another one that attracts demand according to a decreasing function of travel time. Their constraints limit the maximum number of facilities, minimum workload and maximum waiting time calculated using an $M/M/m$ queuing model to represent each facility.

The approaches that these two papers presented to incorporate waiting time are apparently limited to a one-period decision. For a multi-service, multi-period model, however, the best new location may not be the same from one time period to the next (or from one service to another). Patients may request appointments according to a Poisson process but arrivals for actual healthcare service are typically scheduled to be equally-spaced over time so that $M/M/m$ models overstate the actual waiting time and, therefore, find limited application in healthcare.

2.2.3 Optimizing expected excess revenue

A growing body of literature (e.g., (Muthuraman & Lawley, 2008) and (Liu, Ziya, & Kulkarni, 2010)) has dealt with scheduling patients (e.g., at clinics) so as to compensate for no shows. These papers assume that physicians will serve all patients scheduled each day, resulting in overtime if service for one or more patients is not completed by the end of the work day. The typical objective function is to maximize the expectation of excess revenue, the revenue minus the (penalty) cost of patient waiting times and physician overtime. Simulation models were used in numerical tests to evaluate scheduling policies using various distributions (e.g., uniform, exponential, gamma and lognormal) for physician service time and to allow no-shows.

These scheduling models better reflect the healthcare environment than existing queuing

models but are limited in our application, because daily scheduling is an operational-level issue and we deal with the strategic SSHFCP. In addition, data to describe uncertain physician service time is not available. We do note, however, that the objective function used in this healthcare setting highlights the relevancy of our objective function because it is the operational-level analog of the strategic-level one we use, which maximizes expected excess revenue, penalizing expected excess demand and capacity.

2.2.4 Location-allocation model in healthcare applications

Two review papers (Afshari & Peng, 2014; Daskin & Dean, 2004) comprehensively investigates the healthcare facility location problem. The location-allocation model is widely used in healthcare applications, including locating specialized services like MRI (Mahar, Bretthauer, & Salzarulo, 2011), emergency and ambulance vehicles (Cheu, Huang, & Huang, 2008; V. Knight, Harper, & Smith, 2012), organ transplantation hospitals (Bruni, Conforti, Sicilia, & Trotta, 2006), blood service (Şahin, Süral, & Meral, 2007), preventive care (Zhang, Berman, & Verter, 2009), specialty care for veterans (Benneyan, Musdal, Ceyhan, Shiner, & Watts, 2012; M. Côté et al., 2007) and pediatric hospitals (Malczewski & Ogryczak, 1990). Location-allocation model applied on general healthcare facility locations takes particular concerns on different aspects, such as effect of moving population on patient demand (Ndiaye & Alfares, 2008), the equity and efficiency of healthcare benefit allocation (Cho, 1998; Smith, Harper, Potts, & Thyle, 2009), multiple types of health insurance (Benneyan et al., 2012; Griffin, Scherrer, & Swann, 2008) and patient choice (Knight, Williams, & Reynolds, 2012).

Methodologies proposed to solve healthcare facility location-allocation models include Lagrangian relaxation (Syam & Côté, 2010), a quasi-Newton method (Cho, 1998), simulated annealing (Côté, Syam, Vogel, & Cowper, 2007), a hierarchical p-median heuristic (Smith et al.,

2009) and a meta-heuristic (Zhang et al., 2009). Methodologies that deal with patient-demand uncertainty include forecasting and estimating parameters (Côté et al., 2007; Griffin et al., 2008), a queueing model (Knight et al., 2012) and a reliability model (Cheu et al., 2008).

The column-generation heuristic we propose in this dissertation is a novel methodology to solve location-allocation problems, dealing with multi-period stochastic demands in subproblems. The approximation method we propose differs from existing heuristics in that it involves estimating the stochastic location-allocation cost in component of the objective function and providing a bounding procedure to assess solution quality.

CHAPTER III

DSCR*

This chapter presents our DSCR research. This chapter is organized in three sections. Section 3.1 presents our alternative DSCR formulations, addressing the first research objective. Section 3.2 describes results for two test scenarios that promote intuitive interpretation of model results, accomplishing the second research objective. Finally, Section 3.3 reports our computational evaluation, achieving the third and fourth research objectives.

3.1 DSCR model formulation

This section presents our two formulations of DSCR: a traditional MIP, DSCR-T; and a network-based model, DSCR-N. DSCR-T results from traditional logic to relate binary decision variables that prescribe reconfiguration; and DSCR-N utilizes a specialized network to relate binary decision variables to prescribe the same decisions.

3.1.1 Initial MIP formulation of DSCR

This section presents DSCR-T. We first describe our notation, including indices, index sets, parameters, and decision variables and relate assumptions. We use the terms *facility* and *location* interchangeably for convenience to indicate a supplier, plant, DC, or CZ.

---

* Reprinted with permission from "Computational comparison of two formulations for dynamic supply chain reconfiguration with capacity expansion and contraction." by Wilhelm, W. E., Han, X., and Lee, C., 2013, *Computers & Operations Research*, 40, 2340-2356, Copyright 2013 by Elsevier.

To focus discussion and streamline presentation, we fix suppliers, production plants and CZs in the "open" state. Following prior supply chain research, we assume that only DCs can be reconfigured. Our models can be adapted easily, however, to allow plants to be reconfigured and to allow suppliers to be selected and de-selected. Each CZ may experience demand for each product $p \in P$.

The index set L contains four index subsets of locations: suppliers, $L_S \subset L$; plants, $L_P \subset L$; DCs, $L_{DC} \subset L$; and CZs, $L_{CZ} \subset L$. Note that $L_S \bigcup L_P \bigcup L_{DC} \bigcup L_{CZ} = L$ and $L_e \bigcap L_{e'} = \varnothing$, where $e, e' \in E, e \neq e'$, and $E = \{1,2,3,4\}$, representing echelons, $\{S, P, DC, CZ\}$. T is the index set of time periods in the planning horizon.

A DSCR model must reconfigure the supply chain at locations represented by index set L_{DC} over the index set of time periods T. We comment that a facility is reconfigured *in a period* but assume that such actions take place *instantaneously at the beginning of the time period*. Once opened, a facility remains operating until closed. A facility can be opened at a location only once over the planning horizon and cannot be reopened once closed but that a facility can be expanded or contracted once each period after it has been opened and operating for at least one period. Closing in a period indicates that the facility is not operating (i.e., open) in that period or subsequent ones. We assume that the initial (i.e., at time instant 0) and final (i.e., at time instant |T|) inventories (backorders) are zero and each supplier has unlimited capacity. Because single sourcing provides significant advantages (e.g., cost, diminished opportunity for error, consistency), we require that each CZ be supplied with all relevant products by just one DC.

We use K to denote the index set of capacity alternatives, n_j to denote the number of capacity modules associated with alternative j, and β to denote the maximum permissible capacity. We structure capacity alternatives by defining each as an integer multiple of the capacity

of a basic module and define α as the number of products that each module can produce in a period.

The discounted fixed costs associated with each capacity decision are defined as

G_{lt}^{C} : cost to close one unit of capacity

G_{lt}^{D} : cost to contract one unit of capacity

G_{lt}^{E} : cost to expand one unit of capacity

G_{lt}^{M} : cost to operate one unit of capacity

G_{lt}^{O} : cost to open one module.

Discounted variable costs associated with material flows are

$U_{ll't}^{p}$: cost / hundred-miles to ship a unit of product p from location l to location l' in period t

$H_{lt}^{p}(Q_{lt}^{p})$: cost to hold a unit product p in inventory (backorder) at location l at the end of period t

$R_{\ell t}^{p}$: cost to purchase a unit of product p from outside supplier l in period t (i.e., outsourcing).

The demand $b_{\ell t}^{p}$ for product $p \in P$ at facility (i.e., CZ) $l \in L_{CZ}$ during period t must be satisfied by production in the current period, by drawing from inventory, by incurring backorders, and/or by outsourcing. δ_{ℓ}^{p} denotes the workload required to process one unit of product p at facility $l \in L$. The maximum material flow on transportation link ll' in period t is limited by the upper bound capacity $V_{ll't}$, where $l \in L_e$ and $l' \in L_{e+1}$. $d_{ll'}$ denotes the distance between $l \in L_e$ and $l' \in L_{e+1}$.

The available budget for all fixed costs over the entire planning horizon at location $l \in L$

is specified by B_t. The maximum total number of all expansions and contractions for facility l over the planning horizon is limited to N_l^T.

DSCR-T involves five types of binary variables:

a_{jlt}=1 if location l is expanded or contracted to capacity alternative j in time period t, 0 otherwise

c_{lt}=1 if (open) facility l is closed at the start of period t, 0 otherwise

o_{jlt}=1 if location l is opened at capacity alternative j in time period t, 0 otherwise

m_{jlt}=1 if location l is operating at capacity alternative j in time period t, 0 otherwise

$s_{l'lt}$=1 if the transportation link from l to l' is available in period t (i.e., both facilities are operating).

DSCR-T involves five types of continuous variables that prescribe capacities,

$\bar{\bar{k}}_{lt}$ = amount of capacity used at location l in time period t

$\tilde{k}_{lt}$ = amount of capacity configured at location l in time period t

$\bar{k}_{lt}$ = amount of capacity expansion at location l in time period t

$\underline{k}_{lt}$ = amount of capacity contraction at location l in time period t

k_{lt}^C = amount of capacity upon closing location l in time period t,

and four that prescribe material flow, inventories, backorders, and outsourcing,

r_{lt}^p = the amount of product p outsourced by facility l for use in period t

v_{lt}^p = the amount of product p backordered in period t

$x_{ll't}^p$ = the amount of product p shipped from l to l' in period t

y_{lt}^p = the amount of product p held in inventory at l from period t to $t+1$.

Depending on the types of facilities at locations l and l', the flow of product p, $x^p_{ll't}$, can be an amount that is purchased from a supplier, processed, shipped on a link, or distributed from a DC to a CZ.

We now present our **MIP model** $(DSCR-T)$

$$Z^* = \min \sum_{t\in T}\sum_{l\in L_{PDC}}\left(G^O_{lt}\sum_{j\in K} n_j o_{jlt} + G^C_{lt}k^C_{lt} + G^M_{lt}\tilde{k}_{lt} + G^E_{lt}\bar{k} + G^D_{lt}\underline{k}\right)$$
$$+\sum_{t\in T}\sum_{l\in L}\sum_{p\in P}\left(H^p_{lt}y^p_{lt} + Q^p_{lt}v^p_{lt} + R^p_{lt}r^p_{lt} + \sum_{l'\in \bar{L}, l'\neq l} d_{ll'}U^p_{ll't}x^p_{ll't}\right) \quad (1)$$

s.t. $\sum_{j\in K} o_{jlt} + c_{lt} + \sum_{j\in K} a_{jlt} \le 1$ $\quad l\in L_{DC}, t\in T$ (2)

$\sum_{t\in T}\sum_{j\in K} o_{jlt} \le 1$ $\quad l\in L_{DC}$ (3)

$c_{lt} \le \sum_{s=1}^{t-1}\sum_{j\in K} o_{jls}$ $\quad l\in L_{DC}, t\in T\setminus\{1\}$ (4)

$\sum_{j\in K} m_{jlt} = \sum_{s=1}^{t}\left(\sum_{j\in K} o_{jls} - c_{ls}\right)$ $\quad l\in L_{DC}, t\in T$ (5)

$m_{jlt} \ge o_{jlt} + a_{jlt}$ $\quad j\in K, l\in L_{DC}, t\in T$ (6)

$m_{jl,t-1} + a_{jlt} \le 1$ $\quad j\in K, l\in L_{DC}, t\in T\setminus\{1\}$ (7)

$\sum_{t\in T}\left(G^O_{lt}\sum_{j\in K} n_j o_{jlt} + G^C_{lt}k^C_{lt} + G^M_{lt}\tilde{k}_{lt} + G^E_{lt}\bar{k} + G^D_{lt}\underline{k}\right) \le B_l$ $\quad l\in L_{DC}$ (8)

$\sum_{t\in T}\sum_{j\in K} a_{jlt} \le N^T_l$ $\quad l\in L_{DC}$ (9)

$\sum_{l\in L_e} x^p_{ll't} - \sum_{l"\in L_{e-1}} x^p_{l"lt} + y^p_{l,t-1} - y^p_{lt} + v^p_{l,t+1} - v^p_{lt} + r^p_{lt} = b^p_{lt}$ $\quad p\in P, l\in L_e, t\in T$ (10)

$s_{ll't} \le \sum_{j\in K} m_{jlt}$ $\quad l\in L_e, l'\in L_{e+1}, t\in T$ (11)

$s_{ll't} \le \sum_{j\in K} m_{jl't}$ $\quad l\in L_e, l'\in L_{e+1}, t\in T$ (12)

$\sum_{l\in L_{DC}} s_{ll't} = 1$ $\quad l\in L_{CZ}, t\in T$ (13)

$$\sum_{p\in P} x^p_{ll't} \le V_{ll't} s_{ll't} \qquad l\in L_e, l'\in L_{e+1}, t\in T \qquad (14)$$

$$\bar{\bar{k}}_{lt} = \sum_{l'\in L_{e+1}, l'\neq l} \sum_{p\in P} \delta^p_l x^p_{ll't} \qquad l\in L_{DC}, t\in T \qquad (15)$$

$$\bar{\bar{k}}_{lt} \le \beta\left(\sum_{j\in K} m_{jlt}\right) \qquad l\in L_{DC}, t\in T \qquad (16)$$

$$\tilde{k}_{lt} = \sum_{j\in K} \alpha n_j m_{jlt} \qquad l\in L_{DC}, t\in T \qquad (17)$$

$$\bar{\bar{k}}_{lt} \le \tilde{k}_{lt} \qquad l\in L_{DC}, t\in T \qquad (18)$$

$$k^C_{lt} \ge \tilde{k}_{l,t-1} - \beta(1-c_{lt}) \qquad l\in L_{DC}, t\in T\setminus\{1\} \qquad (19)$$

$$k^C_{lt} \le \beta c_{lt} \qquad l\in L_{DC}, t\in T\setminus\{1\} \qquad (20)$$

$$\bar{k}_{lt} \ge \sum_{j\in K} \alpha n_j a_{jlt} - \tilde{k}_{l,t-1} \qquad l\in L_{DC}, t\in T\setminus\{1\} \qquad (21)$$

$$\bar{k}_{lt} \le \beta \sum_{j\in K} a_{jlt} \qquad l\in L_{DC}, t\in T\setminus\{1\} \qquad (22)$$

$$\underline{k}_{lt} \ge \tilde{k}_{l,t-1} - \sum_{j\in K} \alpha n_j a_{jlt} - \beta\left(1-\sum_{j\in K} a_{jlt}\right) \qquad l\in L_{DC}, t\in T\setminus\{1\} \qquad (23)$$

$$\underline{k}_{lt} \le \beta \sum_{j\in K} a_{jlt} \qquad l\in L_{DC}, t\in T\setminus\{1\} \qquad (24)$$

$$\sum_{j\in K} n_j o_{jlt} + \bar{k}_{lt} - \underline{k}_{lt} - k^C_{lt} = \tilde{k}_{lt} - \tilde{k}_{l,t-1} \qquad l\in L_{DC}, t\in T\setminus\{1\} \qquad (25)$$

$$c_{l1} = 0, \sum_{j\in K} a_{jl1} = 0 \qquad l\in L_{DC} \qquad (26)$$

Objective (1) minimizes total cost. Fixed costs associated with facility l include charges for opening G^O_{lt}, closing G^C_{lt}, operating G^M_{lt}, expanding G^E_{lt} and contracting G^D_{lt}. Note that the model does not require that a facility be closed at the end of the planning horizon. Variable costs accrue for holding inventories H^p_{lt}, incurring backorders Q^p_{lt}, outsourcing R^p_{lt}, and transporting products $U^p_{ll't}$.

Inequalities (2) ensure that at most one decision is prescribed during time period t to open

or close facility l, or expand or contract its capacity, preventing two or more such decisions in the same period. Inequalities (3) allow at most one opening at location l over the planning horizon. Constraints (4) allow facility l to be closed in period t only if it were opened in a previous period. Equalities (5) specify that facility l is operating in period t ($M_{lt} = 1$) if it has been opened, but not closed, by that time. Facility l is operating from the time period it is opened until the period it is closed. Inequalities (6) assure that, if facility l is opened at capacity level j or adjusted to capacity level j in period t, then m_{jlt} should be 1 and inequalities (7) assure that, if facility l is operating at capacity alternative j in time period t-1, it cannot be both expanded and contracted to the same alternative j in time period t. Inequalities (8) impose a limit on fixed cost expenditures for reconfiguration. This is plausible in that each potential location might have a limited budget allocation according to the overall long-term plan of the enterprise. Inequalities (9) limit the total number of capacity expansions and contractions allowed over the planning horizon at location l to N_l^T, representing practical management considerations.

Flow conservation constraints (10) ensure that demands in all CZs are met each period. Demands for end-products occur only in CZs. Nodes representing suppliers are sources of flow and thus have positive b_{lt}^p values, nodes representing CZs are flow sinks and have negative b_{lt}^p values, and nodes in intermediate echelons represent plants (or DCs) that process (or store) a product and can be viewed as transshipment nodes, each with $b_{lt}^p = 0$. It is realistic to assume that manufacturing plants and DCs can hold stock from a previous period, receive flow from an outside supplier, and receive backorders from a subsequent period. For each product p, the summation of flow out to downstream nodes, inventory from period t-1, outsourcing in period t, and input of backorders from t+1 minus the summation of flow in from each intermediate node from upstream facilities, inventory at the end of period t, and backorders at the end of period t sum to b_{lt}^p.

Inequalities (11) ((12)) assure that the transportation link from l to l' can be used only if facilities at locations l and l' are both operating. Single sourcing constraints (13) require that only one DC supplies customer zone $l' \in L_{cz}$ with all types of products, representing a common practice in industry. Constraints (14) allow product p to be transported from location l to l' only in time periods during which the transportation link from l to l' is established.

Equalities (15) defines $\overline{\overline{k}}_{lt}$, the amount of capacity used at location l in time period t, relative to $x^p_{ll't}$, the amount shipped out, and inequalities (16) assure that, if facility l is not operating, there is no flow through the facility. Equalities (17) define $\tilde{k}_{lt}$, the amount of capacity configured at location l in time period t, assuring that $\tilde{k}_{lt}$ is zero if facility l is not operating in time period t. Inequalities (18) relate $\overline{\overline{k}}_{lt}$ and $\tilde{k}_{lt}$. Inequalities (19) define the capacity at closing: if c_{lt} =1, then $k^C_{lt} = \tilde{k}_{l,t-1}$; otherwise, $\tilde{k}_{l,t-1} - \beta \le 0$ because β is the maximum capacity a facility can attain, so $k^C_{lt} \ge 0$. Inequalities (20) force c_{lt}=1 if the facility at location l is closed in time period t, and force $k^C_{lt} = 0$ if facility at location l is not closed in time period t. Inequality (21) define $\bar{k}_{lt}$, the amount of capacity expansion in time period t. If there is a capacity adjustment at location l in time period t, $\sum_{j \in K} a_{jlt} = 1$; and, if the adjustment is an expansion, $\sum_{j \in K} n_j a_{jlt} - \tilde{k}_{l,t-1} > 0$, so $\bar{k}_{lt} = \sum_{j \in K} n_j a_{jlt} - \tilde{k}_{l,t-1}$; otherwise, $\bar{k}_{lt} = 0$ if the adjustment is a contraction. When there is no adjustment, $-\tilde{k}_{l,t-1} \le 0$, so $\bar{k}_{lt} = 0$. To assure that $\bar{k}_{lt} > 0 \Rightarrow \sum_{j \in K} a_{jlt} = 1$, constraints (22) are required in addition to (20), which only provide a lower bound on each $\bar{k}_{lt}$. Without (22), $\bar{k}_{lt}$ can still take a positive value with $\sum_{j \in K} a_{jlt} = 0$. Inequalities (23) and (24) work together to define $\underline{k}_{lt}$, the amount of capacity contraction in time period t, in a manner analogous to the way (21)

and (22) define $\bar{k}_{lt}$. Inequalities (25) balance the actions of reconfiguration with the change of capacity from *t*-1 to *t*. They disallow the case in which $\tilde{k}_{lt} - \tilde{k}_{l,t-1} > 0$, but both $\sum_{j \in K} o_{jlt} = 0$ and $\sum_{j \in K} a_{jlt} = 0$. They also disallow the case in which $\tilde{k}_{lt} - \tilde{k}_{l,t-1} < 0$, but both $c_{lt} = 0$ and $\sum_{j \in K} a_{jlt} = 0$. If (25) were not included, the cost of reconfiguration may be computed incorrectly.

Equalities (26) establish certain initial conditions for $l \in L_{DC}$. For clarity, we do not include constraints that zero initial $y_{l0}^{p}\left(v_{l0}^{p}\right)$ and final $y_{l,|T|+1}^{p}\left(v_{l,|T|+1}^{p}\right)$ inventories (backorders); impose binary restrictions on a_{jlt}, c_{lt}, m_{jlt}, o_{jlt} and $s_{ll't}$; or impose non-negativity on material flow variables $r_{lt}^{p}, v_{lt}^{p}, x_{ll't}^{p}$ and y_{lt}^{p} or on capacity defining variables $\bar{k}_{lt}$, $\underline{k}_{lt}$, k_{lt}^{C}, $\bar{\bar{k}}_{lt}$ and $\tilde{k}_{lt}$.

We note that earlier papers have used different types of constraints and different rules to manage facility configuration. Instead of (4) and (5), (Wesolowsky & Truscott, 1975) formulated a dynamic facility location model that defines a relationship between opening, operating, and closing variables:

$$C_{lt} - \sum_{k \in K_l} O_{klt} = M_{l,t-1} - M_{lt} \qquad l \in L, t \in T .$$

Although (Wesolowsky & Truscott, 1975) did not mention it explicitly, $M_{l,t=0}$ must be fixed to zero for each location at which a facility is not operating in time period 0 before opening it at the beginning of time period 1. Without this boundary condition $(i.e., M_{l,t=0} = 0)$, the model allows a facility to operate, even though it has not been opened. Our DSCR-T model assumes that no potential facility is open at the beginning of the planning horizon. If a facility is to operate (i.e., $M_{lt} = 1$), it must first be opened (i.e., $\sum_{k \in K_l} O_{klt} = 1$); thus, if we employ the constraint of (Wesolowsky & Truscott, 1975) instead of (4) and (5), we would need to include constraints $M_{l,t=0} = 0$. If a facility were operating before the beginning of the planning horizon, the boundary

condition of $M_{l,t=0}=1$ would be required. Our preliminary tests showed that our DSCR-T model solves faster using constraints (4) and (5) than with the constraint of (Wesolowsky & Truscott, 1975).

Another approach for prescribing facility opening and closing is to define two location index sets: one set for facilities that can be opened and expanded, and the other set existing facilities that can be contracted and closed (e.g.,(Yolanda Hinojosa et al., 2008; Yolanda Hinojosa et al., 2000; M. T. Melo et al., 2006; P. N. Thanh et al., 2008; Van Roy & Erlenkotter, 1982)). In contrast to our model, this approach does not allow the same facility to expanded and contracted over the planning horizon once opened (and before closure).

3.1.2 Network-based formulation of DSCR

We now present DSCR-N, using notation defined earlier and some additional symbols. We use $K=\{1,\ldots,n\}$ to denote the index set of alternative capacities that might be provided at location $l \in L_{DC}$ and let $n=|K|$. The amount of capacity provided by alternative $k \in K$ is given by $\bar{U}_k$. $\bar{U}_k = \alpha n_k$ for $k \in K$, $\bar{U}_k = 0$ for $k=0$ and $k=n+1$. Let $\beta = \bar{U}_n$ denote the maximum capacity of any alternative at location l.

In DSCR-N, the net capacity after expansion or contraction is defined by the capacity alternative prescribed in time period $t+1$. We define binary decision variable $z_{jklt}=1$ if the facility at location $l \in L_{DC}$ utilizes capacity alternative j in period t and expands or contracts to use capacity alternative k in period $t+1$. Instead of using four binary variables (i.e., o_{jlt}, m_{jlt}, a_{jlt}, and $c_{\ell t}$) for location l in period t, depending on the combination of j and t, we use variable z_{jklt} to represent reconfiguration over time at location l. Note that variable z_{jklt} is defined in our

model only for $l \in L_{DC}$.

Let G_{jklt}, the cost of changing capacity alternative j in time period t to capacity alternative k in time period t+1 at location l, be associated with decision variable z_{jklt}. Then

$$G_{jklt} = \begin{cases} n_k G_{lt}^O & \text{for } j = 0, k \in K, l \in L, t \in T \bigcup \{0\} \\ \alpha n_j G_{lt}^C + \alpha n_j G_{lt}^M & \text{for } j \in K, k = n+1, l \in L, t \in T \setminus \{1\} \\ \alpha\left(n_k - n_j\right) G_{lt}^E + \alpha n_j G_{lt}^M & \text{for } j, k \in K, k \geq j, l \in L, t \in T \\ \alpha\left(n_j - n_k\right) G_{lt}^D + \alpha n_j G_{lt}^M & \text{for } j, k \in K, k < j, l \in L, t \in T \\ \alpha n_j G_{lt}^M & \text{for } j \in K, k = n+1, l \in L, t = \tau. \end{cases}$$

Other constraints and variables are the same as in DSCR-T.

Before presenting DSCR-N, we give an example of the capacity alternative network for location l, which employs binary variable z_{jklt}. Figure 1 depicts the directed, acyclic network $G\left(\bar{N}_l, \bar{A}_l\right)$ associated with location l in which $\bar{N}_l$ is the index set of nodes representing capacity alternatives and $\bar{A}_l$ is the index set of (directed) arcs that connect capacity alternatives that are feasible relative to constraints (1)-(6). Each arc $(i, j, t) \in \bar{A}_l$ denotes a feasible reconfiguration decision at location l. As a result, the structure of network $G\left(\bar{N}_l, \bar{A}_l\right)$ imposes these constraints by introducing a "flow conservation" constraint for variable z_{jklt} to prescribe an optimal capacity-alternative path for location l over all $t \in T$.

Level (i.e., row of nodes) t represents time period t. $\bar{N}_l$ includes a (dummy) start node in level $t = 0$ and a (dummy) end node in level $\tau = |T| + 1$. Nodes in the first (last) column represent the decision to not open (close) facility l in time period t and nodes in each other column represents an alternative capacity that can be prescribed for facility l.

Each arc $(i, j, t) \in \bar{A}_l$ points from capacity alternative node i in level t to capacity

alternative node j in level $t+1$. The network incorporates five types of arcs, each with a corresponding type of cost for opening, operating, expanding, contracting, and closing, respectively. The relevant cost is applied to each of the corresponding arcs. An optimal path from the start node defines a capacity alternative for each time period, equivalently, a series of reconfiguration decisions (e.g., see the two possible paths composed of arcs represented by solid and dash lines in Figure 1).

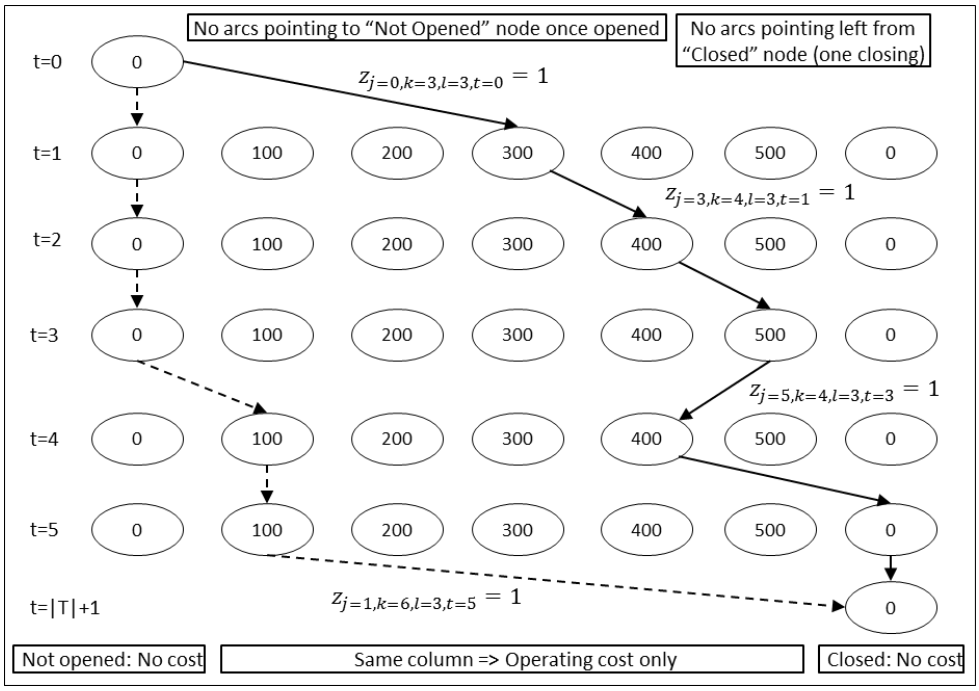

Figure 1. An example DSCR-N capacity network with l=3, $|T|$=5

We now present network-based model DSCR-N:

$$Z^* = \min \sum_{\ell \in L_{PDC}} \sum_{(j,k,t) \in \bar{A}_l} G_{jklt} z_{jklt} + \sum_{t \in T} \sum_{l \in L_e} \sum_{p \in P} [H_{lt}^p y_{lt}^p + Q_{lt}^p v_{lt}^p + R_{lt}^p r_{lt}^p + \sum_{l' \in L_{e+1}, l' \neq l} U_{ll't}^p d_{ll'} x_{ll't}^p] \qquad (27)$$

s.t. Constraints (10), (13), and (14)

$$\sum_{(j,k,t)\in\bar{A}_l} z_{kjlt} - \sum_{(j,k,t)\in\bar{A}_l} z_{jkl,t-1} = \rho_{klt} \qquad l \in L_{DC},\ k \in K,\ t \in T \bigcup \{0,\tau\} \qquad (28)$$

$$\sum_{(j,k,t)\in\bar{A}_l} G_{jklt} z_{jklt} \le B_l \qquad l \in L_{DC} \qquad (29)$$

$$\sum_{j,k\in K, j\neq k,(j,k,t)\in\bar{A}_l} z_{jklt} \le N_l^T \qquad l \in L_{DC} \qquad (30)$$

$$s_{ll't} \le \sum_{j\in K,(j,k,t)\in\bar{A}_l} z_{jklt} \qquad l \in L_e, l' \in L_{e+1}, t \in T \qquad (31)$$

$$s_{ll't} \le \sum_{j\in K,(j,k,t)\in\bar{A}_{l'}} z_{jkl't} \qquad l \in L_e, l' \in L_{e+1}, t \in T \qquad (32)$$

$$\sum_{l'\in L_{CZ}} \sum_{p\in P} \delta_l^p x_{ll't}^p \le \sum_{j\in K\cup\{0,n+1\},\ (j,k,t)\in\bar{A}_l} \bar{U}_j z_{jklt} \qquad l \in L_{DC}, t \in T \qquad (33)$$

Objective (27) minimizes total costs as (1) does in DSCR-T. Flow conservation constraint (28) ensures that reconfiguration decisions are consistent over the planning horizon, paralleling (2)-(7). These constraints formulate the shortest-cost capacity alternative path as a network flow problem in which one unit of flow originates at the start node, travels through the network, and terminates at the end node. This requires a flow balance at each node, so that the flow out of node k minus the flow into it equals ρ_{klt}, where $\rho_{k=0,l,t=0} = +1$, $\rho_{k=n+1,l,t=\tau} = -1$, and $\rho_i = 0$ for $i \in \bar{N}_l \setminus \{start\ node \bigcup end\ node\}$.

Inequalities (29) limit the amount of capital that can be invested at location l over the entire planning horizon for all reconfiguration decisions and (30) limit the number of expansions and contractions that can be made at a location over the planning horizon. The RHSs of (31) and (32) are either zero or one, contributing to forming a tight linear relaxation of DSCR-N. Inequalities (33) ensure that the workload at facility l does not exceed this facility's capacity in each time period.

To streamline presentation, we omit restrictions that require variables z_{jklt} to be binary.

Constraints (10), (13) and (14) are the same as in DSCR-T. Section 4 compares the sizes of these two models (i.e., numbers of constraints and continuous and binary variables) and the run time each requires.

3.2 Demonstration of DSCR model use

We now describe model application to two demand scenarios that promote an intuitive interpretation of our models by demonstrating how time varying demand can lead to reconfiguration over a $|T| = 8$-period planning horizon. We deal with the four-echelon supply chain (i.e., $|EC|$ =4), a commonly studied structure. The first subsection describes the data we synthesize to describe this case study, and the second (third) gives results for the first (second) scenario.

3.2.1 Data

The data we use allows for a ready, intuitive interpretation of results. As depicted in Table 1, Scenario 1 represents a product life cycle for which demand increases in early periods and then decreases as obsolescence approaches, and Scenario 2 represents an economic downturn then recovery with demand decreasing and then increasing.

Table 1. Demand scenario over 8 years

Year	Scenario 1 Product life cycle	Scenario 2 Economic cycle
1	850,000	3,400,000
2	1,700,000	2,550,000
3	2,550,000	1,700,000
4	3,400,000	850,000
5	3,400,000	850,000
6	2,550,000	1,700,000
7	1,700,000	2,550,000
8	850,000	3,400,000

We assume that each module provides the capacity to produce $\alpha = 1{,}000{,}000$ products each time period and deal with four capacity alternatives, $K = \{1,...,4\}$, so that alternative j entails a capacity of $\alpha n_j = \alpha j$ and the maximum capacity is $\beta = |K|\alpha$. We assume that the costs for each module are based on its opening (or expansion) cost of \$2,000,000: 15% of construction cost to operate each year, 10% to contract, and 5% to close. This gives the following costs per unit of capacity:

G_{lt}^O : cost to open one module, $G_{lt}^O = \$2{,}000{,}000$

G_{lt}^C : cost to close one unit of capacity, $G_{lt}^C = \$0.01 = \$100{,}000/\alpha$

G_{lt}^M : cost to operate one unit of capacity, $G_{lt}^M = \$0.3 = \$300{,}000/\alpha$

G_{lt}^E : cost to expand one unit of capacity, $G_{lt}^E = \$2.00 = \$2{,}000{,}000/\alpha$

G_{lt}^D : cost to contract one unit of capacity, $G_{lt}^D = -\$0.2 = -\$200{,}000/\alpha$.

Diagonal elements in Table 2 give the cost to construct each alternative capacity, elements above (below) the diagonal give the cost to expand (contract) from capacity alternative *i* to capacity alternative *j*.

Table 2. Capacity alternative costs

Alt.	1	2	3	4
1	\$2,000,000	\$2,000,000	\$4,000,000	\$6,000,000
2	\$200,000	\$4,000,000	\$2,000,000	\$4,000,000
3	\$400,000	\$200,000	\$6,000,000	\$2,000,000
4	\$600,000	\$400,000	\$200,000	\$8,000,000

The columns of Table 3 give, for each capacity alternative, the fixed cost to open, the annual operating cost, the fixed cost to close, and the annual capacity. Column 6 gives the demand

for which the capacity (column 5) would have a utilization of 85% (column 7), a value commonly targeted in industry. For example, if capacity alternative 1 (capacity = 1,000,000) were prescribed to correspond to demand in time period 1 under scenario 1 (demand = 850,000), the utilization of the facility would be 85%.

Each scenario deals with 5 locations in each echelon, where $L = \{1,\ldots,20\}$: $L_S = \{1,\ldots,5\}$, $L_P = \{6,\ldots,10\}$, $L_{DC} = \{11,\ldots,15\}$ and $L_{CZ} = \{16,\ldots,20\}$. We take the 20 most populace cities in the U.S. to be the locations in set L and assign them in groups of five, in order, to the respective four echelons. We determine the cost of transportation from one location to another as the actual distance between the cities times \$0.01/unit of product/hundred miles. For each scenario and each time period, we assign one fifth of the total demand to each of the five CZs.

Table 3. Data for each capacity alternative

Cap. Alt.	Fixed cost to open	Annual operating cost	Fixed cost to close	Capacity #products/yr	Demand #products/yr	Utilization= 100(Demand/ Capacity)
1	\$2,000,000	\$300,000	\$100,000	1,000,000	850,000	85%
2	\$4,000,000	\$600,000	\$200,000	2,000,000	1,700,000	85%
3	\$6,000,000	\$900,000	\$300,000	3,000,000	2,550,000	85%
4	\$8,000,000	\$1,200,000	\$400,000	4,000,000	3,400,000	85%

Table 4. Locations

Suppliers	New York	Los Angeles	Chicago	Houston	Philadelphia
Plants	San Jose	Indianapolis	San Francisco	Jacksonville	Columbus
DCs	Washington	El Paso	Seattle	Denver	Nashville
CZs	New Orleans	Las Vegas	Cleveland	Long Beach	Albuquerque

3.2.2 Scenario 1 results

Both models (i.e., DSCR-T and DSCR-N) prescribe the same solution for scenario 1, which opens

and operates DC 12 at El Paso, TX, for $4 \leq t \leq 5$ and DC 15 at Nashville, TN, for $1 \leq t \leq 8$, as shown by Figure 2. The solution agrees with intuition, providing sufficient - but not excess - capacity that allows the utilization of total DC facility capacity to be 85% in each time period, a common target in industry. That is, the amount of capacity used, $\sum_{l \in L_{DC}} \bar{\bar{k}}_{lt}$, is 85% of the capacity provided each period, $\sum_{l \in L_{DC}} \tilde{k}_{lt}$. Run times are 103.44 seconds for the DSCR-N model and 150.68 seconds for the DSCR-T model.

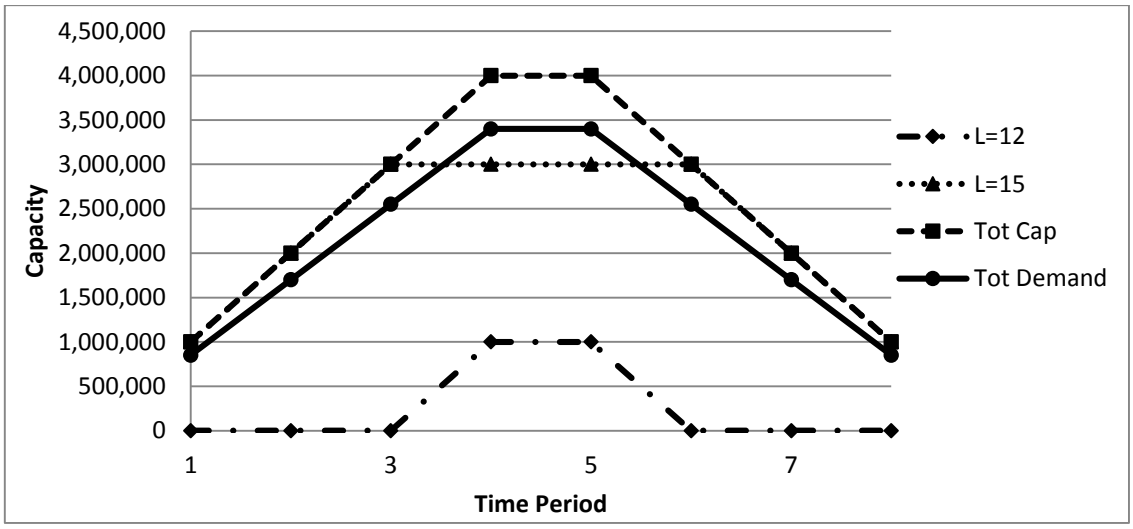

Figure 2. Capacity configuration of DCs vs scenario 1 demand

To lend further insight into how the decision variables in traditional model DSCR-T interact, we report solution values for primary variables in each of the time periods in Tables 5-8. Table 5 shows that DC 15 is never closed (i.e., $c_{lt} = 0$ for $1 \leq t \leq 8$); DC capacity is expanded (contracted) in increments of 1 million in periods 2 and 3 (7 and 8), $\bar{k}_{lt}$, ($\underline{k}_{lt}$); and that $k_{lt}^{C} = 0$ for

$1 \le t \le 8$, because DC 15 is never closed. Table 6 gives values for o_{jlt}, showing that DC 15 is opened at capacity alternative 1 in time period 1 (i.e., $o_{1,15,1} = 1$). Table 7 gives values of a_{jlt}=1 if location *l* is expanded or contracted to capacity alternative *j* in time period *t*, showing that DC 15 is expanded (i.e., $a_{2,15,2} = a_{3,15,3} = 1$) (contracted; i.e., $a_{2,15,7} = a_{1,15,8} = 1$) in periods 2 and 3 (7 and 8). Finally, Table 8 gives values for m_{jlt}, showing that DC 15 is operating at some capacity alternative j in each time period (i.e. $m_{j,15,t} = 1$ for $1 \le t \le 8$).

3.2.3 Scenario 2 results

Both models prescribe the same solution for scenario 2, opening and operating DC 12 at El Paso, TX, for 1≤ t ≤8 and DC 15 at Nashville, TN, for 1≤ t≤8 as shown in Figure 3. The utilization of DC capacity is 85% in the first and last time periods but the capacity level does not change over time, even though demand changes. This result is optimal because demand requires a capacity of at least 4,000,000 in the first and last periods; but the cost of expansion, G_{lt}^{E}, is more than the cost benefit of contraction, $G_{lt}^{D} + G_{lt}^{M}$. Thus, it is not economical to decrease, then increase, capacity levels in this scenario. Run times are 38.36 seconds for DSCR-N and 51.79 seconds for DSCR-T.

Using data from Tables 2 and 3, both scenarios prescribe the same locations for DCs: DC 12 at El Paso, TX, and DC 15 at Nashville, TN are opened and operated during the planning horizon, while the other DC candidate sites are never opened (see Figure 4).

To investigate the sensitivity of the solution to the value of the expansion cost parameter, G_{lt}^{E}, we decrease it to $400,000, which is 1/5 of the original expansion cost, so that it compares favorably with $G_{lt}^{D} + G_{lt}^{M}$. Both models (i.e., DSCR-T and DSCR-N) prescribe the same solution, which opens DC 15 at Nashville, TN, for 1≤ t ≤8 and keeps other DCs closed throughout the

planning horizon. Figure 5 shows that the capacity level changes with demand, first decreasing, then increasing. The utilization is 85% in each of six time periods (1, 2, 3, 6, 7, 8). The reason that the capacity level does not change with demand in time periods 4 and 5 is that we limited the number of expansions and contractions to 4. Run times are 255.01 seconds for DSCR-N and 3098.24 seconds for DSCR-T.

Table 5. Capacity decisions for DC 15 over time [(m: million)]

Dec. Var.	Time t							
	1	2	3	4	5	6	7	8
c_{lt}	0	0	0	0	0	0	0	0
$\bar{\bar{k}}_{\ell t}$	0.85m	1.7m	2.55m	2.72m	2.72m	2.55m	1.7m	0.85m
$\tilde{k}_{\ell t}$	1m	2m	3m	3m	3m	3m	2m	1m
$\bar{k}_{\ell t}$	0	1m	1m	0	0	0	0	0
$\underline{k}_{\ell t}$	0	0	0	0	0	0	1m	1m
k^c_{lt}	0	0	0	0	0	0	0	0

Table 6. o_{jit} decisions to open DC 15 over time

Cap. Level j	Time t							
	1	2	3	4	5	6	7	8
1	1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0

Table 7. a_{jit} decisions to expand and contract capacity at DC 15 over time

Cap. Level j	Time t							
	1	2	3	4	5	6	7	8
1	0	0	0	0	0	0	0	1
2	0	1	0	0	0	0	1	0
3	0	0	1	0	0	0	0	0
4	0	0	0	0	0	0	0	0

Table 8. m_{jit} decisions to operate DC 15 over time

Cap. Level j	Time t							
	1	2	3	4	5	6	7	8
1	1	0	0	0	0	0	0	1
2	0	1	0	0	0	0	1	0
3	0	0	1	1	1	1	0	0
4	0	0	0	0	0	0	0	0

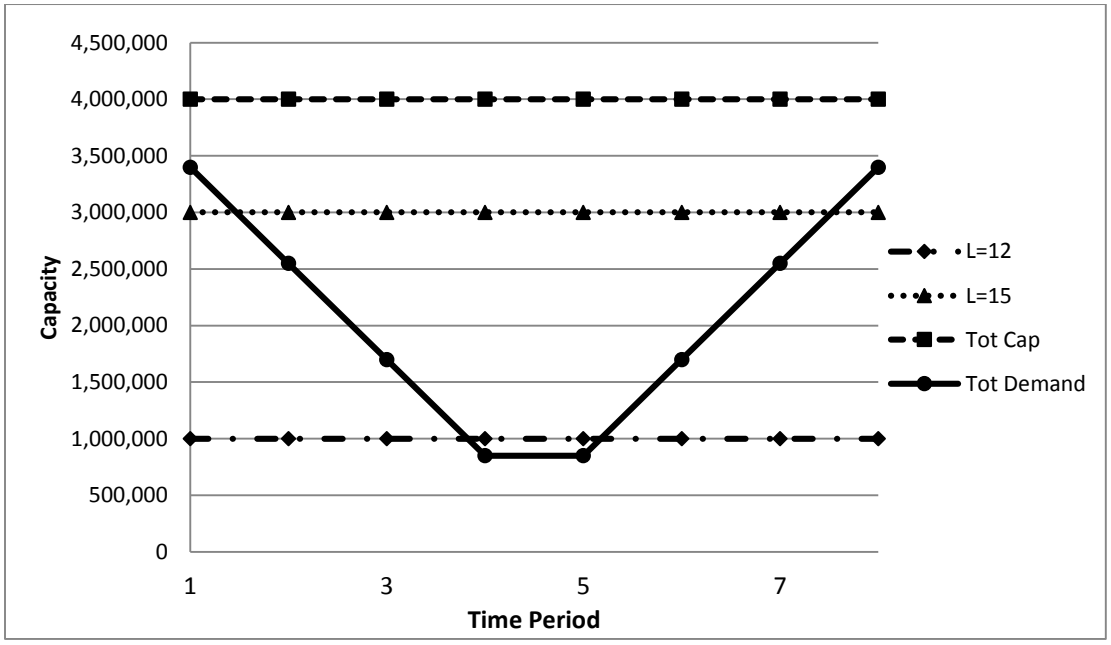

Figure 3. Capacity configuration of DCs vs scenario 2 demand

3.3 Computational evaluation

We now relate our primary computational experiments, designed to fulfill research objective 3 by identifying the computational characteristics of the traditional and network-based models and by determining if one offers superior solvability. We program our DSCR-T and DSCR-N formulations in *AMPL 9.0*® and use the *IBM ILOG CPLEX12.1*® branch-and-bound solver with

Figure 4. Solution description

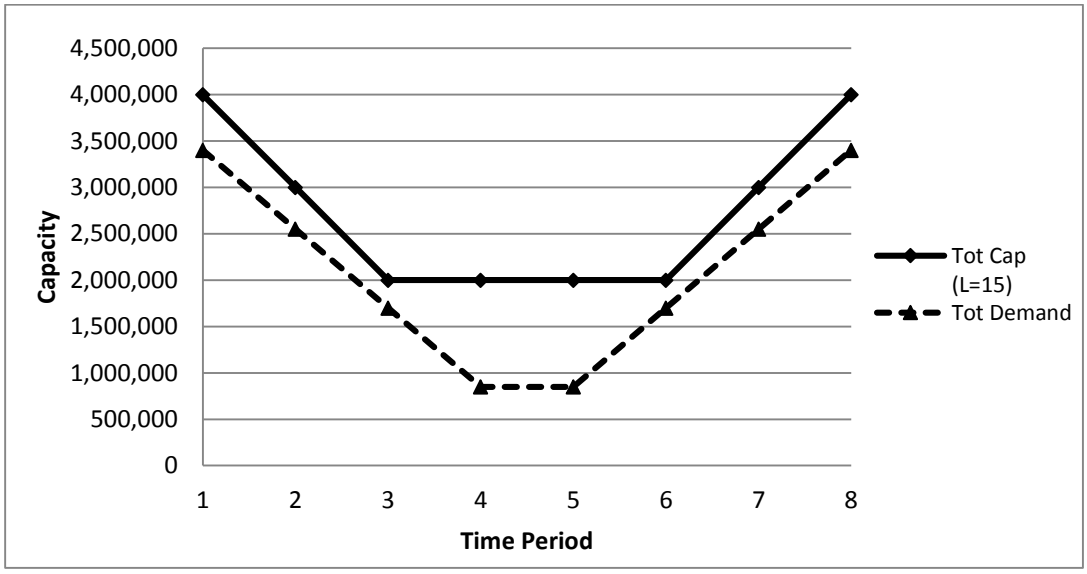

Figure 5. Capacity configuration of DCs vs scenario 2 demand with small expansion cost

default settings. We perform all computational tests on an Intel Xeon CPU E5620 @ 2.40GHz (2 processors), a 64-bit operating system, with 12.0GB RAM. This section comprises three subsections; design of experiments, test results, and analysis of results.

3.3.1 Design of experiments

Our experiment employs six factors, each with a selected number of levels: (1) the number of demand scenarios (Scenarios 1 & 2), (2) the number of potential locations for facilities (20, 24, 28, 32, 36, 40), the number of DCs (5, 6, 7, 8, 9, 10), the maximum number of capacity expansions and contractions (3, 4), the number of capacity alternatives at each location (4, 5), and the number of time periods in the planning horizon (8, 10). Each experiment deals with a single product type.

The first capacity alternative is 1,000,000, and each successive alternative increments capacity by 1,000,000. In scenario 1, total demand is 850,000 in time period 1 and increases 850,000 units each time period until $\lceil |T|/2 \rceil$, then decreases 850,000 units each time period until $|T|$, representing a product life cycle. The total demand in scenario 2 reverses the first and second halves of demand in scenario 1, decreasing then increasing to represent an economic downturn followed by recovery. Total demand in each period is divided equally among CZs. Cost parameters for scenario 1 follow Tables 2 and 3. Scenario 2 takes the same cost parameters, except it uses \$400,000 as expansion cost (G_{lt}^{E}).

3.3.2 Test results

Table 9 gives the sizes of the 24 test cases before and after preprocessing. Columns in Table 9 are organized in three groups: the first group describes the case, giving case number and the level of each factor ($|L|$, $|DC|$, N_{Tot}, $|P|$, $|K|$, $|T|$);the second and third groups give instance size (#

of continuous variables, # of binary variables, # of constraints, # nonzeros in constraints) first before, then after, preprocessing for DSCR-N and DSCR-T, respectively. Table 10 gives the MIP objective, LP optimal objective value (i.e., root-node solution) and % GAP for each case in both scenarios. The %GAP is defined as $\%\text{GAP} = 100(Z^*_{IP} - Z^*_{LP}) / Z^*_{IP}$, in which Z^*_{IP} is the value of the optimal solution to the integer problem and Z^*_{LP} is the value of the optimal solution to its linear relaxation (i.e., the value of the optimal solution at the root node in the B&B search tree). Table 11 (12) gives test results for DSCR-T (DSCR-N) in scenario 1, using CPLEX 12.1 default settings. These tables detail the cuts generated by CPLEX (cover cuts, implied bound cuts, flow cover cuts, MIR cuts, flow path cuts, zero-half cuts, multi-commodity cuts, Gomory fractional cuts, clique cuts, total cuts generated), and gives computational results for CPLEX (run time, number of B&B nodes). Table 13 (14) gives test results for DSCR-T (DSCR-N) in scenario 2, using CPLEX 12.1 default settings.

3.3.3 Analysis of results

This subsection analyzes results, focusing on run time and the number of B&B nodes to assess the solvability of each model. We also analyze run time sensitivity to key parameters.

3.3.3.1 Performance comparison

In each of the cases, DSCR-N has fewer continuous variables and constraints but more binary variables, both before and after pre-processing, than DSCR-T. Preprocessing reduces the numbers of continuous variables, binary variables, constraints and non-zeros by 4.20%, 19.49%, 20.83%, and 4.43% (7.42%, 22.22%, 15.20%, and 6.07%) in the DSCR-N (DSCR-T) model. On average, DSCR-N has 15.73% (12.64%) fewer continuous variables and 22.13% (27.28%) fewer constraints but 36.26% (41.08%) more binary variables and 58.04% (60.76%) more non-zeros

before (after) pre-processing, than DSCR-T.

Interestingly, both models give the same optimal solution as well as the same GAP value in each case. The LP relaxations of both models prescribe the same solutions in continuous space. But our tests show that for the MIP formulations, DSCR-N performs better than DSCR-T with regard to run time.

In 19 out of the 24 scenario-1 cases, DSCR-N requires less run time than DSCR-T. On average, DSCR-N requires 48.76% less run time but 1.15% more branch-and-bound (B&B) nodes. Because DSCR-N closes gaps faster than DSCR-T, it explores, then cuts off, more nodes during the B&B search.

When solving cases in scenario 2, neither DSCR-T nor DSCR-N solve case 23 to optimality, or solve case 24 without exceeding memory. Thus, we only count cases 1 to 22 for test result comparison. In 21 out of the 22 scenario-2 cases, DSCR-N requires less run time than DSCR-T. On average, DSCR-N requires 72.44% less run time and 82.36% fewer branch-and-bound (B&B) nodes. On average, DSCR-N (DSCR-T) requires 101.72% (96.13%) more run time to solve scenario 2 cases than scenario 1 cases. Thus, scenario 2 turns to be more challenging to solve than scenario 1, and DSCR-N performs much better than DSCR-T in these cases.

Run times can be rather long, even though these cases do not appear to be large. However, our DSCR-T (DSCR-N) model contains multiples of $|DC|^2 \cdot |T| + |K| \cdot |DC| \cdot |T| + |P| \cdot |L| \cdot |T|$ constraints, $|K|^2 \cdot |DC| \cdot |T| + |DC|^2 \cdot |T|^2$ $\left(|K|^2 \cdot |DC| \cdot |T| + |DC|^2 \cdot |T|\right)$ binary variables, and $|DC|^2 \cdot |P| \cdot |T|^2$ continuous variables, so that case size grows rapidly with the cardinality of sets, K, L, P and T. As the size of a case increases, run time for DSCR-T increases significantly, compared to that of DSCR-N. Our test results show that DSCR-N provides superior solvability, requiring less run time and fewer B&B nodes than DSCR-T.

Increasing the number of time periods and locations increases case size (i.e., the number of binary and continuous variables as well as constraints), resulting in significant computational challenges. Based on test results, the number of time periods has the most significant impacts on run time.

Since runs times are rather long, we also investigated the rate of convergence to optimality. Figure 6 shows how values of the objective functions of the two models converge to optimality as a function of the number of simplex iterations for a typical case (case 6). We define *MIPGAP* as

$$\text{MIPGAP}=|\text{best integer value - best node value}|/(|\text{best integer value}|+1e-10) \ .$$

3.3.3.2 Run time sensitivity

We now analyze the impacts of key parameters ($|P|$, $|T|$, $|L|$, N_Tot) on run time. Tables 15 – 18 give results. We observe (Table 15) that increasing the number of products increases run time substantially, because the number of continuous variables and constraints both increase with $|P|$. Note that the optimization searches for cases with $|P|=4$ and $|P|=5$ do not terminate (i.e., Gap values are 6.9% and 10.05%, respectively) with optimal solutions after lengthy run times. Tables 16-18 show that run time is also sensitive to $|T|$, somewhat less so relative to $|L|$, and not sensitive with respect to N_Tot.

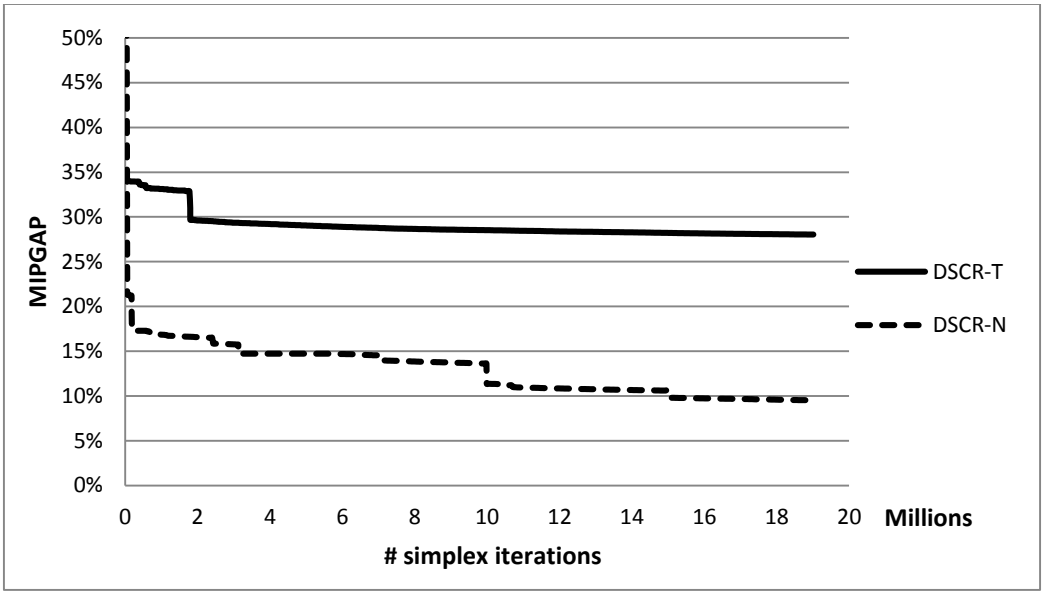

Figure 6. Gap vs. simplex iteration count in DSCR-N and DSCR-T

Table 9. Sizes of test cases for DSCR-N and DSCR-T before and after preprocessing

(a) DSCR-N model

Case	|L|, # Locations	|DC|, #of DCs	N_Tot, Maximum # expansions/ contractions	|P|, # Products	|K|, # Capacity alternatives	|T|, # Time periods	Before Preprocesses				After Preprocessing			
							# Continuous var.	# Binary var.	# Constraints	# Non-zeros	# Continuous var.	# Binary var.	# Constraints	# Non-zeros
1	20	5	3	1	4	8	780	1560	1495	13900	740	1265	1195	13150
2	24	6	3	1	4	8	1080	2016	2088	19176	1032	1638	1656	18180
3	28	7	3	1	4	8	1428	2520	2709	25284	1372	1995	2121	24010
4	32	8	3	1	4	8	1824	3072	3416	32224	1760	2408	2680	30640
5	36	9	3	1	4	8	2268	3672	4203	39996	2196	2853	3303	38070
6	40	10	3	1	4	8	2760	4320	5070	48600	2680	3330	3990	46300
7	20	5	4	1	4	8	780	1560	1495	13820	740	1285	1220	13915
8	24	6	4	1	4	8	1080	2016	2088	20232	1032	1638	1656	19098
9	28	7	4	1	4	8	1428	2520	2709	26516	1372	2023	2156	25081
10	32	8	4	1	4	8	1824	3072	3416	33632	1760	2440	2720	31864
11	36	9	4	1	4	8	2268	3672	4203	41580	2169	2889	3348	39447
12	40	10	4	1	4	8	2760	4320	5070	50360	2680	3370	4040	47830
13	20	5	4	1	4	10	970	1970	1910	18660	903	1652	1520	17700
14	24	6	4	1	4	10	1344	2544	2592	25560	1296	2070	2064	24288
15	28	7	4	1	4	10	1778	3178	3374	33516	1722	2555	2688	31892
16	32	8	4	1	4	10	2272	3872	4256	42528	2208	3080	3392	40512
17	36	9	4	1	4	10	2826	4626	5238	52596	2754	3645	4176	50148
18	40	10	4	1	4	10	3440	5440	6320	63720	3360	4250	5040	60800
19	20	5	4	1	5	10	970	2475	1955	24015	930	2105	1540	23095
20	24	6	4	1	5	10	1344	3150	2646	33078	1296	2646	2088	31854
21	28	7	4	1	5	10	1778	3885	3437	43561	1722	3227	2716	41993
22	32	8	4	1	5	10	2352	4680	4408	55464	2208	3848	3424	53512
23	36	9	4	1	5	10	2916	5535	5409	68787	2754	4509	4212	66411
24	40	10	4	1	5	10	3540	6450	6510	83530	3360	5210	5080	80690

Table 9 Continued
(b) DSCR-T model

Case	|L|, # Locations	|DC|, #of DCs	N_Tot, Maximum # expansions/ contractions	|P|, # Products	|K|, # Capacity alternatives	|T|, # Time periods	Before Preprocesses				After Preprocessing			
							# Continuous	# Binary var.	# Constraints	# Non-zeros	# Continuous	# Binary var.	# Constraints	# Non-zeros
1	20	5	3	1	4	8	980	1120	2090	10335	895	880	1785	9625
2	24	6	3	1	4	8	1320	1488	2748	13554	1200	1170	2334	12654
3	28	7	3	1	4	8	1708	1904	3486	17157	1568	1477	2947	16051
4	32	8	3	1	4	8	2144	2368	4304	21144	1984	1816	3624	19816
5	36	9	3	1	4	8	2628	2880	5202	25515	2448	2187	4365	23949
6	40	10	3	1	4	8	3160	3440	6180	30270	2960	2590	5170	28520
7	20	5	4	1	4	8	980	1120	2090	10335	895	880	1785	9625
8	24	6	4	1	4	8	1320	1488	2748	13554	1200	1170	2334	12654
9	28	7	4	1	4	8	1708	1904	3486	17157	1568	1477	2947	16051
10	32	8	4	1	4	8	2144	2368	4304	21144	1984	1816	3624	19816
11	36	9	4	1	4	8	2628	2880	5202	25515	2448	2187	4365	23949
12	40	10	4	1	4	8	3160	3440	6180	30270	2960	2590	5170	28520
13	20	5	4	1	4	10	1220	1400	2610	13410	1110	1125	2245	12560
14	24	6	4	1	4	10	1644	1860	3432	17532	1512	1470	2934	16452
15	28	7	4	1	4	10	2128	2380	4354	22134	1974	1855	3703	20804
16	32	8	4	1	4	10	2672	2960	5376	27216	2496	2280	4552	25616
17	36	9	4	1	4	10	3276	3600	6498	32778	3078	2745	5481	30888
18	40	10	4	1	4	10	3940	4300	7720	38820	3720	3250	6490	36620
19	20	5	4	1	5	10	1220	1550	2705	15275	1110	1270	2340	14370
20	24	6	4	1	5	10	1644	2040	3546	19890	1512	1644	3048	18744
21	28	7	4	1	5	10	2128	2590	4487	25025	1974	2058	3836	23618
22	32	8	4	1	5	10	2672	3200	5528	30680	2496	2512	4704	28992
23	36	9	4	1	5	10	3276	3870	6669	36855	3078	3006	5652	34866
24	40	10	4	1	5	10	3940	4600	7910	43550	3720	3540	6680	41240

Table 10. MIP solution and gap for both DSCR models in CPLEX12.1

Case	|L|, # Locations	|DC|, # DCs	N_Tot, Max. # expansions/ contractions	|P|, # Products	|K|, # Capacity alternatives	|T|, # Time periods	Scenario 1			Scenario 2		
							MIP Optimal Obj in CPLEX12.1	LP Optimal value	%GAP = 100(IP-LP)IP	MIP Optimal Obj in CPLEX12.1	LP Optimal value	%GAP = 100(IP-LP)IP
1	20	5	3	1	4	8	17,790,633	14,750,975	17.09	19,232,381	15,710,975	18.31
2	24	6	3	1	4	8	16,621,219	14,162,182	14.79	17,790,939	15,122,180	15.00
3	28	7	3	1	4	8	16,802,121	14,098,444	16.09	17,954,690	15,058,444	16.13
4	32	8	3	1	4	8	16,605,912	13,653,609	17.78	17,808,868	14,613,609	17.94
5	36	9	3	1	4	8	16,468,578	13,517,322	17.92	17,696,158	14,477,319	18.19
6	40	10	3	1	4	8	16,480,083	13,478,374	18.21	17,644,359	14,438,374	18.17
7	20	5	4	1	4	8	17,490,633	14,750,975	15.66	19,074,128	15,710,975	17.63
8	24	6	4	1	4	8	16,621,219	14,162,182	14.79	17,790,939	15,122,180	15.00
9	28	7	4	1	4	8	16,802,121	14,098,444	16.09	17,954,690	15,058,444	16.13
10	32	8	4	1	4	8	16,605,912	13,653,609	17.78	17,808,868	14,613,609	17.94
11	36	9	4	1	4	8	16,456,409	13,517,322	17.86	17,657,933	14,477,319	18.01
12	40	10	4	1	4	8	16,480,083	13,478,374	18.21	17,644,359	14,438,374	18.17
13	20	5	4	1	4	10	23,985,341	20,470,213	14.66	26,460,099	21,776,463	17.70
14	24	6	4	1	4	10	22,939,354	19,587,031	14.61	26,970,719	20,893,280	22.53
15	28	7	4	1	4	10	23,048,694	19,491,410	15.43	24,601,262	20,797,657	15.46
16	32	8	4	1	4	10	22,833,502	18,824,163	17.56	24,376,317	20,130,413	17.42
17	36	9	4	1	4	10	22,343,157	18,619,754	16.66	24,142,132	19,926,004	17.46
18	40	10	4	1	4	10	22,599,369	18,542,561	17.95	24,163,645	19,867,561	17.78
19	20	5	4	1	5	10	23,985,341	20,451,463	14.73	29,957,812	21,751,463	27.39
20	24	6	4	1	5	10	22,939,354	19,568,281	14.70	24,509,074	20,868,280	14.85
21	28	7	4	1	5	10	23,048,694	19,472,660	15.52	24,601,262	20,772,658	15.56
22	32	8	4	1	5	10	22,833,502	18,805,413	17.64	24,376,317	20,105,413	17.52
23	36	9	4	1	5	10	22,343,157	18,601,005	16.75	24,142,100^	19,901,004	17.57
24	40	10	4	1	5	10	22,599,369	18,542,561	17.95	24,163,645	19,842,561	17.88

^ best integer

Table 11. Test results of DSCR-N with CPLEX12.1 MIP cuts in scenario 1

Case	|L|, # Locations	|DC|, # DCs	N_Tot, Max. # expansions / contractions	|P|, # Products	|K|, # Capacity alternatives	|T|, # Time periods	Cover cuts	Implied bound cuts	Flow cover cuts	MIR cuts	Flow path cuts	Zero-half cuts	Multicommodity flow cuts	Gomory Fractional cuts	Clique cuts	Total cuts generated	Run time : Total (Sec.)	Nodes of B&B in CPLEX121
1	20	5	3	1	4	8	348	7	355	4	87	3	10	10	0	824	547	302,988
2	24	6	3	1	4	8	231	26	532	2	126	0	3	23	0	943	284	34,057
3	28	7	3	1	4	8	406	30	912	10	216	0	0	15	0	1,589	2,569	183,001
4	32	8	3	1	4	8	330	28	1,059	13	326	0	4	12	0	1,772	2,208	107,626
5	36	9	3	1	4	8	420	24	1,259	7	380	1	0	6	0	2,097	4,467	135,869
6	40	10	3	1	4	8	480	29	1,671	15	508	0	84	4	0	2,791	13,583	241,262
7	20	5	4	1	4	8	377	0	419	21	101	1	26	8	0	953	103	36,615
8	24	6	4	1	4	8	264	21	509	5	128	0	7	16	0	950	137	31,109
9	28	7	4	1	4	8	237	27	757	8	210	0	88	9	0	1,336	1,335	220,487
10	32	8	4	1	4	8	193	34	933	15	313	0	113	22	0	1,623	1,180	48,771
11	36	9	4	1	4	8	240	36	909	4	277	0	110	10	0	1,586	1,841	142,560
12	40	10	4	1	4	8	805	35	1,310	13	416	0	139	2	0	2,720	6,391	394,507
13	20	5	4	1	4	10	769	10	558	19	116	1	39	15	0	1,527	14,626	686,233
14	24	6	4	1	4	10	505	24	817	4	200	0	9	14	0	1,573	8,968	308,063
15	28	7	4	1	4	10	357	28	1,160	25	283	0	76	8	0	1,937	42,045	710,529
16	32	8	4	1	4	10	1,200	87	1,467	22	424	0	143	10	0	3,353	338,721	2,979,490
17	36	9	4	1	4	10	1,371	8	1,570	13	451	0	123	6	0	3,542	237,022	1,398,719
18	40	10	4	1	4	10	1,496	91	1,549	35	414	0	137	20	0	3,742	687,059	4,133,100
19	20	5	4	1	5	10	900	0	595	20	111	0	0	13	0	1,639	2,716	731,757
20	24	6	4	1	5	10	696	15	724	16	161	0	4	23	0	1,639	9,034	2,122,092
21	28	7	4	1	5	10	1,166	8	1,231	26	290	0	6	16	0	2,743	105,079	8,580,809
22	32	8	4	1	5	10	2,585	15	1,559	28	446	0	68	18	0	4,719	182,767	6,200,084
23	36	9	4	1	5	10	1,357	0	1,587	8	428	1	3	7	0	3,391	84,193	2,583,929
24	40	10	4	1	5	10	0	0	840	9	218	0	48	1	0	1,116	523,563	62,667,919

Table 12. Test results of DSCR-T with CPLEX12.1 MIP cuts in scenario 1

Case	|L|, # Locations	|DC|, # DCs	N_Tot, Max. # expansions/ contractions	|P|, # Products	|K|, # Capacity alternatives	|T|, # Time periods	Cover cuts	Implied bound cuts	Flow cover cuts	MIR cuts	Flow path cuts	Zero-half cuts	Multicommodity flow cuts	Gomory Fractional cuts	Clique cuts	Total cuts generated	Run time : Total (Sec.)	Nodes of B&B in CPLEX121
1	20	5	3	1	4	8	234	32	330	16	89	14	0	40	1	756	473	217,710
2	24	6	3	1	4	8	202	48	399	8	118	6	0	25	0	806	408	255,973
3	28	7	3	1	4	8	151	45	672	3	174	7	0	33	0	1,085	1,082	279,279
4	32	8	3	1	4	8	174	45	852	14	274	12	0	29	0	1,400	2,699	417,493
5	36	9	3	1	4	8	154	54	900	8	284	6	0	36	1	1,443	5,428	409,728
6	40	10	3	1	4	8	547	53	1,199	8	338	21	0	23	0	2,189	72,478	1,386,364
7	20	5	4	1	4	8	294	23	336	11	100	14	0	37	0	815	151	80,617
8	24	6	4	1	4	8	110	25	462	8	139	17	0	21	1	783	205	79,558
9	28	7	4	1	4	8	217	36	777	7	202	8	0	48	0	1,295	1,165	231,425
10	32	8	4	1	4	8	206	35	777	6	245	7	0	30	0	1,306	2,581	419,733
11	36	9	4	1	4	8	226	56	929	9	274	3	0	22	0	1,519	6,207	739,231
12	40	10	4	1	4	8	518	31	1,270	11	397	12	0	19	0	2,258	16,608	974,814
13	20	5	4	1	4	10	460	9	397	3	66	18	0	57	1	1,011	24,030	3,361,702
14	24	6	4	1	4	10	410	0	488	13	117	8	0	34	0	1,070	35,694	4,681,569
15	28	7	4	1	4	10	788	0	883	16	184	5	0	39	0	1,915	170,485	4,634,768
16	32	8	4	1	4	10	801	38	1,077	15	306	22	5	32	0	2,296	324,818	6,103,080
17	36	9	4	1	4	10	1,111	0	1,343	12	377	20	0	48	0	2,911	329,646	6,152,054
18	40	10	4	1	4	10	2,189	0	1,560	21	409	30	4	37	0	4,250	1,422,269	23,467,836
19	20	5	4	1	5	10	464	29	435	11	96	4	0	52	0	1,091	5,239	1,245,999
20	24	6	4	1	5	10	436	34	623	6	155	4	0	37	1	1,296	9,421	1,807,843
21	28	7	4	1	5	10	868	0	974	21	207	2	0	38	0	2,110	89,706	7,464,197
22	32	8	4	1	5	10	1,253	0	1,162	24	321	25	0	34	2	2,821	192,952	8,097,857
23	36	9	4	1	5	10	1,015	42	1,226	11	342	7	1	34	0	2,678	217,399	7,816,357
24	40	10	4	1	5	10	2,141	48	1,423	10	391	10	6	22	0	4,051	1,499,433	15,756,424

Table 13. Test results of DSCR-N with CPLEX12.1 MIP cuts in scenario 2

Case	|L|, # Locations	|DC|, # DCs	N_Tot, Max. # expansions/ contractions	|P|, # Products	|K|, # Capacity alternatives	|T|, # Time periods	Cover cuts	Implied bound cuts	Flow cover cuts	MIR cuts	Flow path cuts	Zero-half cuts	Multicommodity flow cuts	Gomory Fractional cuts	Clique cuts	Total cuts generated	Run time : Total (Sec.)	Nodes of B&B in CPLEX121
1	20	5	3	1	4	8	506	22	357	8	82	0	14	10	0	999	860	159,579
2	24	6	3	1	4	8	106	34	472	6	145	0	5	21	0	789	197	18,776
3	28	7	3	1	4	8	414	48	885	10	264	0	58	6	0	1,685	3,334	164,535
4	32	8	3	1	4	8	377	64	967	17	341	0	117	5	0	1,888	8,691	238,914
5	36	9	3	1	4	8	1,006	74	1,348	15	468	0	152	11	0	3,074	59,015	1,535,982
6	40	10	3	1	4	8	895	100	1,668	3	480	0	124	11	0	3,281	50,504	597,896
7	20	5	4	1	4	8	308	23	364	7	99	0	10	11	0	822	255	107,382
8	24	6	4	1	4	8	118	34	554	10	149	1	3	20	0	889	391	28,447
9	28	7	4	1	4	8	492	48	836	13	255	0	51	6	0	1,701	7,177	305,054
10	32	8	4	1	4	8	651	63	973	15	339	0	116	8	0	2,165	24,411	498,274
11	36	9	4	1	4	8	498	74	1,269	6	420	0	160	8	0	2,435	17,372	270,381
12	40	10	4	1	4	8	844	98	1,462	11	429	1	142	2	0	2,989	57,628	565,917
13	20	5	4	1	4	10	2,447	28	645	18	132	1	29	36	0	3,336	124,646	7,004,864
14	24	6	4	1	4	10	386	39	628	44	116	15	20	28	0	1,276	7,197	396,839
15	28	7	4	1	4	10	740	49	1,072	15	287	2	99	12	0	2,276	28,274	888,901
16	32	8	4	1	4	10	1,616	61	1,249	14	417	0	182	16	0	3,555	255,345	3,270,248
17	36	9	4	1	4	10	1,745	71	1,427	5	405	0	166	9	0	3,828	209,714	4,169,010
18	40	10	4	1	4	10	1,684	93	1,608	1	404	0	204	6	0	4,000	442,150	4,291,442
19	20	5	4	1	5	10	426	20	634	13	107	2	64	20	0	1,286	1,465	483,516
20	24	6	4	1	5	10	597	44	766	16	207	0	16	16	0	1,662	10,487	857,907
21	28	7	4	1	5	10	1,116	47	1,008	5	296	2	95	22	0	2,591	27,630	1,429,512
22	32	8	4	1	5	10	1,132	62	1,354	15	489	2	131	6	0	3,191	130,024	1,209,698
23	36	9	4	1	5	10	-	-	-	-	-	-	-	-	-	-	1,564,455	15,701,200
24	40	10	4	1	5	10	1,883	99	1,654	13	426	0	191	18	0	4,284	607,108	4,350,337*

* out of memory

Table 14. Test results of DSCR-T with CPLEX12.1 MIP cuts in scenario 2

Case	|L|, # Locations	|DC|, # DCs	N_Tot, Max. # expansions/ contractions	|P|, # Products	|K|, # Capacity alternatives	|T|, # Time periods	Cover cuts	Implied bound cuts	Flow cover cuts	MIR cuts	Flow path cuts	Zero-half cuts	Multicommodity flow cuts	Gomory Fractional cuts	Clique cuts	Total cuts generated	Run time : Total (Sec.)	Nodes of B&B in CPLEX121
1	20	5	3	1	4	8	325	23	404	4	106	11	0	31	0	904	2,416	472,069
2	24	6	3	1	4	8	299	35	426	2	133	7	0	26	0	928	1,721	258,450
3	28	7	3	1	4	8	358	49	736	3	205	23	2	8	0	1,384	6,533	438,707
4	32	8	3	1	4	8	941	61	770	1	245	12	0	9	0	2,039	67,341	3,988,392
5	36	9	3	1	4	8	653	81	1,041	3	318	17	0	8	0	2,121	105,497	3,625,876
6	40	10	3	1	4	8	1,128	101	1,307	2	376	22	3	5	0	2,944	202,112	5,333,746
7	20	5	4	1	4	8	391	23	373	3	90	13	0	41	0	934	3,098	503,312
8	24	6	4	1	4	8	256	37	478	0	146	14	0	20	0	951	1,925	367,446
9	28	7	4	1	4	8	514	50	761	5	233	21	0	8	0	1,592	25,556	1,223,561
10	32	8	4	1	4	8	421	60	821	1	267	9	0	11	0	1,590	62,122	2,352,476
11	36	9	4	1	4	8	906	82	1,075	2	353	14	0	6	0	2,438	102,368	2,324,195
12	40	10	4	1	4	8	881	101	1,243	4	370	10	0	8	0	2,617	75,433	3,208,305
13	20	5	4	1	4	10	302	21	523	5	95	49	0	30	4	1,029	3,057	418,882
14	24	6	4	1	4	10	681	43	684	7	191	35	3	31	4	1,679	29,938	2,894,254
15	28	7	4	1	4	10	1,705	49	955	1	229	14	0	11	0	2,964	391,777	19,347,585
16	32	8	4	1	4	10	1,750	62	996	2	289	19	2	14	0	3,134	795,549	26,062,302
17	36	9	4	1	4	10	2,416	81	1,350	9	403	11	0	15	0	4,285	1,459,181	55,631,196
18	40	10	4	1	4	10	3,469	99	1,592	7	410	18	49	8	0	5,652	1,457,271	18,736,711
19	20	5	4	1	5	10	262	19	489	3	84	12	1	58	0	928	6,141	1,396,720
20	24	6	4	1	5	10	560	36	628	3	178	8	0	51	0	1,464	34,664	1,984,459
21	28	7	4	1	5	10	620	51	863	2	202	4	2	23	0	1,767	105,122	3,884,971
22	32	8	4	1	5	10	1,215	67	1,109	9	326	14	2	8	0	2,750	383,583	7,101,564
23	36	9	4	1	5	10	-	-	-	-	-	-	-	-	-	-	1,863,708	26,969,400
24	40	10	4	1	5	10	2,284	101	1,596	9	424	4	39	23	0	4,480	1,751,226	19,264,453*

* out of memory

Table 15. Sensitivity with respect to number of products, |*P* | (|*T*|=6, |*L*|=24, |*DC*|=6, *N_Tot* = 4)

|P|	Run time (secs)	MIP GAP (%)	Before Preprocesses				After Preprocessing			
			Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros	Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros
1	28.69	0	816	1488	1584	14712	768	1206	1248	13908
2	8579.85	0	1632	1448	1770	16440	1536	1182	1362	15420
3	2355.55	0	2448	1488	1962	18936	2304	1182	1542	17712
4	1624837.01	6.90	3264	1488	2154	21432	3072	1182	1758	20220
5	1624619.78	10.05	4080	1488	2346	23928	3840	1182	1902	22512
6	731012.65	0	4896	1488	2538	26424	4608	1182	2082	25020

Table 16. Sensitivity with respect to number of periods, |*T* | (|*L*|=36, |*DC*|=9, |*P*|=1, *N_Tot* = 4)

|T|	Run time (secs)	Before Preprocesses				After Preprocessing			
		Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros	Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros
6	9.05	1710	2718	3186	30276	1638	2133	2520	28746
8	1841.11	2268	3672	4203	41580	2169	2889	3348	39447
10	84193.23	2916	5535	5409	63720	2754	4509	4212	60800

Table 17. Sensitivity with respect to number of locations, |*L*| (|*T*|=8, |*P*|=1, *N_Tot* = 4)

|L|	|DC|	Run time (secs)	Before Preprocesses				After Preprocessing			
			Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros	Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros
20	5	103.44	780	1560	1495	13820	740	1285	1220	13915
24	6	136.80	1080	2016	2082	20232	1032	1638	1656	19098

Table 17 Continued

|L|	|DC|	Run time (secs)	Before Preprocesses				After Preprocessing			
			Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros	Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros
28	7	1335.45	1428	2520	2709	26516	1372	2023	2156	25081
32	8	1180.07	1824	3072	3416	33632	1760	2440	2720	31864
36	9	1841.11	2268	3672	4203	41580	2169	2889	3348	39447
40	10	6391.10	2760	4320	5070	50360	2680	3370	4040	47830

Table 18. Sensitivity with respect to N_Tot (|*T*|=8, |*L*|=24, |*DC*|=6, |*P*|=1)

N_Tot	Run time (secs)	Before Preprocesses				After Preprocessing			
		Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros	Number of continuous var.	Number of binary var.	Number of constraints	Number of nonzeros
2	323.69	1080	2016	2088	20232	1032	1638	1656	19098
3	283.80	1080	2016	2088	20232	1032	1638	1656	19098
4	136.80	1080	2016	2088	20232	1032	1638	1656	19098
5	284.81	1080	2016	2088	20232	1032	1638	1656	19098

CHAPTER IV

SSHFCP

This chapter presents our SSHFCP research. This chapter is organized in three sections. Section 4.1 describes our capacity configuration network relative to one service at a given potential location and its extensions to prescribe optimal location for the new facility as well as its multi-service, multi-period capacity configuration. Section 4.2 models attraction of demand from population centers to potential locations in a competitive environment with uncertain demand, defines closed form expressions for recourse (i.e., expected excess demand and capacity) so that associated cost penalties can be incorporated in the objective, formulates SSHFCP as a deterministic RCSPP, and presents our algorithm for solving SSHFCP. Section 4.3 describes our case study and sensitivity analyses, and 6 gives conclusions and ideas for future research.

4.1 Capacity configuration network

The index sets, parameters and decision variables are independently defined in Part I and Part II of this dissertation. If a same notation is defined in both parts, it holds its respective definition within each part.

We define index sets for Part II:

L	potential facility locations	$l \in L$;
P	population centers	$p \in P$;
S	healthcare services of concern	$s \in S$;
T	time periods	$t \in T$;

The index set of potential facility locations, L, must be predetermined and decision makers must identify an index set of relevant healthcare services, S, of concern. The region under consideration must be partitioned into an index set population centers, P; for example, a center may be defined as a census tract (zip code) in configuring a facility within a county (region). Finally, the set of capacity alternatives in which a facility can be configured in each time period must be specified. Because healthcare facilities have long life spans and because the services that a patient requires change over time, we deal with the index set of ordered time periods T to represent a horizon of some 10-20 years for strategic planning. Operational-level decisions that schedule actual patient arrivals to promote smooth patient flow on a daily basis, given the capacity prescribed by our model, are beyond the scope of this dissertation.

This section presents the network aspects of our methodology. It begins by describing the time-staged network on which a RCSPP can be solved to prescribe the capacity configuration for one service at a specified location (i.e., ls). It then extends the ls configuration network to deal with multiple services at a given location (i.e., l, $s \in S$). The last section extends this model to prescribe location as well as capacity configuration (i.e., $l \in L$, $s \in S$). For the purpose of presenting features of our capacity configuration network, we describe a deterministic model in this section, and discuss the stochastic model in Section 4.

4.1.1 *ls* capacity configuration network for one service at a given location

Our capacity configuration model deals with the fundamental features of opening, expansions, contractions and closure over a multi-period planning horizon.

Given service s and potential location l, one of several capacity alternatives can be prescribed in each time period. A cost is incurred if capacity is changed by opening, expanding, contracting or closing at the end of a time period. A traditional MIP model that prescribes capacity

configuration over a planning horizon must 1) specify one capacity alternative in each time period; 2) define the amount of capacity that is opened, expanded, contracted or closed at the end of each time period; and 3) assure that capacity is prescribed to adequately serve demand. Several types of binary variables would be required to prescribe decisions that deal with opening, expanding, contracting and closing, and would require a host of additional constraints to relate them logically.

In contrast to such a traditional MIP model, we propose a more compact, time-staged network formulation that allows a facility to be configured (i.e., opened, expanded, contracted, and closed) over the planning horizon. We propose to use a RCSPP to structure these decisions for a given potential facility l and service s (i.e., ls combination). Figure 7 depicts an ls capacity configuration network with n nodes and m arcs in which the source node 0 (far left) originates one unit of flow and the sink node n (far right) demands one unit. Other nodes are transshipment nodes for which each column represents a time period; and each row, an alternative facility capacity.

Each arc, denoted a, is incident *from* a node representing a known capacity in period $t-1$ *to* a node of known capacity in period t; each horizontal arc indicates that the configuration is not changed from one time period to the next; one that points upwards (downwards) indicates a capacity expansion (contraction) of known amount. Facility closure can be represented by an arc from each node to the sink node or to a level of nodes that represent a closed facility in each time period and would allow re-opening at a later time if such an option is realistic. We do not show this multiplicity of arcs because they would unnecessary complicate Figure 7, which now depicts the fundamental structure of interest.

Note that each arc a represents a known, particular lst combination. To facilitate presentation, we suppress subscripts ls in this subsection. We define the following sets and parameters associated with the ls capacity configuration network of Figure 7.

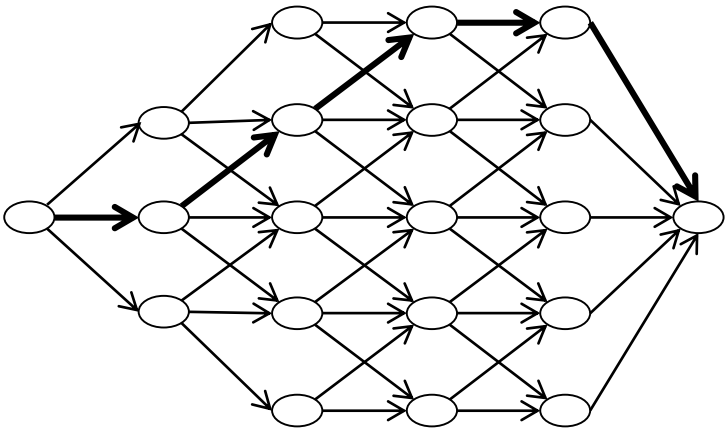

Figure 7. *ls* capacity configuration network

$\bar{A}^t$ index set of arcs incident *from* time period $t-1$ *to* time period t for $t = 1,...,|T|+1$;

N^t index set of nodes (capacity alternatives) in time period t for $t = 0,...,|T|+1$;

$A_i^+ \left(A_i^-\right)$ index set of directed arcs incident from (to) node i, $i \in N^t$ in time period t for $t = 0,...,|T|+1$;

k_a capacity (in terms of number of patient visits/year) at the node *to which* $a \in \bar{A}^t$ is incident;

R index set of limited resources (e.g., budget or number of expansions);

κ_r availability of resource $r \in R$;

φ_{ar} amount of resource $r \in R$ required by arc $a \in \bar{A}^t$;

$b_i = +1$ for $i = 0$ (source node); $b_i = -1$ for $i = n$ (sink node); and $b_i = 0$ for $i \in \bigcup_{t \in T} N^t$.

Now let c_a be the fixed cost associated with arc a, including the land purchase cost (if an opening arc), the closure cost (if a closure arc), the building cost to open, expand or contract (if any), as well as the administration and maintenance cost associated with the node *to which* a is incident (i.e., the capacity in time period t). Upon closure, we allow the building and land to be

sold, incurring revenue as the total value of the land and building. $\tilde{\Omega}_a$, the excess revenue associated with arc a in this deterministic case can be determined as the amount by which revenue exceeds cost (i.e., c_a).

A solution prescribes the path (for example, see bold arcs in Figure 1) taken by the unit of flow through the network, specifying the RCSP and the capacity configuration profile over the planning horizon. Each node on the path represents the capacity prescribed in a time period; and each arc, the excess revenue associated with a particular decision. Using the binary decision variable $x_a = 1$ if a unit of flow is prescribed on arc a (i.e., if arc a is on the RCSP); 0 otherwise, the RCSPP, which inherently minimizes a linear objective function, aims to find the source-to-sink path with maximum excess revenue in the ls configuration network:

$$\tilde{Z}^* = -Min\left\{\sum_{t\in T}\sum_{a\in\bar{A}^t} -\tilde{\Omega}_a x_a : \boldsymbol{x}\in U\right\}, \text{ where}$$

$$U = \left\{\boldsymbol{x}\in\{0,1\}^m : \sum_{a\in A_i^+} x_a - \sum_{a\in A_i^-} x_a = b_i, i\in N^t, t = 0,...,|T|+1; \sum_{t=1}^{|T|}\sum_{a\in\bar{A}^t}\varphi_{ar}x_a \le \kappa_r, r\in R\right\}. \qquad (34)$$

Constraints in (34) invoke flow balance at each node and resource limitations, respectively. The r^{th} resource constraint may, for example, represent a budget limitation for the total fixed cost of configuration over the planning horizon or a limitation on the total number of expansions and contractions at the location. It is clear that $m = \sum_{t=1}^{|T|+1}\left|\bar{A}^t\right|$ and $n = \sum_{t=0}^{|T|+1}\left|N^t\right|$. Because the sink node is a "dummy" used only to invoke network flow balance, $\tilde{\Omega}_a = 0$ for $a\in A^{|T|+1}$; and we omit such terms in the objective function. Although our goal is to maximize excess revenue, we pose (34) as a minimization problem using negative values of excess revenue because an RCSPP is based on finding the *shortest* (i.e., the minimum-length) source-to-sink path in the associated ls configuration network, which is acyclic.

4.1.2 Capacity configuration network for all services at a given location

In this subsection, we extend the ls capacity configuration network to prescribe capacity configuration for a set of services S over the planning horizon at a given potential location l by specifying opening (in time period 1 or later), expansion(s), contraction(s), and possible closure.

For a given location l, ls configuration networks for all services $s \in S$ can be arranged in series as shown in Figure 8, merging the sink node for service $s-1$ and the source node for service s; so a RCSPP can be used to configure the capacity associated with each service $s \in S$.

4.1.3 Facility location and capacity configuration network

This subsection extends the network model further to prescribe an optimal facility location and its capacity configuration for each service $s \in S$. We arrange the set of capacity configuration networks associated with each $l \in L$ (as described in Figure 8) in parallel. Figure 9 shows an example of such an extended network for $|S| = 3$ and $|L| = 3$.

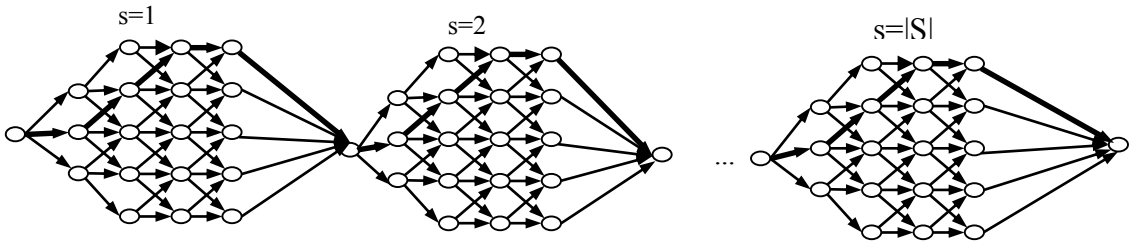

Figure 8. *l, s∈S* capacity configuration network

To prescribe a single location with multiple services, one can solve each of $|L|\bullet|S|$ ls configuration networks, and then determine, by enumeration, the best location

$l^* = \arg max_{l \in L} \left\{ \sum_{s \in S} \tilde{Z}^*_{ls} \right\}$, where $\tilde{Z}^*_{ls}$ is defined in (34). Alternatively, one can solve a single, but larger RCSPP on the extended network (e.g., Figure 9). The solution path from start node to end node prescribes the best location as well as its capacity configuration of each service over the planning horizon (e.g., see the bold path in Figure 9).

The RCSPP that prescribes the optimal location and its capacity configuration is:

$$\tilde{Z}^* = -Min \left\{ \sum_{l \in L} \sum_{s \in S} \sum_{t \in T} \sum_{a \in \bar{A}^t_{ls}} -\tilde{\Omega}_a x_a : \boldsymbol{x} \in \bigcup_{l \in L, s \in S} U_{ls} \right\}, \text{ where } U_{ls} \text{ is defined in (34).} \qquad (35)$$

It is fortuitous that associated revenue and costs can both be related to arc a in an ls configuration network and that the computational advantages of RCSPP can be retained in problems that involve selecting one of multiple potential locations and/or providing multiple services at one location by constructing appropriate configuration networks.

Following recommendations of our healthcare collaborators, the objective is to maximize excess revenue over the planning horizon. This is a "natural" objective for *for-profit* providers, but even *not-for-profit* providers must maintain a positive cash flow to remain financially viable. In addition, all providers will need funds to conform to new requirements mandated by the ACA. Alternatively, a provider could also use our model to minimize total cost by setting revenues to zero.

4.2 Modeling uncertainty in a competitive environment

This chapter evolves our methodology to deal with uncertainty under competition. The first subsection proposes our prototypical MCI demand attraction model, which can be adapted for various healthcare applications. The second subsection models demand uncertainty; the third formulates expected excess revenue as the objective function in our stochastic model; the fourth derives closed-form expressions for expected excess demand and capacity. Finally, the fifth subsection presents the equivalent deterministic formulation of our stochastic model and the

algorithm we propose to solve SSHFCP.

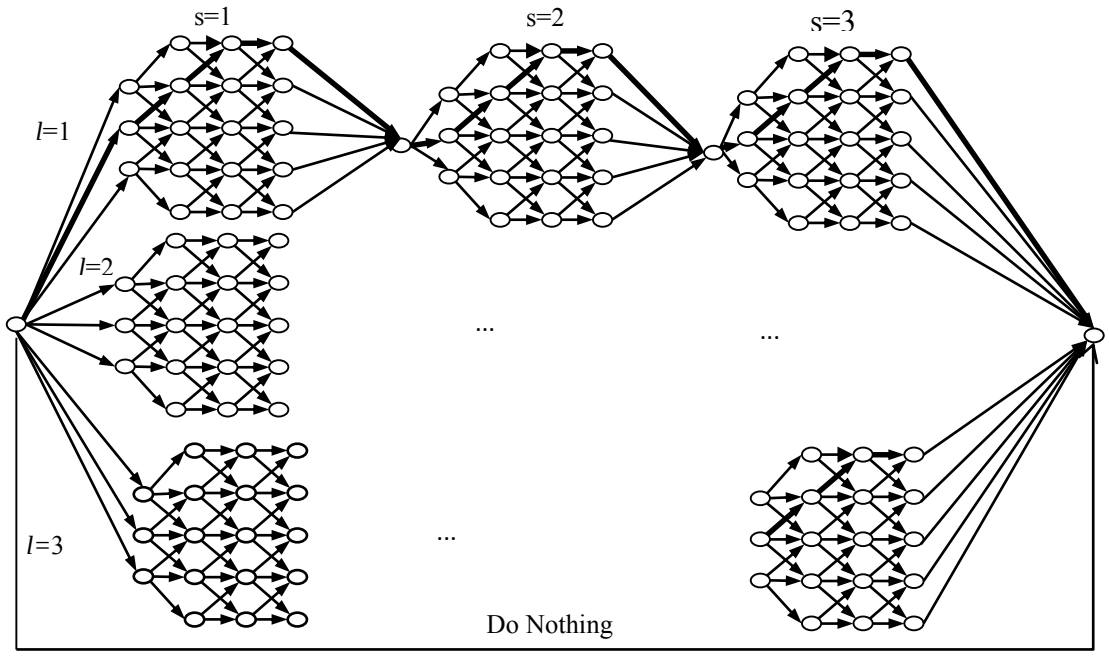

Figure 9. Extended network for |L|=3 and |S|=3

To model competition, we define an index set

C locations of competitors' facilities $l \in C$;

to represent demand in a micro scale, we define index set

G demographic groups (each $\{$age, gender, race/ethnicity$\}$ combination) $g \in G$;

assume that, based on a recent census and a forecast of demographics evolution, we obtain

$\overline{n}_{gpt}$ number of people who are members of demographic group g and reside in population center p in time period t;

and that, based on historical data, one can obtain

$\bar{p}_{gs}$ probability that a member of demographic group g requires healthcare service s.

We assume mutual statistical independence among locations, population centers, services, time periods and demographic groups. Now, we introduce the following random variables:

$\tilde{D}_{gpst}$ Number of members of demographic group g who reside in population center p and seek healthcare service s in time period t;

$\hat{D}_{pst}$ Total demand (in term of number of patient visits/year) originating from population center p for healthcare service s in time period t;

$D_{lst}(k_a)$ Total demand (in term of number of patient visits/year) attracted to facility l from all population centers ($p \in P$) for healthcare service s in time period t, given that capacity alternative k_a, $a \in \bar{A}_{ls}^{t}$, is prescribed.

4.2.1 MCI demand attraction model

Instead of "allocating" all patients in a population center to a particular facility as do location-allocation models (Harper et al., 2005; Syam, 2008; Syam & Côté, 2010), we propose to use an attraction model, reflecting the fact that each person can select the healthcare facility of his/her own choice. We present this as a prototypical model because it demonstrates how relevant healthcare factors can be used to estimate the attractiveness of each location and because it can be readily adapted to other healthcare applications, as we demonstrate in Section 5.

Let $l'' \in L$ be the location of the new facility. Our attraction model adapts the form of an MCI model to define $\xi_{lpst}(k_a)$, the probability that a resident in population center p will be attracted to facility $l \in C \bigcup \{l''\}$ for service s in period t given that capacity k_a is prescribed, using

$$\xi_{lpst}(k_a) = \frac{\Gamma_{lst}^{\alpha} k_a^{\gamma} d_{lp}^{-\beta} B_{lp}^{\delta}}{\sum_{l' \in C \cup \{l''\}} \Gamma_{l'st}^{\alpha} k_a^{\gamma} d_{l'p}^{-\beta} B_{l'p}^{\delta}}, \qquad \text{where competition factor} \qquad (36)$$

Γ_{lst} is a measure of facility attractiveness (e.g., physician rating) relative to healthcare considerations such as patient surveys, survival score, patient safety, and care-related evaluations;

k_a is the capacity at the node *to which* arc $a \in A_{ls}^t$ is incident;

d_{lp} is the shortest-path distance from (the centroid of) population center p to location l;

f_{lp} is the time to travel the fastest route from (the centroid of) population center p to location l;

$\bar{D}_l = \sum_{p \in P} d_{lp} / |P|$ is the average distance of shortest paths from population centers to location l;

$\bar{F}_l = \sum_{p \in P} f_{lp} / |P|$ is the average time of fastest routes from population centers to location l;

$B_{lp} = \left(d_{lp} / \bar{D}_l\right) \div \left(f_{lp} / \bar{F}_l\right)$ is an objective index of accessibility that reflects patient convenience.

Attraction model (36) can be used to calculate $\xi_{lpst}(k_a)$ for each facility $l \in C \bigcup \{l''\}$. Our case study of section 5 is concerned with a single new facility, l''. Should $\hat{m}$ new facilities be configured from $|L|$ potential locations, (36) would have to be formed for each of $\binom{|L|}{\hat{m}}$ possible site combinations, effectively limiting its applicability to locating a single new facility.

We include capacity as a factor in our attraction model to reflect the fact that larger facilities tend to offer more capable physicians and, therefore, their collective reputations (e.g. Mayo Clinic) tend to attract more patients. One could also think of the capacity factor as a surrogate for customer waiting: larger facilities can offer more physicians (i.e., servers) to reduce patient waiting time. In any event, capacity alternatives are known before the model is optimized

so that costs can be calculated to label the expected excess revenue associated with each arc in the ls configuration network. This allows us to model multiple time periods to prescribe globally optimal configurations.

Luo and Wang (2003) proposed measures of accessibility for primary care facilities to represent the relative ease with which a patient can reach a facility for service. We initiate the definition of B_{lp}. Ratio $d_{lp} / \bar{D}_l$ ($f_{lp} / \bar{F}_l$) gives the "standardized" distance (time) from population center p to location l, representing the *relative* ease of access from a particular population center to a specific location. Both distance and accessibility may be important in attracting a patient to a facility. For example, a short distance may require a long travel time due to traffic lights and other bottlenecks (giving poor accessibility). A longer distance on an uncongested freeway that allows faster travel speed (giving good accessibility) may be preferred.

The parameters (i.e., exponents) in (36) represent nonlinearities associated with attractiveness, α; capture rate decay with increasing distance, β; capacity, γ; and accessibility, δ. For each $p \in P$, $s \in S$ and $t \in T$, we note that $0 \le \xi_{lpst} \le 1$ for each $l \in C \bigcup \{l''\}$ and that $\sum_{l \in C \bigcup \{l''\}} \xi_{lpst} = 1$, assuming that $\Gamma^{\alpha}_{lpst} k_a^{\gamma} > 0$ for at least one $l \in C \bigcup \{l''\}$ and that $d_{lp} > 0$ and $\bar{B}_{lp} > 0$ for all $l \in C \bigcup \{l''\}$. We make measures commensurate by scaling each to the range [1, 10].

We assume that factor Γ_{lpst} and parameters α, β, γ, and δ are the same for all demographic groups in each population center, although they could be made specific to each group if data permitted. Attraction model (36) can be used in both deterministic and stochastic cases. In this section, we use it to determine $D_{lst}(k_a)$ as a function of $\hat{D}_{pst}$, $p \in P$:

$$D_{lst}(k_a) = \sum_{p \in P} \xi_{lpst}(k_a) \hat{D}_{pst} \ . \qquad (37)$$

4.2.2 Demand

It is clear that the number of patients $\tilde{D}_{gpst}$ follows a binomial distribution $B\left(\overline{n}_{gpt}, \overline{p}_{gs}\right)$ with mean $\tilde{\lambda}_{gpst} = \overline{n}_{gpt}\overline{p}_{gs}$. The distribution of $\tilde{D}_{gpst}$ can be well-approximated by the Poisson with mean rate $\tilde{\lambda}_{gpst}$ because $\overline{n}_{gpt}$ is sufficiently large and $\overline{p}_{gs}$ is sufficiently small.

Summing up st demand from all demographic groups $g \in G$, we model the number of patients originating in population center p who seek service s in time period t, $\sum_{g \in G} \tilde{D}_{gpst}$, as a Poisson random variable with mean rate $\sum_{g \in G} \tilde{\lambda}_{gpst}$. Using probability $\xi_{lpst}(k_a)$, which is determined by our attraction model (36), we split the pst Poisson patient stream into independent Poisson streams, one for each primary care service $s \in S$, time period $t \in T$ and location $l \in C \bigcup \{l''\}$. Thus, the number of patients for each lst combination, which translates demand *from* population center p *to* location l, is Poisson with mean rate $\sum_{p \in P} \sum_{g \in G} \xi_{lpst}(k_a) \tilde{\lambda}_{gpst}$.

The standard measure of demand and workload used in the industry is the number of patient *visits/year*. Each patient who arrives may require multiple visits/year. We define parameter:

v_{gs} number of visits/year that a patient in demographic group g requires for service s.

Demand in terms of number of visits/year is no longer Poisson, but a random variable with mean $v_{gs}\tilde{\lambda}_{gpst}$ and variance $v_{gs}^2\tilde{\lambda}_{gpst}$, assuming, as we do, that v_{gs} is known deterministically. When

these independent random variables are summed over $g \in G$ to get $\hat{D}_{pst}$, then over $p \in P$ to get $D_{lst}(k_a)$ as described in (37), the distribution of $D_{lst}(k_a)$ can, by the Central Limit Theorem (Casella & Berger, 2001), be approximated by the normal distribution with mean μ_{lst} and variance σ^2_{lst} as defined in (38) and (39), respectively:

$$\mu_{lst}(k_a) = \sum_{p \in P} \sum_{g \in G} \xi_{lpst}(k_a)\, v_{gs} \tilde{\lambda}_{gpst} \tag{38}$$

$$\sigma^2_{lst}(k_a) = \sum_{p \in P} \sum_{g \in G} \xi_{lpst}(k_a)\, v^2_{gs} \tilde{\lambda}_{gpst}\,. \tag{39}$$

4.2.3 Expected excess revenue

In contrast to the *deterministic* excess revenue (i.e., $\tilde{\Omega}_a$) described in Section 3, we define Ω_a as the *expected* excess revenue associated with arc a in the face of uncertain demand. This subsection models the components that constitute Ω_a.

Since a deterministic capacity, k_a, will not equal demand under all possible outcomes, we define the following random variables in terms of the number of patient visits per year as functions of the capacity (i.e., k_a) represented by the node *to which* $a \in \bar{A}^t_{ls}$ incident:

$K^+_{lst}(k_a)$ Excess capacity for service s in time period t at location l;

$D^+_{lst}(k_a)$ Excess demand for service s in time period t at location l;

$\hat{N}_{lst}(k_a)$ Actual number of visits for service s in time period t at location l.

Let Π_{st} be the revenue per visit for service s in time period t. If we define $\Omega_a = \left\{ \Pi_{st} E\left[\hat{N}_{lst}(k_a) \right] - c_a \right\}$, the objective function would not include cost penalties for expected excess demand (i.e., unserved patient visits) or capacity (i.e., resource idleness). Therefore, we

introduce a measure of recourse that penalizes both excessive demand (D^+_{lst}) and capacity, (K^+_{lst}) , using (40) to define $K^+_{lst}(k_a)$ and $D^+_{lst}(k_a)$:

$$D_{lst}(k_a) - D^+_{lst}(k_a) + K^+_{lst}(k_a) = k_a \qquad l \in L, s \in S, t \in T, a \in \bar{A}^t_{ls} . \tag{40}$$

Clearly, $\hat{N}_{lst}(k_a)$ is the minimum of demand and capacity:

$$\hat{N}_{lst}(k_a) = \min\{D_{lst}(k_a), k_a\} , \qquad l \in L, s \in S, t \in T, a \in \bar{A}^t_{ls} . \tag{41}$$

Ω_a reflects recourse by assessing penalties for the expected amounts of excess demand, $E\lfloor D^+_{lst}(k_a)\rfloor$, and capacity, $E\lfloor K^+_{lst}(k_a)\rfloor$, where

$$E\left[D^+_{lst}(k_a)\right] = E\left[\max\left(D_{lst}(k_a) - k_a, 0\right)\right] = E\left\lfloor \left(D_{lst}(k_a) - k_a\right)^+ \right\rfloor , \tag{42}$$

$$E\left[K^+_{lst}(k_a)\right] = E\left[\max\left(k_a - D_{lst}(k_a), 0\right)\right] = E\left\lfloor \left(k_a - D_{lst}(k_a)\right)^+ \right\rfloor . \tag{43}$$

For each *lst* combination, let c^K_{lst} be the time-discounted cost per unit of excess capacity; that is, under-utilized resources, or equivalently, the opportunity cost associated with investments that could provide other services that are in limited supply; c^D_{lst} be the time-discounted cost per unit of excess demand, which can represent patient delay and/or health ramifications if a patient must travel to a distant location or forego healthcare. ACA requires that providers be measured relative to patient satisfaction, so excess demand is a valid measure of the level of patient-centered service offered; large values may result in an undesirable loss of patients seeking repeat services.

Knowing the nodes *from* and *to* which arc a is incident allows it to be labeled with fixed cost c_a appropriate for the associated capacity change, if any, in transitioning *from* the tail node *to* the head node (i.e., opening, expanding, contracting, or closing), as well as revenue and penalties for excess demand and capacity, all of which relate to the capacity k_a (i.e., the maximum number

of patient visits/year that can be served) for service s at location l in period t that is associated with the node *to which* $a \in \bar{A}_{ls}^{t}$ is incident. Now, Ω_a for each $a \in \bar{A}_{ls}^{t}$, $l \in L$, $s \in S$, $t \in T$ is defined as

$$\Omega_a = \Pi_{st} E\left[\hat{N}_{lst}(k_a)\right] - c_{\bar{a}} - c_{lst}^{K} E\left[K_{lst}^{+}(k_a)\right] - c_{lst}^{D} E\left[D_{lst}^{+}(k_a)\right], \tag{44}$$

leading to the stochastic formulation for SSHFCP associated with the ls configuration network:

$$(\text{SHFCP}_{ls}): \qquad Z_{ls}^{*} = -Min\left\{\sum_{t\in T}\sum_{a\in\bar{A}_{ls}^{t}} -\Omega_a x_a : (7),(8),(9),(10) \text{ for } a \in \bar{A}_{ls}^{t},\, t \in T; \text{ and } x_a \in U_{ls}.\right\},$$

where Ω_a is defined in (44) and U_{ls} is defined in (34).

4.2.4 Closed form expressions for recourse

This section establishes a closed-form model of recourse that represents *all* possible demand outcomes instead of, for example, using a two-stage stochastic program that represents outcomes statistically using a (typically small) sample of scenarios.

Recall that $E\left\lfloor D_{lst}^{+}(k_a)\right\rfloor$, $E\left\lfloor K_{lst}^{+}(k_a)\right\rfloor$ and $E\left[\hat{N}_{lst}(k_a)\right]$ are functions of the prescribed capacity k_a for each $a \in \bar{A}_{ls}^{t}$ in the ls configuration network, $l \in L$, $s \in S$ and $t \in T$. For simplicity in deriving the closed form expression for recourse, in this subsection, we suppress the subscript a for each capacity alternative as well as the subscript lst. We use $f(k)$ and $F(k)$ to denote the probability density and cumulative distribution function, respectively, of the normal distribution evaluated at capacity k. Proposition 1 establishes that $E[K^{+}]$ is a linear function of $E[D^{+}]$, regardless of the distribution of D:

Proposition 1: Regardless of the distribution of demand D, $E[K^{+}] = E[D^{+}] - E[D] + k$.

Proof: $K^+ = \max(k-D,0) = k - D + \max(0,-k+D) = k - D + D^+$

$\Rightarrow E[K^+] = E[D^+] - E[D] + k$. □

Proposition 2 shows that $E[\hat{N}]$ can be determined as a function of $E[D^+]$ or of $E[K^+]$, regardless of the distribution of D:

Proposition 2: Regardless of the distribution of demand D, $E[\hat{N}] = E[D] - E[D^+] = k - E[K^+]$.

Proof: $\hat{N} = \min(D,k) \quad = -\max(-D,-k) = D - D - \max(-D,-k) = D - \max(0, D-k) = D - D^+$

$= D - (K^+ - k + D) = k - K^+$. Linear transformation on expectation gives the results. □

We consider approximating demand D using the normal distribution $N(\mu,\sigma^2)$. Proposition 3 gives $E\lfloor D^+ \rfloor$, given that demand is normally distributed.

Proposition 3: If demand D is $N(\mu,\sigma^2)$, $E\lceil D^+ \rceil = \sigma\phi(z) + (\mu - k)\lceil 1 - \Phi(z) \rceil$,

where $z = (k-\mu)/\sigma$, $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$, and $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ denote, respectively, standard normal and its probability density and cumulative distribution functions.

Proof: $E[D^+] = E[(D-k)^+] = \int_k^\infty (w-k) f(w) dw$

$$= \int_k^\infty (w-\mu) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(w-\mu)^2}{2\sigma^2}} dw + \int_k^\infty (\mu - k) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(w-\mu)^2}{2\sigma^2}} dw$$

$$= \sigma^2 f(k) + (\mu - k)\lceil 1 - F(k) \rceil = \sigma\phi(z) + (\mu - k)[1 - \Phi(z)]. \square \qquad (45)$$

Invoking Propositions 1 and 2, Corollary 4 presents $E\lfloor K^+ \rfloor$ and $E\lceil \hat{N} \rceil$ for $D \sim N(\mu,\sigma^2)$:

Corollary 4: If demand D is $N(\mu,\sigma^2)$,

$E\lfloor K^+ \rfloor = E\lfloor D^+ \rfloor - \mu + k$ and

$$E\left[\hat{N}\right] = \mu - E\left[D^+\right] = k - E\left[K^+\right].$$

Capacity may actually be a random variable instead of a fixed number. Let $\hat{k}$ be the expected number of patient visits/year that one physician can serve. If the capacity of one physician is $N(\hat{k}, \rho^2\hat{k}^2)$, the capacity of η physicians is a random variable, $K \sim N(\eta\hat{k}, \eta\rho^2\hat{k}^2)$. The random variable $D-K$ is normally distributed and $D^+ = \max(0, D-K)$. Corollary 5 defines $E[D^+]$ by replacing k with $\eta\hat{k}$ and σ with $\sqrt{\sigma^2 + \eta\rho^2\hat{k}^2}$ in Proposition 3.

Corollary 5: If demand $D \sim N\left(\mu, \sigma^2\right)$ and $K \sim N(\eta\hat{k}, \eta\rho^2\hat{k}^2)$,

$$E\left[D^+\right] = \zeta\phi(z) + \left(\mu - \eta\hat{k}\right)\left[1 - \Phi(z)\right], \text{ where } \zeta = \sqrt{\sigma^2 + \eta\rho^2\hat{k}^2} \text{ and } z = \left(\eta\hat{k} - \mu\right)\Big/\zeta.$$

4.2.5 Solution methodology

The network formulation offers the advantage that revenue, which is a function of $E\left[\hat{N}_{lst}\right]$, and all costs, including c_a, and the penalties associated with $E\left\lfloor D^+_{lst}\right\rfloor$ and $E\left\lfloor K^+_{lst}\right\rfloor$, are related to the capacity k_a prescribed in period t by arc a. Using closed-forms of (41), (42) and (43), constraints (40) can be eliminated and expected excess revenue Ω_a in (44) can be calculated a priori and used to label arc a. This fortuitously avoids nonlinearity in the objective function and allows the stochastic SSHFCP to be solved as a deterministic RCSPP. Proposition 6 defines the closed-form SSHFCP$_{ls}$:

Proposition 6: For a given ls configuration network with known mean (i.e., $\mu_{lst}(k_a)$) and variance (i.e., $\sigma^2_{lst}(k_a)$) of demand, SHFCP$_{ls}$ is

$$Z^*_{ls} = -Min\left\{\sum_{t\in T}\sum_{a\in\bar{A}^t_{ls}} -\Omega_a x_a : x_a \in U_{ls}\right\}, \text{ where}$$

$$\Omega_a = \left(\Pi_{st} + c^K_{lst}\right)\mu_{lst}(k_a) - c^K_{lst}k_a - c_a - (\Pi_{st} + c^K_{lst} + c^D_{lst})E\left[D^+_{lst}(k_a)\right], \qquad (46)$$

$E\left[D^{+}_{lst}(k_a)\right]$ is defined by (45) and U_{ls} is defined by (34).

Proof: The maximum expected excess revenue associated with an *ls* configuration network is

$$Max\sum_{t\in T}\sum_{a\in\bar{A}^{t}_{ls}}\left(\Pi_{st}E\left[\hat{N}_{lst}(k_a)\right]-c_{\bar{a}}-c^{K}_{lst}E\left[K^{+}_{lst}(k_a)\right]-c^{D}_{lst}E\left[D^{+}_{lst}(k_a)\right]\right)x_a\ ,$$

where expected revenue $\Pi_{st}E\left[\hat{N}_{lst}(k_a)\right]$ and penalties for expected excess capacity and demand, $c^{K}_{lst}E\left[K^{+}_{lst}(k_a)\right]$ and $c^{D}_{lst}E\left[D^{+}_{lst}(k_a)\right]$, are represented explicitly as functions of k_a; that is, of decision variable x_a.

From Propositions 1 and 2, $E\left[\hat{N}_{lst}(k_a)\right]=\lambda_{lst}-E\left[D^{+}_{lst}(k_a)\right]$ and $E\left[K^{+}_{lst}(k_a)\right]=E\left[D^{+}_{lst}(k_a)\right]-\mu_{lst}+k_a$; by substituting, we obtain

$$\Omega_a=\left(\Pi_{st}+c^{K}_{lst}\right)\mu_{lst}(k_a)-c^{K}_{lst}k_a-c_{\bar{a}}-(\Pi_{st}+c^{K}_{lst}+c^{D}_{lst})E\left[D^{+}_{lst}(k_a)\right]\ ;$$

and the objective function of SSHFCP results:

$$Z^{*}_{ls}=Max\sum_{t\in T}\sum_{a\in\bar{A}^{t}_{ls}}\Omega_a x_a=-Min\sum_{t\in T}\sum_{a\in\bar{A}^{t}_{ls}}-\Omega_a x_a\ .\ \square$$

If the capacity impact is eliminated from attraction model (36) (i.e., $\gamma=0$), μ_{lst} and σ^{2}_{lst} are constants with respect to decision variable x_a, and the objective function can be further simplified:

Corollary 7: If the capacity factor is eliminated from attraction model (36) (i.e., $\gamma=0$), SHFCP$_{ls}$ reduces to:

$$Z^{*}_{ls}=\sum_{t\in T}\left(\Pi_{st}+c^{K}_{lst}\right)\mu_{lst}-Min\left\{\sum_{t\in T}\sum_{a\in\bar{A}^{t}_{ls}}\hat{\Omega}_a x_a : x_a\in U_{ls}\right\},\ \text{where}$$

$$\hat{\Omega}_a=\left(c^{K}_{lst}k_a+c_a+(\Pi_{st}+c^{K}_{lst}+c^{D}_{lst})E\left[D^{+}_{lst}(k_a)\right]\right),$$

$E\left[D^{+}_{lst}(k_a)\right]$ is defined by (45) and U_{ls} is defined by (34).

Figure 10 shows how $|L|\bullet|S|$ *ls* configuration networks can be used in an enumerative

process to prescribe a single optimal location $l^* = \arg max_{l \in L} \left\{ \sum_{s \in S} Z_{ls}^* \right\}$ and the capacity configuration of services it offers. The algorithm runs in pseudo-polynomial time.

I. For each $l \in L$, $s \in S$:

1. For each time period $t \in T$ and each capacity alternative k_a, $a \in \bar{A}_{ls}^t$:
   (1) Compute $\xi_{lpst}(k_a)$, the probability that a resident in population center p seeks service s at facility l if capacity alternative k_a is prescribed in time period t using (3);
   (2) Compute mean (i.e., $\mu_{lst}(k_a)$) and variance (i.e., $\sigma_{lst}^2(k_a)$) of demand attracted facility l if capacity alternative k_a is prescribed in time period t using (5) and (6)
   (3) Compute the expected excess demand ($E\left[D_{lst}^+(k_a) \right]$) associated with arc $a \in A_{ls}^t$ using (12) in Proposition 3.
   (4) Label the expected excess revenue Ω_a on arc $a \in A_{ls}^t$ using (13) in Proposition 6.
2. Solve the RCSPP (1) associated with the ls configuration network, obtaining the optimal configuration and the expected excess revenue for the ls combination, Z_{ls}^* .

II. Prescribe the best location l^* for the new facility: $l^* = \arg max_{l \in L} \left\{ \sum_{s \in S} Z_{ls}^* \right\}$.

Figure 10. Steps in solving SSHFCP

4.3 Case study

To demonstrate managerial use of our model, we now describe a case study that prescribes the configuration of a new primary care center that can provide up to three services (family practice (FP), internal medicine (IM), and pediatrics (P)) over a ten-year planning horizon. We base this hypothetical, yet realistic, case on publically available data, expecting that a provider conducting

an actual study could supplement this data, for example, with proprietary, historical data and/or by purchasing more detailed information such as forecasts from a demographics service.

Our case study uses the extended network as described in Figure 9 for a compact formulation. We follow step 1 in Figure 10 to label Ω_a on arc $a \in \bigcup_{l\in L, s\in S, t\in T} \bar{A}_{ls}^t$, then solve a single RCSPP for SSHFCP:

$$\text{(SSHFCP): } Z^* = -Min\left\{ \sum_{l\in L}\sum_{s\in S}\sum_{t\in T}\sum_{a\in A_{ls}^t} -\Omega_a x_a : x_a \in U_{ls} \text{ for } a \in \bigcup_{t\in T} \bar{A}_{ls}^t, l \in L, s \in S \right\},$$

where U_{ls} is defined in (34).

We assume that personnel are downsized in a contraction but that the building is not changed, because it would not be possible to lease or sell part of the building. In our case study, our ls configuration network is extended to define each capacity alternative as a (number of physicians, building size) pair, so, if a contracted facility were later expanded, the correct building size would be represented. As a result, the fixed configuration cost (i.e., c_a) embodies the economies of scale.

This section comprises four subsections. The first describes the case setting and the data we use. The second presents a sensitivity analysis that we conduct to assess the relative impacts of competition factors and key parameters on location and configuration decisions. The third describes a sensitivity analysis relative to the variance of physician capacity; and the fourth, penalty cost values. Because run times are not significant, we do not include a detailed analysis of run time as a function of parameters.

4.3.1 Case setting and data

This subsection describes the setting upon which we base our case study as well as the data sources we employ. The following discussion reviews the sites of existing primary care centers and

potential locations for a new one; current and forecast demand by demographic group in each year; analysis of primary care centers to specify capacity alternatives, including staffing levels and costs (e.g., land, building, annual operating, expanding, contracting and closing); revenue per visit; and attraction model (36) parameters.

To reflect practical features with fidelity, our model deals with a variety of relevant data. Table 19 relates data to our publically available sources. Column 1 gives the set or parameter; Column 2, our source (typically a government website); and Column 3, an index we use in this section to refer to each source that is identified by this table. We report data specific to the Brazos County in this section. The interested readers can obtain the rather voluminous data we have obtained from national sources on the authors' website (**http://ise.tamu.edu/nsfhc**).

Our case study focuses on rural Brazos County, TX (population 194,851). The cities of College Station (population 93,857) and the adjoining Bryan (population 76,201) form a long, narrow metropolis that runs southeast to northwest within the county (Figure 11). Although a few medical offices are scattered around, the three principal centers of primary care are in the center-north of Bryan and the center-north and center of College Station, as depicted by three flags (C={E1, E2, E3}) in Figure 11. College Station has grown rapidly over the last decade, mostly southward, creating potentially underserved areas. Our case study seeks to identify the best configuration for a new primary care center, given the three existing centers. Figure 5 also shows eight tear drops (L1-L8) that identify potential locations $l'' \in L$ where a new primary care center could be sited. We identified these parcels of land by consulting with city planning offices (s1) to learn about development plans and areas where growth is expected, and by consulting local real estate data bases (s7) to determine available parcels and their prices (Table 19).

This case study involves a relatively small county, we define each population center as a census tract. By definition, each is a geographic area within a county that has between 2,500 and

8,000 residents and boundaries that follow physical features (e.g., roads, highways, rivers). The U.S. Census Bureau (s2) partitions Brazos County into 40 census tracts. We use ArcGIS to determine the centroid of each tract p and Google Maps to compute the shortest travel distance and time from it to each existing primary care center ($l \in C$) and to each potential location ($l \in L$).

Table 19. Source for each type of data

Sets or Parameters	Source	Index
L	http://ims.bryantx.gov/gis/website/new_development/viewer.htm	s1
2000 and 2010 census	http://www.census.gov/;	s2
Population growth projection	http://txsdc.utsa.edu/data/TPEPP/ Projections/Index.aspx	s3
$\bar{p}_{gs}$ and v_{gs}	http://www.cdc.gov/nchs/data/ahcd/namcs_summary/2010_namcs_web_tables.pdf	s4
Π_{st}	www.amnhealthcare.com/industry-research/ 2010-Physician-Inpat *FP ($1,622,832), IM ($1,678,341), and P ($856,154)*	s5
Γ_{lst}	https://www.healthtap.com/ (Table 3)	s6
Land purchase cost (part of c_a)	https://propaccess.trueautomation.com/clientdb/?cid=65 (Table 2)	s7
Capacity: # visits/year/physician	http://www.medscape.com/sites/ public/ physician-comp/2012 *FP (3836), IM (3564), and P (3754)*	s8
Staff associated with physician	https://www.vendorportal.ecms.va.gov	s9
Clinic & parking-lot size standards	http://iss.unm.edu/PCD/docs/UNMH-Clinic-Standards_Revised04-12-10.pdf	s10
Construction cost (part of c_a)	http://www.license.state.tx.us/ab/2012abtas5.htm;	s11
Physician and staff salaries (part of c_a)	www.salary.com *Physician: FP ($178,000), IM ($189,000) and P ($174,000), Registered nurses ($80,372), Licensed practical nurses ($52,685), and Physician assistants ($107,934)*	s12

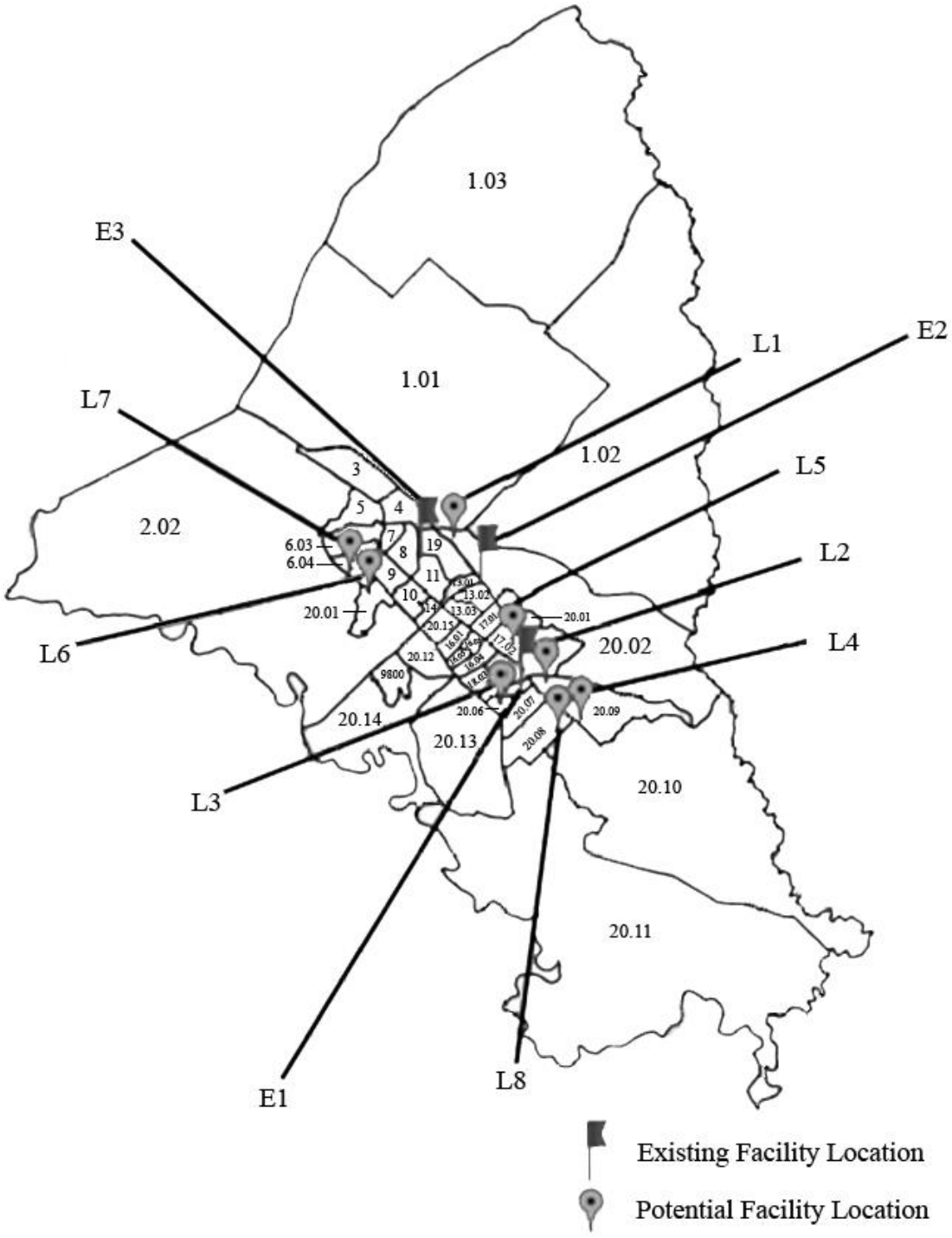

Figure 11. Census Tracts within Brazos County

Table 20. Land cost at potential locations

Potential Location	Cost ($/acre)	Potential Location	Cost ($/acre)
1	326,699.9	5	87,120.2
2	134,747.8	6	39,426.4
3	87,119.8	7	13,721.3
4	209,088.1	8	282,269.6

The U.S. Census Bureau (s2) reports the total population for each of the 40 tracts, as well as the number in each demographic group, each of which comprises one of the possible combinations of (age, gender, race/ethnicity)

- Age <15, 15-24, 25-44, 45-64, 65-74, ≥ 75
- Gender male or female
- Race/Ethnicity African American, Caucasian, Hispanic or Latino, Asian;

giving $|G| = 6 \times 2 \times 4 = 48$ (age, gender, race/ethnicity) groups altogether. In order to forecast the annual demand for service s from within population center p during time period t for a $|T|$-year planning horizon, we combined data from the U.S. Census Bureau (s2), the Texas State Data Center (s3) and the Center for Disease Control and Prevention (s4), which give, respectively, the 2010 population for each census tract within Brazos County by demographic group; the population projections for Brazos County by age, sex, and race/ethnicity; and the national number of annual visits *by age* for primary care services. Some 50,000 students attend Texas A&M University in College Station, but we do not include them as they typically use healthcare facilities on campus.

Tracts in the center of Brazos County have increased populations over the years and are now "saturated," providing little space for new homes, so we assume they will experience growth rates less than 5% per year. We assume that residents of saturated tracts have lived there for some time so that 50% are committed to their PCP and will not consider changing physicians. However, we assume

that 50% of the residents in each of these tracts are amenable to our attraction model because they may be selecting a PCP for the first time or may have purchased an existing home and moved into the tract recently. For tracts with populations that are projected to grow by more than 5% per year, we assume that 100% of the population is amenable to our attraction model because these residents are typically young and may be moving into the area as new houses are built.

The Center for Disease Control and Prevention (s4) provides the number of office visits/year/100 persons for primary care (i.e., FP, IM, and P) for each patient-age category (i.e., aggregating visits for genders and race/ethnicities in each age group). To obtain μ_{lst}, we adapt our model by first adding our population forecasts for genders and race/ethnicity in each age group and then, for each service, multiplying by the annual number of office visits for each age group.

Next, we analyze primary care facilities to determine meaningful capacity alternatives, basing each on the number of physicians it provides. For each ls configuration network, we define the capacity alternatives (nodes) in each time period as comprising 0, 1, 2, 3, 4, 5, 8, 10, 15 or 20 physicians. We measure the capacity of each type of primary care using the number of visits/year with which each PCP can deal. We set Π_{st}, the revenue/visit for service s by dividing the revenue/year/PCP reported by a 2010 survey (s5) by the average number of visits/year each serves (s8). We do not employ a budget limitation but do restrict the number of expansions and contractions to four for each service over the planning horizon. In particular, we define $|R|=1$; $\varphi_{a1}=1$ if a is an expansion/contraction arc, 0 otherwise, and $\kappa_1 = 4$ in (34).

Components of a primary care center can be estimated, given the number of physicians. For example, primary care clinics average 2.17 clinical support staff and 3 rooms for each physician (s9). We estimate the number of square feet needed for the building associated with each capacity alternative using these ratios along with typical clinic and parking-lot size standards (s10) and then determine construction cost using a web site designed for that purpose (s11). Using linear

regression, we define the fixed cost for a building with an area that can accommodate as many as η physicians and their supporting staff and equipment as $\$1,221,400+(78,910)\eta$, reflecting economies of scale. We include a table on our website (**http://ise.tamu.edu/nsfhc**) that describes each alternative, including economies of scale as capacity increases, the numbers of physicians and clinical support staff, the numbers of rooms and square feet, and the capacity (i.e., number of patient visits/year that can be served).

We use average annual compensation of family practitioners, internists, pediatricians, and support staff, including registered nurses, licensed practical nurses, and physician assistants in Bryan/College Station (s12) in determining the annual administration cost for each capacity alternative. Annual operating cost includes physician and staff remuneration plus maintenance at 5% of building cost. Because the land and building costs differ by location and operating costs are different for each service, this large amount of data, can be found at our website (**http://ise.tamu.edu/nsfhc**).

We use physician ratings in Table 21 (s6) to set the reputation factor Γ_{lst} in our attraction model (36). We assume that the new facility tends to hire PCPs with respected reputations so that they share the same PCP ratings as at E2, which are higher than those at E1 and E3.

Table 21. Physician ratings

Facility	FP	IM	P
E1	7.37	7.03	7.77
E2	7.94	7.2	8.05
E3	7.16	6.97	7.32

Table 22 gives the annual growth rate we assume for each of the 40 tracts. We determine the growth rates by measuring the evolution of census tracts from 2000 to 2010 census (s2), and

validate the compound growth that results over the planning horizon with the growth rate projected by the county (s3). We assume that each existing facility increases capacity by 4% each year to correspond with the overall growth rate and to reflect competitor reactions to the new facility. As an example, Table 22 compares the demand for FP attracted from each tract to each location in 2015 before/after a new facility at L3 enters the market with $\alpha = 1$, $\beta = 2$, $\delta = 1$ and $\gamma = 1$. A new facility at L3 attracts a larger market share in growing tracts (20.06-20.13) than in the saturated tracts (7-9, 11, 13), which are more distant and have a lower portion of residents willing to consider a new PCP. The sum of demands in each tract attracted away from existing facilities equals the demand attracted to L3.

4.3.2 Sensitivity analyses

We now conduct sensitivity analyses with the goal of assessing the relative impacts on configuration decisions of key competition factors (e.g., physician rating, distance); parameters $\alpha, \gamma, \beta, \delta$; and percentage of the population amenable to attraction model (36). Initially (base case 1), we set $\alpha = \delta = 1$ and $\beta = 2$, emphasizing travel distance, but subsequently conduct sensitivity analyses relative to these parameters as described in the next subsection. We evaluate the impact of including capacity in attraction model (36) by comparing results with $\gamma = 0$ and $\gamma = 1$ in each case. We also employ test cases 6 and 7, which fix one new facility and prescribe the best location for a second new facility.

Table 23 gives results for our 11 test cases. Columns 1 and 2 give, respectively, the case number and the setting that defines each case. Columns 3-6 (7-10) report results when capacity is not $(i.e., \gamma = 0)$ (is $(i.e., \gamma = 1)$) incorporated in attraction model (36), including the optimal location, and expected excess revenue (x10^8), and the number of PCP years prescribed; and the run time (in

seconds). Each case 2-11 differs from base case 1 in one setting to assess its impact. The impact of including capacity in the attraction model can be assessed by comparing results for a given case

Table 22. Demand attraction before and after the new facility is located at L3

Census Tracts	Growth rate/year	E1		E2		E3		L3
		before	after	before	after	before	after	new
1.01	3.07%	841	826	1564	1536	1835	1802	76
1.02	2.47%	666	651	1095	1070	1306	1277	68
1.03	3.26%	209	203	303	295	348	339	21
2.01	0.76%	320	313	1059	1035	848	829	49
2.02	3.18%	555	548	1874	1853	1056	1045	38
3	1.32%	503	492	1571	1539	1187	1162	68
4	0.39%	225	223	1285	1275	778	772	19
5	0.01%	236	234	1450	1435	632	625	23
6.03	0.57%	231	229	1604	1589	601	596	22
6.04	1.54%	307	303	1790	1767	821	811	38
7	0.07%	47	47	955	952	185	184	4
8	0.96%	47	47	1957	1953	208	208	5
9	-1.55%	65	65	601	596	217	215	8
10	2.71%	187	185	1360	1344	664	656	26
11	-0.72%	79	79	1236	1232	533	532	6
13.01	3.48%	30	29	135	135	814	812	2
13.02	0.00%	57	57	65	65	875	871	4
16.04	4.11%	596	551	207	192	428	395	93
16.05	0.64%	364	352	172	166	389	377	29
16.06	0.64%	474	446	114	107	200	188	47
17.01	3.13%	250	246	136	133	837	824	20
17.02	3.14%	376	363	100	96	240	231	25
18.01	-0.59%	1405	1335	81	77	152	144	82
18.03	0.27%	1510	1278	89	76	151	128	269
18.04	0.00%	335	308	8	8	19	18	30
19	0.57%	91	91	1836	1829	453	451	10
20.01	0.64%	959	930	242	235	681	661	56
20.02	4.39%	1848	1766	340	325	692	661	129
20.06	6.02%	1470	492	75	25	134	45	1117
20.07	5.40%	3044	1545	318	162	692	351	1996
20.08	6.40%	2823	2423	420	361	938	805	592
20.09	6.65%	2655	2514	585	554	1204	1140	236
20.10	15.00%	1189	1137	394	377	785	751	103
20.11	15.00%	1074	1028	570	546	866	829	108
20.13	6.16%	4026	3595	774	691	1291	1153	652
20.14	6.15%	377	363	382	368	527	508	47

with $\gamma = 0$ (columns 3-6) and with $\gamma = 1$ (columns 7-10). To provide a single, summary measure to compare the various cases over the planning horizon, we add the numbers of physicians prescribed in FP, IM and P practices each year over the 10 years and report this summary measure (#PCP years) in Table 23.

We use *boost* (**http://www.boost.org**) to compute standard normal $\Phi(z)$ values, *AMPL 9.0*®, and the *IBM ILOG CPLEX12.1*® branch-and-bound solver with default settings. We perform all computational tests on an Intel Xeon CPU E5620 @ 2.40GHz (2 processors), a 64-bit operating system, with 12.0GB RAM. Our case study RCSPP represents 40 census tracts, 48 demographic groups, 3 existing facilities, and 8 potential new locations; it involves 225,984 binary variables and 11,082 constraints.

Table 23. Settings and test results for 11 cases

Case	Setting	W/O capacity impact $\gamma = 0$				With capacity impact $\gamma = 1$			
		Opt Loc	Opt Obj x10^8	#PCP years	Run time secs	Opt Loc	Opt Obj x10^8	#PCP years	Run time secs
1	$\alpha = 1,\ \beta = 2,\ \delta = 1$	L3	\$1.31	168	23.2	L3	\$0.37	62	22.1
2	$\alpha = 2$, enhance physician ratings	L3	\$1.34	172	23.8	L3	\$0.40	68	21.0
3	Attraction model applies to everyone	L5	\$1.84	217	23.0	L7	\$0.80	136	25.3
4	4000 visits/PCP/year for each service	L3	\$1.35	155	24.7	L3	\$0.31	54	20.6
5	Physician ratings now E1's not E2's	L3	\$1.28	164	24.7	L3	\$0.33	60	21.4
6	0 expansion/contraction for each service	L3	\$1.24	140	6.49	L3	\$0.36	60	6.05
7	Make L8 an existing facility	L3	\$1.18	151	19.9	L3	\$0.32	59	17.1
8	Make L3 an existing facility	L5	\$0.93	109	19.8	L4	\$0.08	20	14.9
9	$\alpha = 0, \beta = 2,\ \delta = 0$, distance only	L3	\$1.29	163	24.8	L3	\$0.30	49	20.8
10	$\alpha = 0, \beta = 0,\ \delta = 1$, accessibility only	L4	\$1.15	148	24.2	L4	\$0.10	40	21.4
11	$\alpha = 0, \beta = 2,\ \delta = 1$, distance & accessibility	L3	\$1.28	164	24.5	L3	\$0.33	60	20.1

Tables 24 ($\gamma = 0$) and 25 ($\gamma = 1$) detail optimal configurations over the planning horizon, in case 1. Columns give the year and (# of Physicians (PC), Building Capacity (BC)) for each of FP, IM, and P. At most 3 expansions over the planning horizon are observed for both $\gamma = 0$ and $\gamma = 1$. Our web site (**http://ise.tamu.edu/nsfhc)** gives configuration tables for all cases.

Table 26 breaks down the expected excess revenue reported in Table 23 for case 1 with $\gamma = 0$ $(\gamma = 1)$ in column 2 (3). Rows give E[Revenue], E[Excess demand penalty], E[Excess capacity penalty], facility operating cost (physicians, staff and maintenance costs), and facility configuration cost (land purchase, opening and expansion building costs). Elements in columns 2 (3) sum to give expected excess revenue for $\gamma = 0$ ($\gamma = 1$): \$1.31 (\$0.37) $\times 10^8$. Table 27 shows the detailed capacity configurations for cases 2 to 11, each involving two settings, γ=0 and γ=1.

Table 24. Optimal configuration, case 1, $\gamma = 0$

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	5	5	5	5	3	3
2012	5	5	5	5	3	3
2013	5	5	5	5	4	4
2014	5	5	5	5	4	4
2015	8	8	5	5	4	4
2016	8	8	5	5	4	4
2017	8	8	5	5	4	4
2018	8	8	8	8	5	5
2019	8	8	8	8	5	5
2020	8	8	8	8	5	5

Table 25. Optimal configuration, case 1, $\gamma = 1$

Period t	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	1	1	2	2	1	1
2012	2	2	3	3	1	1
2013	2	2	3	3	1	1
2014	2	2	3	3	1	1
2015	2	2	3	3	1	1
2016	2	2	3	3	1	1
2017	2	2	4	4	1	1
2018	2	2	4	4	1	1
2019	2	2	4	4	1	1
2020	2	2	4	4	1	1

Table 26. Case 1 revenue & costs $\times \$10^8$

Objective breakdown (10^8)	$\gamma = 0$	$\gamma = 1$
E[Revenue]	2.253	0.791
E[Excess demand penalty]	-0.114	-0.005
E[Excess capacity penalty]	-0.191	-0.157
Facility operating cost	-0.544	-0.230
Configuration cost	-0.090	-0.031

Case 1. Including capacity in the attraction model allows a facility to attract more demand as its capacity increases relative to other facilities. However, the capacity of existing facility E1 is so large, it dominates competitors when $\gamma = 1$, so that the new facility cannot attract enough demand to be very large. In contrast, if capacity is not included in the attraction model $(\gamma = 0)$, E1 loses the competitive advantage conferred by its larger capacity so that it cannot attract as much demand, allowing other facilities to attract larger market shares. New location L3 is close to growing tracts in southern College Station; it competes directly with E1 because the two locations are not far

from each other. As a result, new location L3 attracts more demand and enjoys a larger expected excess revenue when $\gamma = 0$.

Case 2. Set α = 2, enhancing the impact of physician ratings. The physician ratings assumed at the new facility are high, so L3 gains a competitive advantage when the ratings are emphasized, allowing it to attract even more demand than in case 1.

Case 3. Everyone is amenable to the attraction model; no one refuses to consider a new PCP. Existing facilities lose competitive advantage if long-time residents consider new PCPs; total demand for each service at each potential new location increases and, correspondingly, the number of PCP years increases. Long-term residents in the saturated tracts (mostly in the north and center of the county) pull the optimal location northward to L5 at the center of the county when $\gamma = 0$, and to L7 at the north of the county to exploit the lack of nearby existing facility capacity for $\gamma = 1$.

Case 4. Each physician can handle 4000 visits/year. As the capacity of each physician increases, the number of PCP years decreases. The capacity increment for IM (FP) is the largest (from 3564 to 4000) (smallest (from 3836 to 4000)) among the three services, so, as expected, it is the most (least) sensitive to this factor change.

Case 5. Physician ratings at the new location decrease from those as high as E2 to those at E1. The portion of demand it attracts from each tract decreases, reducing the number of PCP years.

Case 6. Limit the number of expansions/contractions to 0. In other cases, we restrict the number of expansions and contractions to 4. It turns out that this constraint is not tight, so RCSPP reduces to an SPP and CPLEX solves all cases at the root node. In Case 6, the CPLEX pre-solver eliminates expansion/contraction arcs, resulting in a network with only horizontal arcs (i.e., $\bar{A}^t_{ls}, l \in L, s \in S$, $t \in T$); it is solved at the root node in little run time. Providing larger capacities in early years does not induce increases in demands for IM or P that offset the associated costs, so this case foregoes expansions in later years, reducing the number of PCP years.

Case 7. Make location L8 an existing facility with 4 FP, 0 IM, 0 P, and PCP rating the same as firm E2. Because L8 does not provide IM or P, it does not compete with these services offered at L3. The market share for FP at L3 is smaller than in base Case 1, because demand is cannibalized by facilities E2 and L8, which have the same physician ratings.

Case 8. Fix L3 as an existing facility with the optimal configuration prescribed in base case 1. This case demonstrates use of our model as a heuristic to configure multiple new facilities, specifying one location at a time. The optimal additional location for the second facility is L5 (L4) with $\gamma = 0$ $(\gamma = 1)$.

Case 9. Set $\beta = 2$ and $\alpha = \delta = 0$, so each patient selects the closest facility. Because physician ratings are not considered, L3 loses an advantage it had in competing with E1, reducing PCP years.

Case 10. Set $\delta = 1$ and $\alpha = \beta = 0$, so the attraction model considers only accessibility. L4 is along a major highway amidst growing tracts in College Station, so it has a competitive advantage.

Case 11. Set $\beta = 2$, $\delta = 1$ and $\alpha = 0$, so the attraction model considers both distance and accessibility. L3 is close to growing tracts and not too far from a state highway; it attracts less demand (requiring fewer PCP years) than in base case 1 but more than in case 9, because it also has a favorable accessibility.

Cases 1, 2, 4-7, 9, and 11 prescribe L3, indicating that it is quite robust. For case 3 in which everyone is amenable to considering a new PCP, L5 (L7) is optimal when $\gamma = 0$ $(\gamma = 1)$. In case 8, after L3 is fixed as an existing facility, L5 (L4) is the optimal choice for a second facility when $\gamma = 0$ $(\gamma = 1)$. In case 10, which uses only the accessibility competition factor, L4 is optimal whether $\gamma = 0$ or $\gamma = 1$.

Table 27. Optimal capacity configuration for cases 2 -11

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	5	5	5	5	3	3
2012	5	5	5	5	3	3
2013	5	5	5	5	4	4
2014	5	5	5	5	4	4
2015	8	8	5	5	4	4
2016	8	8	5	5	4	4
2017	8	8	8	8	5	5
2018	8	8	8	8	5	5
2019	8	8	8	8	5	5
2020	8	8	8	8	5	5

(a) case 2, γ=0

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	2	2	2	2	1	1
2012	2	2	3	3	1	1
2013	2	2	3	3	1	1
2014	2	2	3	3	1	1
2015	2	2	3	3	1	1
2016	2	2	3	3	1	1
2017	2	2	4	4	1	1
2018	2	2	4	4	2	2
2019	2	2	5	5	2	2
2020	2	2	5	5	2	2

(b) case 2, γ=1

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	8	8	8	8	4	4
2012	8	8	8	8	5	5
2013	8	8	8	8	5	5
2014	8	8	8	8	5	5
2015	8	8	8	8	5	5
2016	8	8	8	8	5	5
2017	8	8	8	8	5	5
2018	10	10	8	8	5	5
2019	10	10	8	8	5	5
2020	10	10	10	10	5	5

(c) case 3, γ=0

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	2	2	5	5	3	3
2012	2	2	8	8	4	4
2013	2	2	8	8	4	4
2014	2	2	8	8	4	4
2015	2	2	8	8	4	4
2016	2	2	8	8	4	4
2017	2	2	8	8	4	4
2018	2	2	8	8	4	4
2019	2	2	8	8	4	4
2020	2	2	8	8	4	4

(d) case 3, γ=1

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	5	5	4	4	3	3
2012	5	5	4	4	3	3
2013	5	5	5	5	3	3
2014	5	5	5	5	4	4
2015	5	5	5	5	4	4
2016	8	8	5	5	4	4
2017	8	8	5	5	4	4
2018	8	8	5	5	4	4
2019	8	8	5	5	5	5
2020	8	8	8	8	5	5

(e) case 4, γ=0

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	1	1	2	2	1	1
2012	2	2	2	2	1	1
2013	2	2	2	2	1	1
2014	2	2	2	2	1	1
2015	2	2	2	2	1	1
2016	2	2	3	3	1	1
2017	2	2	3	3	1	1
2018	2	2	3	3	1	1
2019	2	2	3	3	1	1
2020	2	2	3	3	1	1

(f) case 4, γ=1

Table 27 Continued

Period *t*	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	5	5	5	5	3	3
2012	5	5	5	5	3	3
2013	5	5	5	5	3	3
2014	5	5	5	5	4	4
2015	5	5	5	5	4	4
2016	8	8	5	5	4	4
2017	8	8	5	5	4	4
2018	8	8	8	8	5	5
2019	8	8	8	8	5	5
2020	8	8	8	8	5	5

(g) case 5, γ=0

Period *t*	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	1	1	2	2	1	1
2012	2	2	2	2	1	1
2013	2	2	3	3	1	1
2014	2	2	3	3	1	1
2015	2	2	3	3	1	1
2016	2	2	3	3	1	1
2017	2	2	3	3	1	1
2018	2	2	4	4	1	1
2019	2	2	4	4	1	1
2020	2	2	4	4	1	1

(h) case 5, γ=1

Period *t*	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	5	5	5	5	4	4
2012	5	5	5	5	4	4
2013	5	5	5	5	4	4
2014	5	5	5	5	4	4
2015	5	5	5	5	4	4
2016	5	5	5	5	4	4
2017	5	5	5	5	4	4
2018	5	5	5	5	4	4
2019	5	5	5	5	4	4
2020	5	5	5	5	4	4

(i) case 6, γ=0

Period *t*	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	2	2	3	3	1	1
2012	2	2	3	3	1	1
2013	2	2	3	3	1	1
2014	2	2	3	3	1	1
2015	2	2	3	3	1	1
2016	2	2	3	3	1	1
2017	2	2	3	3	1	1
2018	2	2	3	3	1	1
2019	2	2	3	3	1	1
2020	2	2	3	3	1	1

(j) case 6, γ=1

Period *t*	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	4	4	5	5	3	3
2012	4	4	5	5	3	3
2013	5	5	5	5	4	4
2014	5	5	5	5	4	4
2015	5	5	5	5	4	4
2016	5	5	5	5	4	4
2017	5	5	5	5	4	4
2018	5	5	8	8	5	5
2019	5	5	8	8	5	5
2020	8	8	8	8	5	5

(k) case 7, γ=0

Period *t*	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	1	1	2	2	1	1
2012	1	1	3	3	1	1
2013	1	1	3	3	1	1
2014	1	1	3	3	1	1
2015	2	2	3	3	1	1
2016	2	2	3	3	1	1
2017	2	2	4	4	1	1
2018	2	2	4	4	1	1
2019	2	2	4	4	1	1
2020	2	2	4	4	1	1

(l) case 7, γ=1

Table 27 Continued

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	3	3	3	3	2	2
2012	4	4	3	3	2	2
2013	4	4	4	4	2	2
2014	4	4	4	4	2	2
2015	4	4	4	4	2	2
2016	4	4	4	4	3	3
2017	5	5	4	4	3	3
2018	5	5	5	5	3	3
2019	5	5	5	5	3	3
2020	5	5	5	5	3	3

(m) case 8, γ=0

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	0	0	1	1	0	0
2012	0	0	1	1	0	0
2013	0	0	1	1	0	0
2014	0	0	1	1	0	0
2015	0	0	2	2	0	0
2016	0	0	2	2	0	0
2017	1	1	2	2	0	0
2018	1	1	2	2	0	0
2019	1	1	2	2	0	0
2020	1	1	2	2	0	0

(n) case 8, γ=1

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	5	5	4	4	3	3
2012	5	5	5	5	3	3
2013	5	5	5	5	3	3
2014	5	5	5	5	4	4
2015	5	5	5	5	4	4
2016	8	8	5	5	4	4
2017	8	8	5	5	4	4
2018	8	8	8	8	5	5
2019	8	8	8	8	5	5
2020	8	8	8	8	5	5

(o) case 9, γ=0

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	1	1	2	2	0	0
2012	2	2	2	2	0	0
2013	2	2	2	2	0	0
2014	2	2	3	3	0	0
2015	2	2	3	3	0	0
2016	2	2	3	3	0	0
2017	2	2	3	3	0	0
2018	2	2	4	4	0	0
2019	2	2	4	4	0	0
2020	2	2	4	4	0	0

(p) case 9, γ=1

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	5	5	5	5	3	3
2012	5	5	5	5	3	3
2013	5	5	5	5	3	3
2014	5	5	5	5	3	3
2015	5	5	5	5	4	4
2016	5	5	5	5	4	4
2017	5	5	5	5	4	4
2018	8	8	5	5	4	4
2019	8	8	5	5	4	4
2020	8	8	8	8	4	4

(q) case 10, γ=0

Period	FP		IM		P	
t	PC	BC	PC	BC	PC	BC
2011	0	0	2	2	0	0
2012	0	0	3	3	0	0
2013	0	0	3	3	0	0
2014	0	0	3	3	0	0
2015	0	0	4	4	0	0
2016	0	0	5	5	0	0
2017	0	0	5	5	0	0
2018	0	0	5	5	0	0
2019	0	0	5	5	0	0
2020	0	0	5	5	0	0

(r) case 10, γ=1

Table 27 Continued

Period t	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	5	5	5	5	3	3
2012	5	5	5	5	3	3
2013	5	5	5	5	3	3
2014	5	5	5	5	4	4
2015	5	5	5	5	4	4
2016	8	8	5	5	4	4
2017	8	8	5	5	4	4
2018	8	8	8	8	5	5
2019	8	8	8	8	5	5
2020	8	8	8	8	5	5

(s) case 11, γ=0

Period t	FP		IM		P	
	PC	BC	PC	BC	PC	BC
2011	1	1	2	2	1	1
2012	2	2	2	2	1	1
2013	2	2	3	3	1	1
2014	2	2	3	3	1	1
2015	2	2	3	3	1	1
2016	2	2	3	3	1	1
2017	2	2	3	3	1	1
2018	2	2	4	4	1	1
2019	2	2	4	4	1	1
2020	2	2	4	4	1	1

(t) case 11, γ=0

4.3.3 Sensitivity with respect to the variability of physician capacity

We have not found publically available data to estimate the variance of physician capacity, so we analyze sensitivity parametrically with respect to its coefficient of variation, ρ, using results of Corollary 5. We test five values of ρ (0.0, 0.1, 0.2, 0.5 and 1.0) using Case 1 with ρ=0 as a benchmark; all cases in this subsection have $\gamma = 0$ and PCP capacity ~ $N\left(\eta\hat{k}, \eta\rho^2\hat{k}^2\right)$. Columns of Table 28 give case number, ρ, prescribed location, expected excess revenue, number of PCP years, and run time (in seconds).

Increasing values of ρ (0.0, 0.1, 0.2, 0.5 and 1.0) have no effect on location decisions. When capacity changes from a fixed number $\eta\hat{k}$ to a random variable following $N(\eta\hat{k}, \eta\rho^2\hat{k}^2)$, the variance of $D-K$ increases from σ^2 to $\sigma^2 + \eta\rho^2\hat{k}^2$. As a result, recourse values $E\lfloor D^+ \rfloor$ and $E\lfloor K^+ \rfloor$ increase, while the expected actual workload $E\left[\hat{N}\right]$ decreases, leading expected excess revenue to decrease. Expected excess revenue (Table 28) is quite sensitive to ρ, reducing relative

to the expected excess revenue of case 1 (with $\rho = 0$) by 1.3%, 4.9%, 24.6%, 65.8% as ρ increases in cases 12-15, respectively.

Table 28. Sensitivity to physician-capacity coefficient of variation

Case	Coefficient of variation for physician capacity	Opt. Loc.	Expected excess revenue ($\times 10^8$)	#PCP years	Run time (secs)
12	$\gamma = 0$, $\rho = 0.1$	L3	\$1.30	168	22.7
13	$\gamma = 0$, $\rho = 0.2$	L3	\$1.25	168	23.2
14	$\gamma = 0$, $\rho = 0.5$	L3	\$0.99	169	23.7
15	$\gamma = 0$, $\rho = 1.0$	L3	\$0.45	124	23.4

When ρ is small (e.g., cases 12 and 13), optimal capacity configurations stay the same as in case 1. As ρ increases, the distribution of capacity flattens out, causing larger penalties for both expected excess capacity (due to large values in the right tail of the capacity distribution occurring with higher likelihood) and demand (similarly, due to small values of capacity in the left tail). Increasing values of ρ have no effect on capacity configurations for FP and IM in our case studies. But because the revenue/visit for P is smallest, it is more sensitive than FP and IM to these larger penalty costs. In case 14, the optimal capacity of P adds one physician in 2017 to better balance penalty cost ramifications of the larger ρ value. In case 15, it is better not to open the P service at all because the capacity distribution is so flattened that the expected excess demand and capacity penalty costs exceed revenue dramatically.

4.3.4 Sensitivity with respect to penalty costs

To characterize sensitivity of expected excess revenue relative to penalties c^D and c^K, we normalize relative to revenue/visit, evaluating the sensitivity of base case 1 ($\gamma = 0$) with respect

to c^K/Π vs c^D/Π. The response surface (optimal, expected excess revenue) of Figure 5 is smooth, convex and nearly planar. Table 29 details the sensitivity of the number of PCP years to c^K/Π vs c^D/Π, showing that PCP years increases with c^D/Π (penalizing excess demand more) or decreasing c^K/Π (penalizing excess capacity less).

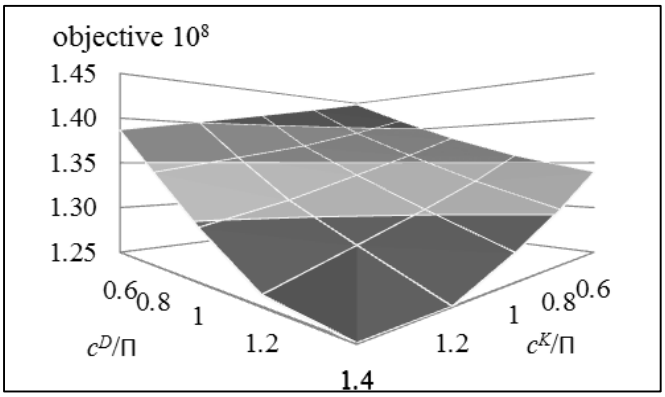

Figure 12. Objective sensitivity with respect to c^D and c^K

Table 29. Objective sensitivity w/r c^D and c^K

c^D/Π	c^K/Π				
	0.6	0.8	1	1.2	1.4
0.6	168	164	163	157	156
0.8	171	165	164	163	157
1	172	171	165	163	163
1.2	172	171	168	165	163
1.4	176	172	171	165	165

CHAPTER V

SMHFCP

This chapter presents our SMHFCP research. This chapter is organized in four sections. Section 5.1 presents our formulation of SMHFCP. Section 5.2 (3) proposes our column-generation heuristic (approximation method) and demonstrates the application of our algorithm on a numerical example. Our single facility model described in Chapter III is employed by each of our two solution methods to estimate the expected excess revenue for each *lp* assignment, $l \in L$, $p \in P$. Section 5.4 reports computational experiments of these two methods on a series of case studies that locate primary care centers in mid-Texas rural area from 20 zip codes to 101 zip codes.

5.1 SMHFCP formulation

The goal of this section is to prescribe our model for the multi-facility case, SMHFCP. We first introduce the decision variable and demand modeling in SMHFCP. Next, we describe a graph that defines adjacency of population centers, which we use to restrict assignments of population centers to facilities; and finally we present our location-allocation model, a mixed integer programming (MIP).

For the location-allocation model, we define binary decision variable

y_{lp} to be 1 if population center p, $p \in P$, is assigned to facility l, $l \in L$, and to be 0 otherwise;

and let $\mathbf{y}_l = (y_{lp})_{p \in P}$ be the vector of p assignments to facility l.

Given the assignment vector $\boldsymbol{y}_l = (y_{lp})_{p \in P}$ and following the derivation of stochastic patient demand in Section 4.2, we claim that the demand (in terms of number of patient visits/year) in facility l for service s in time period t, $D_{lst}(\boldsymbol{y}_l)$, can be well approximated by a normal distribution with mean μ_{lst} and variance σ_{lst} defined in (47) and (48) respectively:

$$\mu_{lst}(\boldsymbol{y}_l) = \sum_{p \in P} \sum_{g \in G} y_{lp} \; v_{gs} \overline{n}_{gpt} \overline{p}_{gs} \tag{47}$$

$$\sigma^2_{lst}(\boldsymbol{y}_l) = \sum_{p \in P} \sum_{g \in G} y_{lp} v^2_{gs} \overline{n}_{gpt} \overline{p}_{gs} \quad . \tag{48}$$

From our research for single-facility case (Chapter IV), If demand $D_{lst}(\boldsymbol{y}_l)$ is $Normal\left(\mu_{lst}(\boldsymbol{y}_l), \sigma^2_{lst}(\boldsymbol{y}_l)\right)$, then we can find the closed-form $\Omega_a(\boldsymbol{y}_l)$:

$$\Omega_a(\boldsymbol{y}_l) = \left(\Pi_{st} + c^K_{lst}\right)\mu_{lst}(\boldsymbol{y}_l) - c^K_{lst} k_a - c_a - (\Pi_{st} + c^K_{lst} + c^D_{lst}) E\left[D^+_{lst}(k_a)\right], \text{ where} \tag{49}$$

$$E\left[D^+_{lst}(k_a)\right] = \sigma_{lst}\phi(z) + \left(\mu_{lst}(\boldsymbol{y}_l) - k_a\right)\left[1 - \Phi(z)\right], \text{ where} \tag{50}$$

$z = \left(k_a - \mu_{lst}\right) / \sigma_{lst}$, $\phi(z) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$, and $\Phi(z) = \displaystyle\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ denote, respectively, standard normal and its probability density and cumulative distribution functions.

For a given l capacity configuration network with a given assignment $\boldsymbol{y}_l$, the stochastic, single facility capacity configuration problem for a given l, SSHFCP$_l$, is defined as

$$\overline{C}^*_l(\boldsymbol{y}_l) = -Min\left\{\sum_{s \in S} \sum_{t \in T} \sum_{a \in \overline{A}^t_{ls}} -\Omega_a(\boldsymbol{y}_l) x_a : x_a \in U_{ls}, s \in S\right\}, \tag{51}$$

where U_{ls} is defined in (34).

Next, we introduce the adjacency graph. Two population centers are defined to be adjacent if they share a border with each other. We represent adjacency by constructing graph $\tilde{G}(P, \tilde{A})$, with each node representing a population center and each edge (p, p') representing adjacency of population centers p and p'. For each $l \in L$, we construct an adjacency tree by specifying *the*

population center in which facility l is located (denoted as p_l) as the root node and performing a breadth-first-search (BFS) on $\tilde{G}$. The resulted adjacency tree for l is denoted as BFS_l. In BFS_l, each path from p_l to a descendant node p represents the shortest path from p_l to population center p in $\tilde{G}$, that is, containing the smallest number of edges between p and p_l.

We now introduce parameters:

q_{lp} ***depth*** of population center p in BFS_l, which reflects the number of population centers (i.e., edges) that a patient in p must traverse to seek service at facility l located in p_l (i.e., root node of BFS_l).

c_{lp} expected excess revenue that would result from assigning p to l throughout the planning horizon.

We assume that, if population center p' is assigned to facility l, then each $p \in P$ with $q_{lp} < q_{lp'}$ must also be assigned to facility l. This assignment assumes that no residents of a population center that is closer to p_l than some p that is assigned to p_l will want to travel to a more distant location. p_l We restrict the assignment of population centers to l by imposing a depth limit in BFS_l. Therefore, we define parameter

$\bar{q}_l$ maximum depth at which a population center can be assigned to facility l.

We now define our location-allocation model for SMHFCP:

(P): $$\max Z^* = \sum_{l \in L} \sum_{p \in P} c_{lp} y_{lp} \tag{52}$$

s.t. $$\sum_{l \in L} y_{lp} = 1 \qquad p \in P \tag{53}$$

$$q_{lp} y_{lp} \le \bar{q}_l \qquad l \in L \ , \ p \in P \tag{54}$$

$$y_{lp} \ge y_{lp'} \qquad l \in L \ , \ p, p' \in P \ , \ q_{lp} < q_{lp'} \tag{55}$$

$$y_{lp} \in \{0,1\} \qquad l \in L\ ,\ p \in P\ . \tag{56}$$

(52) maximizes the expected excess revenue from lp assignments; (53) makes sure that each population center p is assigned to some facility l; (54) imposes the depth limit; (55) requires a facility that serves a population center at a given depth also serves each one that is at a lesser depth.

Given the assignment vector $\boldsymbol{y}_l$, we can compute $\mu_{lst}(\boldsymbol{y}_l)$ and $\sigma^2_{lst}(\boldsymbol{y}_l)$ using (47) and (48), and calculate the total expected excess revenue associated with facility l relative to assignment $\boldsymbol{y}_l$ using (51), $\bar{C}(\boldsymbol{y}_l)$. However, we cannot determine c_{lp} *a prior* for each lp pair because the capacity prescribed at l can be determined after $\boldsymbol{y}_l$ is determined (i.e., all open facilities are known and all population centers have been assigned). The cost function is not separable for each $p \in P$. Thus, we propose to use $\bar{C}(\boldsymbol{y}_l)$ to replace $\sum_{p\in P} c_{lp} y_{lp}$. Define

$\upsilon_l^k = 1$ if assignment vector $\boldsymbol{y}_l^k \in K_l$ is selected, 0 otherwise, where

$$K_l = \left\{ \begin{array}{ll} \boldsymbol{y}_l \in \{0,1\}^{|P|} : q_{lp} y_{lp} \le \bar{q}_l & \text{for} \quad p \in P \\ \qquad\qquad\qquad\quad y_{lp} \ge y_{lp'} & \text{for} \quad p, p' \in P,\ q_{lp} < q_{lp'} \end{array} \right\}. \tag{57}$$

The master problem (MP) in our column-generation scheme is equivalent to (P):

$$\text{(MP): } \max Z^{MP} = \sum_{l\in L} \sum_{k\in K_l} \left(\sum_{p\in P} c_{lp} y_{lp}^k \right) \upsilon_l^k = \sum_{l\in L} \sum_{k\in K_l} \bar{C}(\boldsymbol{y}_l^k) \upsilon_l^k \tag{58}$$

$$\text{s.t.} \qquad \sum_{l\in L} \sum_{k\in K_l} (y_{lp}^k) \upsilon_l^k = 1 \qquad p \in P \tag{59}$$

$$\sum_{k\in K_l} \upsilon_l^k = 1 \qquad l \in L \tag{60}$$

$$\upsilon_l^k \in \{0,1\} \qquad l \in L\ ,\ k \in K_l\ . \tag{61}$$

5.2 Column-generation heuristic for SMHFCP

The goal of this section is to describe our column-generation heuristic. Subsections present the design of our algorithm, a special condition for optimality and a case study that demonstrates the use of our column-generation heuristic.

Because c_{lp}'s are not known, it is not helpful to branch on assignment variable y_{lp} for each lp pair in (P). The linear relaxation of (MP) is amenable to a Dantzig-Wolfe decomposition; and, therefore, branch and price (B&P). Note that each K_l includes the nullspace of $\mathbb{R}^{|P|}$, so it allows no population centers to be assigned to l, forcing l to stay closed. We form the linear relaxation of (58)-(61), further relaxing (59) from equalities to inequalities, so that the slack variable associated with population center p in (63) indicates whether p is either partially or fully served by facilities in L. The restricted master problem (RMP) we use is defined as

$$\text{(RMP): } \max Z^{RMP} = \sum_{l\in L}\sum_{k\in\bar{K}_l}\left(\sum_{p\in P} c_{lp}y_{lp}^k\right)\upsilon_l^k = \sum_{l\in L}\sum_{k\in\bar{K}_l}\bar{C}(\boldsymbol{y}_l^k)\upsilon_l^k \tag{62}$$

$$\text{s.t.} \quad \sum_{l\in L}\sum_{k\in\bar{K}_l}(y_{lp}^k)\upsilon_l^k \le 1 \qquad p\in P \tag{63}$$

$$0\le \upsilon_l^k \le 1 \qquad l\in L,\ k\in K_l \tag{64}$$

and (60),

where $\left\{\boldsymbol{y}_l^k, k\in\bar{K}_l\right\}\subseteq K_l$ is the subset of columns in K_l that have been entered in (RMP).

5.2.1 Column generation procedure

Let w_p^*, $p\in P$, and α_l^*, $l\in L$, be the dual solutions associated with (63) and (60), respectively, in (RMP). The objective of the subproblem associated with each l, (SP_l), is to prescribe the column that maximizes reduced cost to identify the most improving column to enter RMP:

$$(\text{SP}_l):\ \max\left\{Z_l^{SP}(\boldsymbol{y}_l)=\sum_{p\in P}(c_{lp}-w_p^*)y_{lp}-\alpha_l^*=\bar{C}(\boldsymbol{y}_l)-\boldsymbol{w}^*\boldsymbol{y}_l-\alpha_l^*:\boldsymbol{y}_l\in K_l\right\},\tag{65}$$

where K_l is defined in (57).

Note that $\bar{C}(\boldsymbol{y}_l)$ cannot be computed before $\boldsymbol{y}_l$ is specified; so it is difficult to tell if $Z_l^{SP}(\boldsymbol{y}_l)\le 0$ for all $\boldsymbol{y}_l\in K_l$ without enumerating all $\boldsymbol{y}_l\in K_l$, making (SP_l) challenging to solve to optimality. As a result, in a B&P framework, a stopping rule for column generation in each SP_l is difficult to define. So, Type II column generation (Wilhelm, 2001) is not appropriate. Therefore, we resort to type I column generation (Wilhelm, 2001). Thus, we generate a set of attractive columns and select the best.

From the solution of RMP, $\tilde{\upsilon}_l^k$, we set parameter ζ_p to be 1 if the p^{th} constraint in (63) is a strict inequality, and to be 0 otherwise. According to a duality relationship between the primal and dual problems of RMP, if $\zeta_p=1$ (i.e., $\sum_{l\in L}\sum_{k\in K_l}(y_{lp}^k)\tilde{\sigma}_l^k<1$), then the dual variable corresponding to the p^{th} row in (63) must be zero (i.e., $w_p^*=0$). Note that $\zeta_p=1$ indicates that population center p is not 100% assigned to facilities, but only a fraction is. We want to generate columns, each of which serves 100% of the population center, giving $\zeta_p=1$. Thus, the column generation problem associated with facility l is defined as

$$(\text{CGP}_l)\ \max\left\{Z_l^{CG}=\sum_{p\in P}\zeta_p y_{lp}:\boldsymbol{y}_l\in K_l\right\},\tag{66}$$

where K_l is defined in (57).

Figure 13 details the steps to perform our column-generation heuristic. Note that (SP_l) cannot be solved to optimality without enumerating on $\boldsymbol{y}_l\in K_l$. Even when the reduced cost for each $\boldsymbol{y}_l$ generated from (CGP_l) is non-positive for $l\in L$, we still cannot tell if every column in K_l

has a non-positive reduced cost; i.e., $Z_l^{SP}(\boldsymbol{y}_l) \le 0$. Thus, our method is a heuristic. After the column generation procedure terminates, some population centers may each remain unassigned to a particular facility. Thus, we need step 6 to assign such remaining population centers to the nearest opened facility.

Step 1. Initialization. For each $l \in L$, construct and enter the column that covers depth-1 population centers in BFS_l to (RMP).

Step 2. Solve (RMP) and obtain its dual variables.

Step 3. Solve (CGP$_l$). For each $l \in L$, if the optimal solution, $\boldsymbol{y}_l$, is not null, generate a column $\boldsymbol{y}_l$. If there is no column generated, go to Step 6.

Step 4. Solve (51) to obtain $\bar{C}(\boldsymbol{y}_l)$ for each $l \in L$ and compute the reduced cost of this column as defined in (65); i.e., $Z_l^{SP}(\boldsymbol{y}_l) = \bar{C}(\boldsymbol{y}_l) - \boldsymbol{w}^* \boldsymbol{y}_l - \alpha_l^*$.

Step 5. For each $l \in L$, if $Z_l^{SP}(\boldsymbol{y}_l) > 0$, enter $\boldsymbol{y}_l$ to RMP (i.e., add $\boldsymbol{y}_l$ as a column in the coefficient matrix of RMP and add a decision variable associated with $\boldsymbol{y}_l$). If there is no improving column to enter RMP, go to Step 6. Otherwise, go to Step 2.

Step 6. Examine the solution of RMP. If there exists a $p \in P$ that is not assigned to a particular facility, assign p to the nearest facility that is opened (i.e., in set L'). Otherwise, STOP.

Figure 13. Column-generation heuristic framework

5.2.2 Optimality of RMP under special conditions

Because of the difficulty of solving (SP_l) to optimality, we want to establish additional properties of expected excess revenue and reduced cost functions that would allow us to claim optimality of (RMP) upon termination.

Let $\bar{\boldsymbol{y}}_l^q$ be the vector that assigns all population centers up to depth q at facility l. Define $\bar{N}_{lq} = \sum_{p \in P} \bar{y}_{lp}^q$ to denote the total number of population centers assigned to facility l by vector $\bar{\boldsymbol{y}}_l^q = \{\bar{y}_{lp}^q\}_{p \in P}$. The average reduced cost per population center associated with q is computed as $(\bar{C}(\bar{\boldsymbol{y}}_l^q) - \boldsymbol{w}^* \bar{\boldsymbol{y}}_l^q - \alpha_l^*) / \bar{N}_{lq}$.

Claim: If the following three conditions are met for every $l \in L$, RMP is solved to optimality:

(i) $\bar{C}(\bar{\boldsymbol{y}}_l^q)$ is an increasing function of q;

(ii) $(\bar{C}(\bar{\boldsymbol{y}}_l^q) - \boldsymbol{w}^* \bar{\boldsymbol{y}}_l^q - \alpha_l^*) / \bar{N}_{lq} \le 0$ for $q = 1,..., \bar{q}_l$; and

(iii) The optimal solution of (CGP_l), $\hat{\boldsymbol{y}}_l$, provides a non-positive reduced cost in (RMP) (i.e., $Z_l^{SP}(\hat{\boldsymbol{y}}_l) = \bar{C}(\hat{\boldsymbol{y}}_l) - \boldsymbol{w}^* \hat{\boldsymbol{y}}_l - \alpha_l^* \le 0$).

Condition (i) requires that the expected excess revenue at one facility increases when population centers of greater and greater depths (i.e., more and more population centers) are assigned to this facility. Condition (ii) requires that on average, upon assigning one more population center to a particular facility l, the increment of dual objective associated with l is greater than or equal to the increment of expected excess revenue associated with l. Condition (iii) requires that the optimal solution of (CGP_l) provides a non-improving column to RMP. We want to show that these three conditions can guarantee that there is no improving column for any $l \in L$, indicating that the current (RMP) solution is optimal.

Proof: Let $\boldsymbol{y}_l^k \in K_l$ be an arbitrary feasible assignment vector to facility l. Then there are two

cases of the largest depth in this assignment.

- **Case 1**: $\boldsymbol{y}_l^k$ is a column with $\max_{p\in P}\{q_{lp}y_{lp}^k\} \le \max_{p\in P}\{q_{lp}\hat{y}_{lp}\}$ (i.e., a column with largest depth of population centers assigned to facility l that is no greater than the largest depth assigned according to $\hat{\boldsymbol{y}}_l$).

  From condition (i), $\bar{C}(\boldsymbol{y}_l^k) \le \bar{C}(\hat{\boldsymbol{y}}_l)$. Also, $\boldsymbol{y}_l^k$ must include a $p'l$ assignment that is not prescribed by $\hat{\boldsymbol{y}}_l$, so that $\zeta_{p'} = 0$. Otherwise, if $\zeta_{p'} = 1$, then p' must be assigned to l in $\hat{\boldsymbol{y}}_l$. $\zeta_{p'} = 0$ indicates that $w_{p'}^* \ge 0$, leading to the fact that $\boldsymbol{w}^*\boldsymbol{y}_l^k \ge \boldsymbol{w}^*\hat{\boldsymbol{y}}_l$. Therefore, $Z_l^{SP}(\boldsymbol{y}_l^k) = \bar{C}(\boldsymbol{y}_l^k) - \boldsymbol{w}^*\boldsymbol{y}_l^k - \alpha_l^* \le \bar{C}(\hat{\boldsymbol{y}}_l) - \boldsymbol{w}^*\hat{\boldsymbol{y}}_l - \alpha_l^* = Z_l^{SP}(\hat{\boldsymbol{y}}_l) \le 0$.

- **Case 2**: $\boldsymbol{y}_l$ is a column with $\max_{p\in P}\{q_{lp}y_{lp}^k\} > \max_{p\in P}\{q_{lp}\hat{y}_{lp}\}$ (i.e., a column with largest depth of population centers assigned to facility l that is greater than the largest depth assigned according to $\hat{\boldsymbol{y}}_l$).

  Condition (ii) shows that the average reduced cost per population center up to any depth is always negative, so the reduce cost will decrease when any additional population center at a larger depth is assigned to l in comparison with $\hat{\boldsymbol{y}}_l$. Therefore, assigning each additional population center in $\boldsymbol{y}_l^k$ - $\hat{\boldsymbol{y}}_l$ cannot make $Z_l^{SP}(\boldsymbol{y}_l^k)$ any larger than $Z_l^{SP}(\hat{\boldsymbol{y}}_l)$.

Any $\boldsymbol{y}_l^k \in K_l$ falls into either case 1 or case 2 so that it has a non-positive reduced cost. If these three conditions hold for every $l \in L$, we can conclude that there is no improving column to enter (RMP); so the current (RMP) is optimal. □

5.2.3 Numerical example

This subsection presents a case study that applies our column-generation heuristic to locate and configure capacity of primary care centers in Brazos County, TX. We use the Brazos County data

applied by the single-facility case (Chapter IV). Figure 14 is the map of Brazos County. In this numerical example, population centers are defined as census tracts. So the total number of population centers in this case study is $|P|=36$. 8 potential locations are identified ($|L|=8$) in Figure 14 denoted as drop tears. Three existing facilities are denoted as hospital signs.

q_{lp}, the depth of population center p in BFS_l, is shown in Table 30 for each $l \in L$. We set $\bar{q}_l = \bar{q} = 3$ for each $l \in L$, so that at most three population centers can be traversed by a patient seeking service at any facility.

Initial columns. We generate a set of initial columns; each assigns all depth $q=1$ population centers in BFS_l to facility l, $l \in L$. This procedure informs RMP which population centers are closest to each potential site, and leaves some population centers unassigned. We also generate all zero columns for each $l \in L$ with negative expected excess revenue (because all demand is excess) to represent the case in which no population centers are assigned to location l (because no facility is located at l).

Optimality conditions. In Table 31, we compute $\bar{C}(\bar{\mathbf{y}}_l^q)$, $\mathbf{w}^*\bar{\mathbf{y}}_l^q - \alpha_l^*$, $\bar{N}_q$ and $(\bar{C}(\bar{\mathbf{y}}_l^q) - \mathbf{w}^*\bar{\mathbf{y}}_l^q - \alpha_l^*)/\bar{N}_q$ for $q=1,2,3$. This shows that $\bar{C}(\bar{\mathbf{y}}_l^q)$ is monotonically increasing with q (depicted in Figure 15), and $(\bar{C}(\bar{\mathbf{y}}_l^q) - \mathbf{w}^*\bar{\mathbf{y}}_l^q - \alpha_l^*)/\bar{N}_q \le 0$ for all $l \in L$ and $q=1,...,\bar{q}$. Thus, Conditions (i) and (ii) are met. If we cannot find any improving column using (CGP_l), then we can terminate the column generation procedure and claim optimality of (RMP).

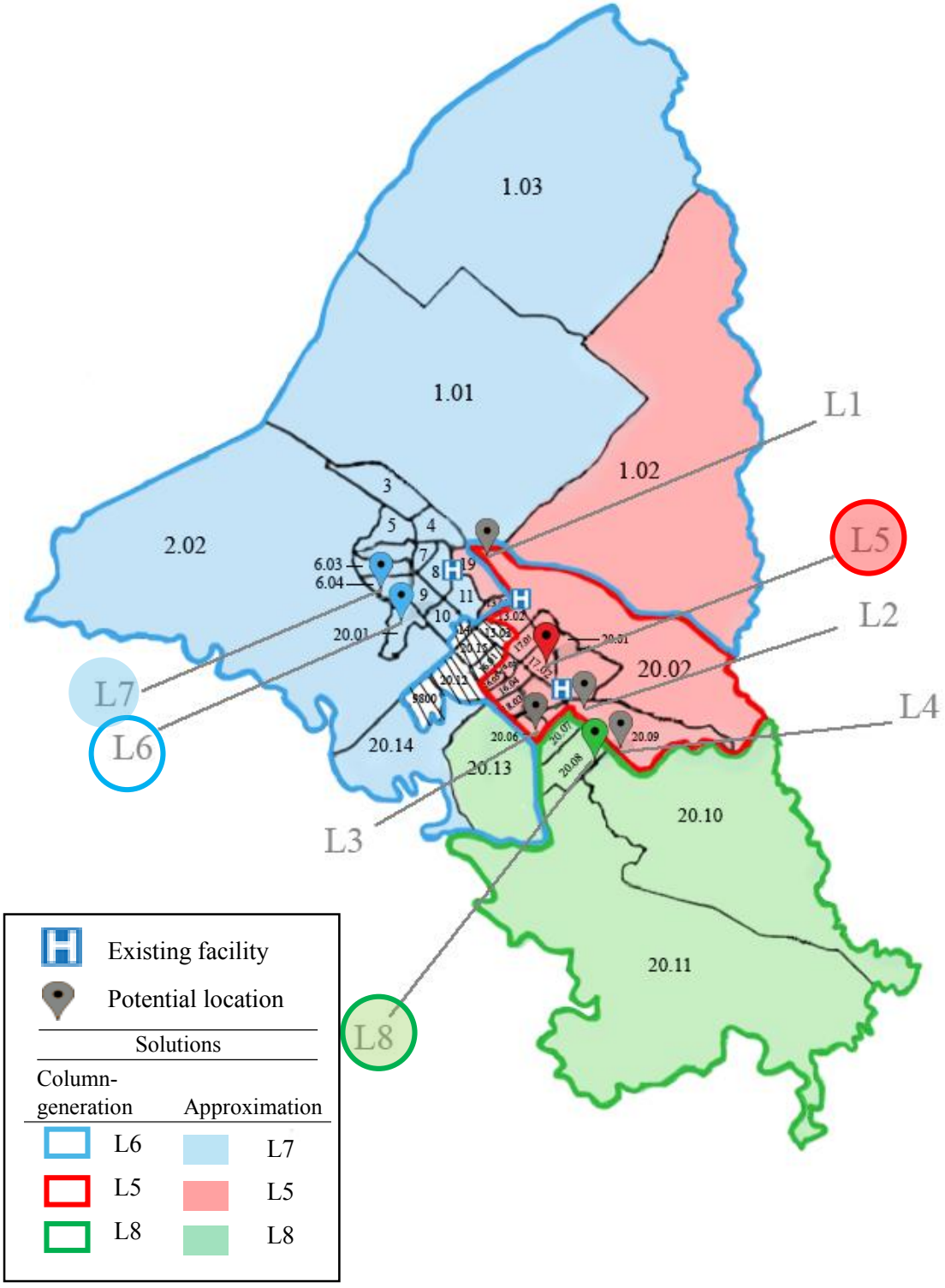

Figure 14. Column-generation heuristic and approximation method solutions

Table 30. q_{lp} in each BFS_l

l=0		*l=1*		*l=2*		*l=3*		*l=4*		*l=5*		*l=6*		*l=7*	
p	q_{lp}	*p*	q_{lp}	*p*	q_{lp}	*p*	q_{lp}	*p*	q_{lp}	*p*	q_{lp}	*p*	q_{lp}	*p*	q_{lp}
25	0	24	0	28	0	31	0	21	0	3	0	9	0	30	0
6	1	22	1	23	1	24	1	17	1	4	1	3	1	29	1
11	1	23	1	24	1	26	1	20	1	9	1	4	1	31	1
14	1	28	1	29	1	27	1	22	1	12	1	8	1	32	1
15	1	29	1	34	1	29	1	26	1	13	1	12	1	33	1
27	1	31	1	17	2	30	1	18	2	0	2	13	2	34	1
0	2	17	2	22	2	32	1	19	2	5	2	0	2	24	2
5	2	21	2	31	2	22	2	23	2	7	2	5	2	28	2
7	2	26	2	30	2	23	2	16	2	8	2	7	2	26	2
10	2	34	2	33	2	28	2	24	2	35	2	35	2	27	2
12	2	30	2	35	2	20	2	27	2	10	2	10	2	35	2
13	2	27	2	18	3	21	2	31	2	11	2	11	2	23	2
16	2	32	2	19	3	0	2	28	3	14	2	14	3	22	3
1	2	18	3	21	3	1	2	34	3	1	3	1	3	20	3
26	2	19	3	26	3	15	2	15	3	2	3	2	3	21	3
31	2	20	3	27	3	16	2	29	3	6	3	6	3	0	3
32	2	33	3	32	3	25	2	0	3	27	3	27	3	1	3
2	3	35	3	4	3	34	2	1	3	33	3	33	3	15	3
4	3	0	3	20	4	33	2	25	3	34	3	34	3	16	3
8	3	1	3	0	4	17	3	32	3	25	3	25	3	25	3
3	3	15	3	1	4	19	3	30	3	15	3	15	4	4	3
9	3	16	3	15	4	2	3	33	4	16	4	16	4	17	3
20	3	25	3	16	4	4	3	35	4	26	4	26	4	19	4
21	3	4	4	25	4	5	3	14	4	31	4	31	4	2	4
22	3	2	4	3	4	6	3	2	4	32	4	32	4	5	4
24	3	5	4	5	4	14	3	4	4	30	4	30	4	6	4
29	3	6	4	7	4	11	3	5	4	23	4	23	4	14	4
30	3	14	4	8	4	35	3	6	4	28	4	28	4	11	4
33	3	11	4	9	4	18	4	11	4	29	4	29	4	3	4
35	4	3	5	13	4	3	4	13	5	20	5	20	5	7	4
19	4	7	5	2	5	7	4	3	5	21	5	21	5	8	4
17	4	8	5	6	5	8	4	7	5	22	5	22	5	9	4
23	4	9	5	14	5	9	4	8	5	24	5	24	5	13	4
28	4	13	5	11	5	13	4	9	5	17	5	17	5	18	4
34	4	10	5	12	5	10	4	10	5	19	6	19	6	10	5
18	5	12	5	10	5	12	4	12	5	18	6	18	6	12	5

Table 31. Components of average reduced cost

l	$\bar{C}(\bar{\mathbf{y}}_l^q)$ (xE+06)			$\mathbf{w}^*\bar{\mathbf{y}}_l^q - \alpha_l^*$ (xE+06)			$\bar{N}_q$			$(\bar{C}(\bar{\mathbf{y}}_l^q) - \mathbf{w}^*\bar{\mathbf{y}}_l^q - \alpha_l^*)/\bar{N}_q$ (xE+06)		
	q=1	q=2	q=3	q=1	q=2	q=3	q=1	q=2	q=3	q=1	q=2	q=3
0	432	168	297	54.1	217	335	6	17	29	-1.83	-2.87	-1.30
1	71.8	173	248	91.7	199	266	6	13	23	-3.33	-2.03	-0.80
2	55.8	143	214	55.8	143	244	5	11	18	0.00	0.00	-1.68
3	117	229	289	148	266	313	7	19	28	-4.39	-1.96	-0.86
4	14.4	106	222	42.2	112	260	5	12	21	-5.57	-0.48	-1.81
5	43.8	122	215	65.9	122	219	5	13	21	-4.42	0.00	-0.17
6	47.9	119	214	71.0	122	217	5	12	20	-4.63	-0.21	-0.14
7	128	179	253	134	183	305	6	12	22	-0.97	-0.37	-2.35

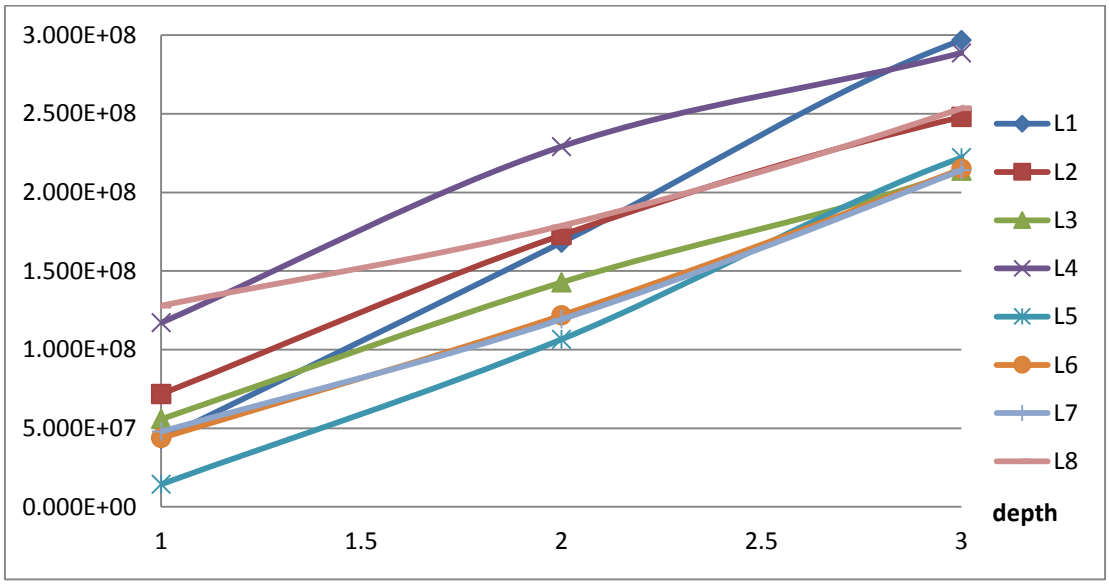

Figure 15. Expected excess revenue up to each depth level for each facility

Numerical results. For each l, we generate columns according to (CGP$_l$) with $\bar{q} = 1,2,3$ and enter the columns that have positive reduced costs into (RMP). This provides more choices of attractive columns. So at most three columns associated with facility l will be entered. We stop the column generation procedure when $Z_l^{SP}(\boldsymbol{y}_l) \le 0$ for all $l \in L$ and all $\bar{q} \in \{1,2,3\}$.

In our example, fortuitously, the optimal solution to RMP is integral. Therefore, it is the optimal solution to (MP). This optimal solution opens facilities L5, L6 and L8. The optimal assignment of population centers to each new facility is highlighted in Table 32. The expected excess revenues associated with L5, L6 and L8 are \$1.14E+08, \$1.64E+08 and \$8.36E+07, respectively, giving a total of Z^{RMP*} = \$3.62E+08. The run time is 323.53 seconds.

Table 32. Assignments using column-generation heuristic

L5		L6		L8	
p	q_{lp}	p	q_{lp}	p	q_{lp}
21	0	9	0	30	0
17	1	3	1	29	1
20	1	4	1	31	1
22	1	8	1	32	1
26	1	12	1	33	1
18	2	13	2	34	1
19	2	0	2	24	2
23	2	5	2	28	2
16	2	7	2	26	2
24	2	35	2	27	2
27	2	10	2	35	2
31	2	11	2	23	2
28	3	14	3	22	3
34	3	1	3	20	3
15	3	2	3	21	3
29	3	6	3	0	3
0	3	27	3	1	3
1	3	33	3	15	3
25	3	34	3	16	3
32	3	25	3	25	3
\$1.14E+08		\$1.64E+08		\$8.36E+07	

Figure 14 shows the assignments of population centers to facilities on the map of Brazos County. The population centers outlined in red, blue and green are assigned to L5, L6 and L8 respectively.

Figure 16 shows the average expected excess revenue per population center up to each depth ($q = 1,2,3$) for each potential location. An interesting finding is that population centers assigned to each l^* in the optimal solution is up to the depth that has the largest average expected excess revenue per population center. For example, L5 and L6 both have the largest average expected excess revenue per population center at depth 3, and the largest depth of population centers assigned to L5 and L6 is 3. Also, L8 has the largest average expected excess revenue per population center at depth 1, and the largest depth of population centers assigned to L8 is 1.

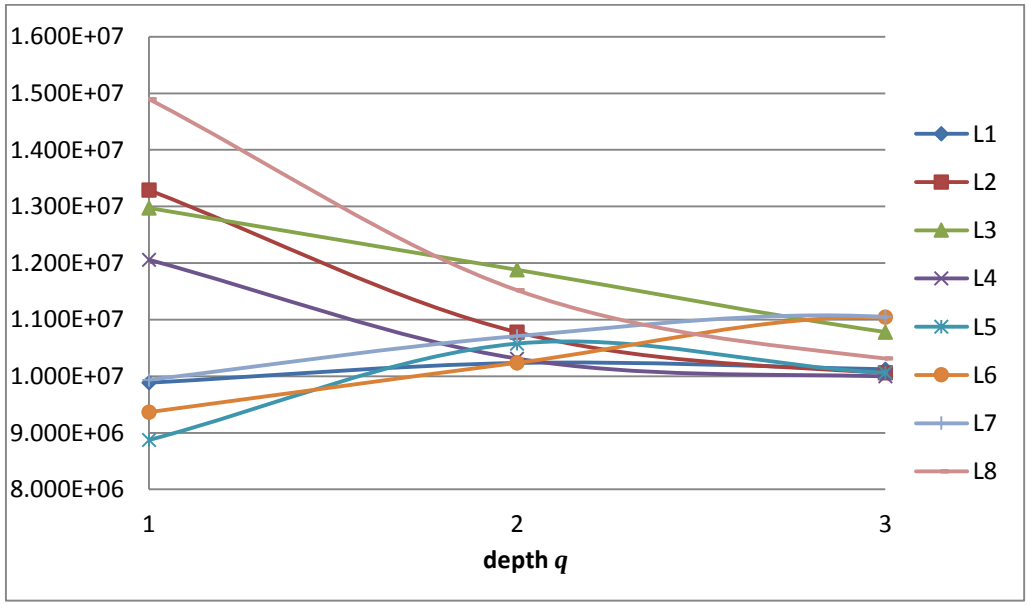

Figure 16. Average expected excess revenue per population center up to depth q for each facility

5.3 Approximation method

The goal of this section is to describe our approximation method. Subsections present the detailed algorithm, the bounding procedure and demonstrate the use of our approximation method using the case as described in Section 6.2.3.

Because c_{lp} cannot be defined knowing just an lp pair, we transform (P) into an equivalent formulation (MP) that gives the expected excess revenue of assignment vector $\boldsymbol{y}_l$ relative to l, $\bar{C}(\boldsymbol{y}_l)$ and Type I column generation (Wilhelm, 2001). Thus, this section describes a new method to approximate the values of c_{lp}'s so that we can solve (P) directly.

5.3.1 Approximation steps

Our approach is based on finding the average incremental expected excess revenue associated with each facility, which approximates the value of c_{lp} for each lp pair.

We approximate c_{lp} at each depth in BFS_l, $l \in L$. First, compute the average incremental expected excess revenue per population center at depth q in BFS_l, $\tilde{c}_{ql}$:

$$\tilde{c}_{ql} = \frac{\bar{C}(\bar{\boldsymbol{y}}_l^q) - \bar{C}(\bar{\boldsymbol{y}}_l^{q-1})}{\bar{N}_q - \bar{N}_{q-1}}, \quad q = 1,2,...\bar{q} \tag{67}$$

Second, replace c_{lp} with $\tilde{c}_{ql}$ if $q_{lp} = q$; i.e., all population centers at depth q in BFS_l adopt the same cost per population center.

Let $\boldsymbol{y}_l$ be an assignment vector with $\max_{p \in P}\{q_{lp} y_{lp}\} = q'$. Then, if we approximate $\bar{C}(\boldsymbol{y}_l)$ with a series of $\tilde{c}_{ql}$'s up to depth q', we can make sure that $\bar{C}(\boldsymbol{y}_l^{q'-1}) = \sum_{q=1}^{q'-1} \tilde{c}_{ql}$. The approximation is accurate for the aggregation of population centers up to depth $q'-1$ and may

incorporate errors only relative to the population centers at depth q'.

Third, solve (MP) using the approximated objective coefficients.

5.3.2 Bounding

This section derives two upper bounds on the optimal solution as well as the gap between our approximation method solution and the optimal solution.

For a given $\boldsymbol{y}_l = (y_{lp})_{p \in P}$, the demand attracted to facility l for service s in time period t in terms of number of patient visits/year, D_{lst}, follows a normal distribution with mean and variance defined in (47) and (48). For each $\bar{a} \in \bar{A}_{ls}^{t}$, $l \in L$, $s \in S$, $t \in T$, the excess revenue associated with $\bar{a}$ is defined in (49) as:

$$\Omega_{\bar{a}}(\boldsymbol{y}_l) = \Pi_{st} E\left[\hat{N}_{lst}(k_{\bar{a}})\right] - c_{lst}^{K} E\left[K_{lst}^{+}(k_{\bar{a}})\right] - c_{lst}^{D} E\left[D_{lst}^{+}(k_{\bar{a}})\right] - c_{\bar{a}}.$$

Upper Bound 1. We want to find an upper bound on $\Pi_{st} E\left[\hat{N}_{lst}(k_{\bar{a}})\right]$ and a lower bound on $c_{lst}^{K} E\left[K_{lst}^{+}(k_{\bar{a}})\right] + c_{lst}^{D} E\left[D_{lst}^{+}(k_{\bar{a}})\right]$, so as to derive an upper bound on $\Omega_{\bar{a}}(\boldsymbol{y}_l)$ for each arc $\bar{a}$, and, further, an upper bound on $\bar{C}(\boldsymbol{y}_l^{q})$ for each $q = 1,...,\bar{q}$.

Claim: Given $\boldsymbol{y}_l$, $E\left[D_{lst}^{+}(k_{\bar{a}})\right]$ is convex in $k_{\bar{a}}$.

Proof. As $D_{lst} \sim Normal(\lambda_{lst}, \sigma_{lst}^{2})$, $E\left[D_{lst}^{+}(k_{\bar{a}})\right] = \sigma_{lst}\phi(z) + (\lambda_{lst} - k_{\bar{a}})[1 - \Phi(z)]$, where $z = \dfrac{k_{\bar{a}} - \lambda_{lst}}{\sigma_{lst}}$, $\phi(z) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^{2}}$, and $\Phi(z) = \int_{-\infty}^{z} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z^{2}}{2}} dz$ denote, respectively, standard normal and its probability density and cumulative distribution functions (Chapter IV).

The first derivative is $\dfrac{dE\left[D_{lst}^{+}(k_{\bar{a}})\right]}{dk_{\bar{a}}} = \sigma_{lst}\phi'(z)\dfrac{dz}{dk_{\bar{a}}} - 1 + \Phi(z) - (\lambda_{lst} - k_{\bar{a}})\phi(z)\dfrac{dz}{dk_{\bar{a}}}$.

Note that $\phi'(z) = -z\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2} = -z\phi(z)$ and $\frac{dz}{dk_{\bar{a}}} = \frac{1}{\sigma_{lst}}$, so that

$$\begin{aligned}\frac{dE\left[D^+_{lst}(k_{\bar{a}})\right]}{dk_{\bar{a}}} &= -z\sigma_{lst}\phi(z)\frac{1}{\sigma_{lst}} - 1 + \Phi(z) - (\lambda_{lst} - k_{\bar{a}})\phi(z)\frac{1}{\sigma_{lst}} \\ &= -z\phi(z) - 1 + \Phi(z) + z\phi(z) \\ &= -1 + \Phi(z) \le 0 \text{ for } z \in [-\infty, +\infty].\end{aligned} \tag{68}$$

The second derivative is

$$\frac{dE^2\left[D^+_{lst}(k_{\bar{a}})\right]}{dk^2_{\bar{a}}} = \phi(z)\frac{dz}{dk_{\bar{a}}} = \frac{\phi(z)}{\sigma_{lst}} \ge 0\,. \tag{69}$$

Thus, $E\left[D^+_{lst}(k_{\bar{a}})\right]$ is convex in $k_{\bar{a}}$. □

Now, we substitute $E\left[K^+_{lst}(k_{\bar{a}})\right]$ with $E\left[D^+_{lst}(k_{\bar{a}})\right]$ using $E\left[K^+_{lst}(k_{\bar{a}})\right] = E\left[D^+_{lst}(k_{\bar{a}})\right] - \lambda_{lst} + k_{\bar{a}}$, as derived in Chapter IV:

$$\begin{aligned}c^K_{lst}E\left[K^+_{lst}(k_{\bar{a}})\right] + c^D_{lst}E\left[D^+_{lst}(k_{\bar{a}})\right] &= c^K_{lst}\left(E\left[D^+_{lst}(k_{\bar{a}})\right] - \lambda_{lst} + k_{\bar{a}}\right) + c^D_{lst}E\left[D^+_{lst}(k_{\bar{a}})\right] \\ &= (c^K_{lst} + c^D_{lst})E\left[D^+_{lst}(k_{\bar{a}})\right] + c^K_{lst}k_{\bar{a}} - c^K_{lst}\lambda_{lst}\,.\end{aligned} \tag{70}$$

From (69), it is obvious that the second derivative of (70) with respect to $k_{\bar{a}}$ is $(c^K_{lst} + c^D_{lst})\frac{\phi(z)}{\sigma_{lst}} \ge 0$. Thus, the penalty cost on arc $\bar{a}$, $c^K_{lst}E\left[K^+_{lst}(k_{\bar{a}})\right] + c^D_{lst}E\left[D^+_{lst}(k_{\bar{a}})\right]$, is a convex function of $k_{\bar{a}}$ as well. Given l, s, t, and assuming that physician capacity is continuous on $[0, +\infty)$ and using the claim, we can find a lower bound on the penalty cost. Using (68), we can derive that the first derivative of (70) with respect to $k_{\bar{a}}$ is

$$\frac{d\left(c^K_{lst}E\left[K^+_{lst}(k_{\bar{a}})\right] + c^D_{lst}E\left[D^+_{lst}(k_{\bar{a}})\right]\right)}{dk_{\bar{a}}} = (c^K_{lst} + c^D_{lst})(-1 + \Phi(z)) + c^K_{lst}\,. \tag{71}$$

By setting (71) to 0, we can solve for k_{lst}^* such that $k_{lst}^* = \arg\min\left\{c_{lst}^K E\left[K_{lst}^+(k_{\bar{a}})\right] + c_{lst}^D E\left[D_{lst}^+(k_{\bar{a}})\right], k_{\bar{a}} \geq 0\right\}$. $z_{lst}^* = \dfrac{k_{lst}^* - \lambda_{lst}}{\sigma_{lst}}$, then $\Phi(z_{lst}^*) = \dfrac{c_{lst}^D}{c_{lst}^K + c_{lst}^D}$, and the minimum value of the penalty cost is

$$
\begin{aligned}
& c_{lst}^K E\left[K_{lst}^+(k_{\bar{a}}^*)\right] + c_{lst}^D E\left[D_{lst}^+(k_{\bar{a}}^*)\right] \\
&= (c_{lst}^K + c_{lst}^D) E\left[D_{lst}^+(k_{\bar{a}}^*)\right] + c_{lst}^K k_{\bar{a}}^* - c_{lst}^K \lambda_{lst} \\
&= (c_{lst}^K + c_{lst}^D)\left(\sigma_{lst}\phi(z_{lst}^*) + (\lambda_{lst} - k_{\bar{a}}^*)[1 - \Phi(z_{lst}^*)]\right) + c_{lst}^K k_{\bar{a}}^* - c_{lst}^K \lambda_{lst} \\
&= (c_{lst}^K + c_{lst}^D)\sigma_{lst}\phi(z_{lst}^*) + (c_{lst}^K + c_{lst}^D)(\lambda_{lst} - k_{\bar{a}}^*)\frac{c_{lst}^K}{c_{lst}^K + c_{lst}^D} + c_{lst}^K k_{\bar{a}}^* - c_{lst}^K \lambda_{lst} \\
&= (c_{lst}^K + c_{lst}^D)\sigma_{lst}\phi(z_{lst}^*) + c_{lst}^K(\lambda_{lst} - k_{\bar{a}}^*) + c_{lst}^K(k_{\bar{a}}^* - \lambda_{lst}) \\
&= (c_{lst}^K + c_{lst}^D)\sigma_{lst}\phi(z_{lst}^*).
\end{aligned}
\tag{72}
$$

Note that $\hat{N}_{lst}(k_{\bar{a}}) = \min\{D_{lst}, k_{\bar{a}}\}$ indicates that $E\left[\hat{N}_{lst}(k_{\bar{a}})\right] \leq E\left[D_{lst}\right] = \lambda_{lst}$ and $E\left[\hat{N}_{lst}(k_{\bar{a}})\right] \leq k_{\bar{a}}$, so an upper bound on $E\left[\hat{N}_{lst}(k_{\bar{a}})\right]$ is at hand:

$$
E\left[\hat{N}_{lst}(k_{\bar{a}})\right] \leq \min\{\lambda_{lst}, k_{\bar{a}}\} \tag{73}
$$

Using (72) and (73), we define $\hat{\Omega}_{\bar{a}}(\boldsymbol{y}_l)$ as an upper bound on $\Omega_{\bar{a}}(\boldsymbol{y}_l)$ for $\bar{a} \in \bar{A}_{ls}^t$:

$$
\begin{aligned}
\Omega_{\bar{a}}(\boldsymbol{y}_l) &= \Pi_{st} E\left[\hat{N}_{lst}(k_{\bar{a}})\right] - c_{lst}^K E\left[K_{lst}^+(k_{\bar{a}})\right] - c_{lst}^D E\left[D_{lst}^+(k_{\bar{a}})\right] - c_{\bar{a}} \\
&\leq \Pi_{st} \min\{\lambda_{lst}, k_{\bar{a}}\} - c_{lst}^K E\left[K_{lst}^+(k_{\bar{a}}^*)\right] - c_{lst}^D E\left[D_{lst}^+(k_{\bar{a}}^*)\right] - c_{\bar{a}} \\
&= \Pi_{st} \min\{\lambda_{lst}, k_{\bar{a}}\} - (c_{lst}^K + c_{lst}^D)\sigma_{lst}\phi(z_{lst}^*) - c_{\bar{a}} \\
&:= \hat{\Omega}_{\bar{a}}(\boldsymbol{y}_l).
\end{aligned}
\tag{74}
$$

Let $\hat{C}(\boldsymbol{y}_l)$ be the maximum expected excess revenue using $-\hat{\Omega}_{\bar{a}}(\boldsymbol{y}_l)$ as the cost associated with arc $\bar{a} \in \bar{A}_{ls}^t$, $s \in S$, which is prescribed by the shortest path $\hat{p}$. Let $\bar{p}$ be the shortest path that prescribes $\bar{C}(\boldsymbol{y}_l)$ using $-\Omega_{\bar{a}}(\boldsymbol{y}_l)$ as the cost associated with arc $\bar{a} \in \bar{A}_{ls}^t$, $s \in S$. So $\hat{C}(\boldsymbol{y}_l) = \sum_{\bar{a} \in \bar{p}} \hat{\Omega}_{\bar{a}}(\boldsymbol{y}_l) \geq \sum_{\bar{a} \in \bar{p}} \hat{\Omega}_{\bar{a}}(\boldsymbol{y}_l) \geq \sum_{\bar{a} \in \bar{p}} \Omega_{\bar{a}}(\boldsymbol{y}_l) = \bar{C}(\boldsymbol{y}_l)$. The first inequality holds

because $\hat{p}$ is the shortest path using $\hat{\Omega}_{\bar{a}}$ as the cost associated with $\bar{a}$. The second inequality holds because $\hat{\Omega}_{\bar{a}}$ is an upper bound on $\Omega_{\bar{a}}$.

Upper Bound 2. Upon defining penalty costs, we do not set $c_{lst}^K = c_{lst}^D$ because that would make the total penalty the same for all possible outcomes of demand. Utilizing the differences between c_{lst}^K and c_{lst}^K, we show that $\breve{\Omega}_{\bar{a}}(\boldsymbol{y}_l)$ is an upper bound on $\Omega_{\bar{a}}(\boldsymbol{y}_l)$ for $\bar{a} \in \bar{A}_{ls}^t$:

$$\begin{aligned}\Omega_{\bar{a}}(\boldsymbol{y}_l) &= \Pi_{st} E\left[\hat{N}_{lst}(k_{\bar{a}})\right] - c_{lst}^K E\left[K_{lst}^+(k_{\bar{a}})\right] - c_{lst}^D E\left[D_{lst}^+(k_{\bar{a}})\right] - c_{\bar{a}} \\ &\le \Pi_{st} E\left[\hat{N}_{lst}(k_{\bar{a}})\right] - \min\left\{c_{lst}^K, c_{lst}^D\right\} E\left[K_{lst}^+(k_{\bar{a}}^*)\right] - \min\left\{c_{lst}^K, c_{lst}^D\right\} E\left[D_{lst}^+(k_{\bar{a}}^*)\right] - c_{\bar{a}} \qquad (75) \\ &:= \breve{\Omega}_{\bar{a}}(\boldsymbol{y}_l).\end{aligned}$$

Let $\breve{C}(\boldsymbol{y}_l^q)$ be the maximum expected excess revenue using $\breve{\Omega}_{\bar{a}}(\boldsymbol{y}_l)$ as the cost associated with arc $\bar{a} \in \bar{A}_{ls}^t$, $s \in S$. Using the same argument to establish UB1, we can claim that $\breve{C}(\boldsymbol{y}_l^q)$ is an upper bound on $C(\boldsymbol{y}_l^q)$.

Gap. Figure 17 demonstrates the bounds for facility l as a function of depth q. The green line comprises straight line segments between $\bar{C}(\bar{\boldsymbol{y}}_l^q)$ and $\bar{C}(\bar{\boldsymbol{y}}_l^{q+1})$, in which the slope of segment q represents the average expected excess revenue per population in depth p in BFS_l. The blue (red) dashed line segments shows the upper bound defined by $\hat{C}(\bar{\boldsymbol{y}}_l^q)$ ($\breve{C}(\bar{\boldsymbol{y}}_l^q)$). The black curve shows the actual expected excess revenue as more and more population centers are assigned to l; it is bounded by both the blue and red line segments.

5.3.3 Numerical example

Using the Brazos County example as described in Chapter IV, the solution from our approximation method selects L5, L7 and L8 as the new facility sites. Table 33 presents the $\tilde{c}_{ql}$ values that are computed and used in the approximation method. Table 34 shows the assignment of population

centers to each facility. The expected excess revenues for L5, L7 and L8 are \$1.13E+08, \$1.22E+08 and \$9.97E+07, respectively, for a total expected excess revenue of \$3.35E+08. The gap between the objective function values prescribed by our approximation (\$3.35E+08) and our column -generation heuristic (\$3.62E+08) is 7.46%, showing that our approximation method gives a close-to-optimal solution. The run time is 57.26 seconds, which is only 17.7% of the run time required by our column-generation heuristic.

The gap for each lq combination is defined as:

$$\min\left\{\frac{\hat{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)},\frac{\breve{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)}\right\},$$

and the gap associated with facility l is

$$\max_{q=1,\ldots,q_l}\left\{\min\left\{\frac{\hat{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)},\frac{\breve{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)}\right\}\right\}. \tag{76}$$

Now, because each new facility may serve a portion of $|P|$ population centers, the gap that the approximation method prescribes is a convex combination of the gaps of all new facilities, which will be no larger than the maximum gap prescribed by all facilities:

$$\text{gap\_app}=\max_{l\in L'}\left\{\max_{q=1,\ldots,q_l}\left\{\min\left\{\frac{\hat{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)},\frac{\breve{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)}\right\}\right\}\right\}. \tag{77}$$

The assignment of population centers to locations prescribed by our approximation method is shown by Figure 14. The population centers colored in red, blue and green blocks are assigned to L5, L7 and L8 respectively. The assignment of population centers to each new facility is highlighted in Table 34.

5.4 Computational experiments

The goal of this section is to compare our column-generation heuristic and approximation method

in application to a case study, which is based on a mid-Texas rural area that comprises 101 zip codes across 15 counties (Figure 18). We use each zip code as a population center and identify 26 potential locations. Figure 8 also shows the population density level of each zip code. We identify more potential locations in population-dense zip codes than in sparsely populated ones. Fixed cost c_a on arc $a \in A_{ts}^t$ and demand associated with each zip code are determined as described in Chapter IV. We assume that there is no competitor in the market.

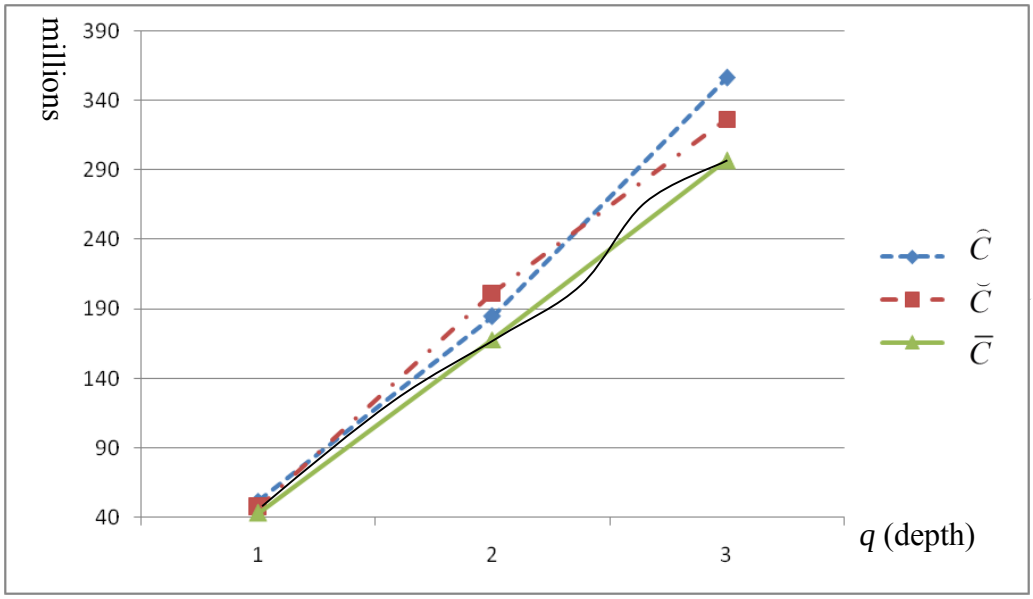

Figure 17. Bounding for approximation method

Table 33. Average incremental expected excess revenue per population center

	$\bar{C}(\bar{\mathbf{y}}_l^q)$			$\tilde{c}_{ql}$		
l	q=1	q=2	q=3	q=1	q=2	q=3
0	4.32E+07	1.68E+08	2.97E+08	7.192E+06	1.14E+07	1.07E+07
1	7.18E+07	1.73E+08	2.48E+08	1.196E+07	1.44E+07	7.52E+06
2	5.58E+07	1.43E+08	2.14E+08	1.115E+07	1.45E+07	1.02E+07
3	1.17E+08	2.29E+08	2.89E+08	1.673E+07	9.33E+06	6.63E+06

Table 33 Continued

4	1.44E+07	1.06E+08	2.22E+08	2.872E+06	1.32E+07	1.28E+07
5	4.38E+07	1.22E+08	2.15E+08	8.753E+06	9.75E+06	1.17E+07
6	4.79E+07	1.19E+08	2.14E+08	9.575E+06	1.02E+07	1.19E+07
7	1.28E+08	1.79E+08	2.53E+08	2.133E+07	8.47E+06	7.47E+06

Table 34. Assignments using the approximation method

L5		L7		L8	
p	q_{lp}	p	q_{lp}	p	q_{lp}
21	0	9	0	30	0
17	1	3	1	29	1
20	1	4	1	31	1
22	1	8	1	32	1
26	1	12	1	33	1
18	2	13	2	34	1
19	2	0	2	24	2
23	2	5	2	28	2
16	2	7	2	26	2
24	2	35	2	27	2
27	2	10	2	35	2
31	2	11	2	23	2
28	3	14	3	22	3
34	3	1	3	20	3
15	3	2	3	21	3
29	3	6	3	0	3
0	3	27	3	1	3
1	3	33	3	15	3
25	3	34	3	16	3
32	3	25	3	25	3
$1.13E+08		$1.22E+08		$9.97E+07	

5.4.1 Case study results

To evaluate the run times of our methods relative to problem size as measured by the number of zip codes, we start with zip code 78502 (i.e., the zip code in which L18 is located as a "seed") and construct five adjacency graphs, forming five test cases by taking the first 20, 40, 60, 80 and 101 population centers (i.e., zip codes) in BFS order, respectively. Each successive case includes the zip codes used in the previous case as well as an additional set of locations.

Table 35 summarizes the size of problem in each case using each method. The first column identify the case number. The second column gives $|P|$, the number of population centers in each case. The third column gives $|L|$, the number of potential locations in each case. A pair of columns is associated with each method, showing the number of decision variables (labeled "# variables") and number of constraints (labeled "# constraints"), respectively. The number of decision variables in our column-generation heuristic is the same as the number of columns that have been generated. The number of constraints in our column-generation heuristic is $|L|+|P|$, the total number of rows in constraints (60) and (63). The number of decision variables in our approximation method is $|L|*|P|$, because y_{lp} is defined for each lp pair. The number of constraints in our approximation method is $O(|L|*|P|^2)$ because of the set of constraints in (55).

Table 36 (37) describes the solutions from our column-generation heuristic (approximation method) in these five cases. The first column in Table 36 (37) lists potential locations. Five pairs of columns follows, one for each test case. Each pair of columns gives the case number, and the number of new facilities prescribed, $|L'|$, for the associated case. The first column is # PCP years, which is the sum of the number of primary care physicians prescribed each year throughout the planning horizon. The second column is c_bar, which is the total expected excess revenue calculated using (51). If, in a particular case, a potential location is in a zip code that is not included in this case, its #PCP years and c_bar entries are both encoded by "-". If, in a

particular case, a location has its #PCP years equal to zero (0), it is a potential location, but is not selected in the solution, so its c_bar must be 0. Row 29, labeled "Total", gives the total # PCP years and total expected excess revenue in each case. Row 30, labeled "Run time", records the run time in each case in seconds. In Table 37, row 31 (32), labeled "# PCP years" ("c_bar_gap"), gives the differences between #PCP year (total expected excess revenue) of our two methods as a percentage of #PCP years (total expected excess revenue) prescribed by our column-generation heuristic.

Table 35. Problem sizes

Case	$\|P\|$	$\|L\|$	Column generation heuristic		Approximation method	
			# variables	# constraints	# variables	# constraints
			(# cols)	($\|L\|+\|P\|$)	($\|L\|*\|P\|$)	O ($\|L\|*\|P\|^2$)
1	20	6	65	26	120	553
2	40	11	339	51	440	3456
3	60	17	495	77	1020	10332
4	80	23	717	103	1840	21218
5	101	26	1425	127	2626	31817

In each of the test cases, our column-generation heuristic prescribes a larger total expected excess revenue but fewer new facilities than our approximation method. Except case 1 ($|P|=20$), the gap of # PCP years prescribed by our two methods is very little ($\leq \pm 2\%$) and the gap between total expected excess revenues is relatively small ($\leq 10.8\%$). Our approximation method requires an average of 80.1% less run time than our column-generation heuristic.

Our approximation method tends to prescribe several locations with respective population centers that are adjacent to each other, while our column-generation heuristic tends to prescribe a single location to serve nearby population centers. Our approximation method (column-generation heuristic) gives an average $|L'|/|L|$ value over all cases of 50.6% (27.1%). Our approximation

method prescribes clusters of nearby facilities when they have similar adjacency BSF trees and similar average incremental expected excess revenue per population center. It cannot differentiate between the locations, so it prescribes all of them. For example, in case 2, our approximation method prescribes L18, L19 and L20. On the other hand, our column-generation heuristic prescribes L12, which is surrounded by L18, L19 and L20. This phenomenon deteriorates the solution quality of our approximation method because it does not "see" the "big picture" that our column-generation heuristic does, especially for a small-scale problem like case 1.

As the number of population centers increases, our approximation method prescribes more scattered locations of new facilities and the $|L'|/|L|$ ratio monotonically decreases from 83.3% (in case 1) to 34.6% (in case 5). Because opening many facilities that are close to each other is not cost-effective for serving a larger area, clusters of locations are less attractive in larger scale instances. On the other hand, our column-generation heuristic gives an $|L'|/|L|$ ratio that increases monotonically form 18.2% (in case 2) to 34.6% (in case 5), with the exception of 33.3% in case 1. As the number of population centers increase, our column-generation heuristic would prescribe additional locations when the current set of selected locations cannot make the total expected excess revenue any larger by serving more population centers, because the increasing expected excess demand poses a larger penalty cost on the total expected excess revenue.

In our largest-scale case (i.e., case 5), both of our methods prescribe similar solutions, each with nine locations; the gap between the total expected excess revenues (#PCP years) prescribed by our two methods is only 3.18% (-1.32%). Figure 19 (20) presents the solution of our column-generation heuristic (approximation method) in case 5. Six (L5, L19, L22, L11, L17, L13) of nine locations are prescribed by both methods. The other three locations prescribed by our two methods can be paired to be counterparts because the pair of counterparts are located close to each other (i.e., located in adjacent zip codes). For example, L24 prescribed by our column-generation

heuristic is close to L1 prescribed by our approximation method. The other two pairs of counterparts prescribed by our column-generation heuristic (approximation method) are L18 (L23) and L26 (L9). Both of the two methods prescribe four new facilities in Williams and Travis Counties to serve these population-dense areas as well as rural areas in Burnet County. The rural area in Lampasas is assigned to L19 (in Bell County) by our column-generation heuristic and to L1 (in McLennan County) by our approximation method.

5.4.2 Bounding procedure

We specify the depth limit $\bar{q}_l$ as the smallest depth that contains over 30 population centers in BFS_l. Table 38 shows results of the bounding procedure used in our approximation method for case 5. The first column gives locations and is followed by five triples of columns, one for each depth. Each triple gives: N_q, the total number of population centers up to depth q; Gap1, $\frac{\hat{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)}$; and Gap2, $\frac{\breve{C}(\bar{\boldsymbol{y}}_l^q)-\bar{C}(\bar{\boldsymbol{y}}_l^q)}{\bar{C}(\bar{\boldsymbol{y}}_l^q)}$. The last column gives Gap, which is computed using (76). The maximal Gap in the last column defines gap_app (defined in (77)), which is 8.4%.

Because our approximation method runs much faster than our column-generation heuristic, we can use the gap derived for our approximation method (i.e., gap_app) to infer gap_cgh, the gap between expected excess revenues prescribed by our column-generation heuristic and the optimal solution. For example, in case 5, gap_cgh is computed as:

$$\text{gap\_cgh} = \frac{\bar{C}_{APP}(1+8.4\%)-\bar{C}_{CGH}}{\bar{C}_{APP}(1+8.4\%)} = \frac{\bar{C}_{APP}(1+8.4\%)-\bar{C}_{APP}(1+3.2\%)}{\bar{C}_{APP}(1+8.4\%)} = 1-\frac{1+3.2\%}{1+8.4\%} = 4.8\%,$$

where $\bar{C}_{APP}$ ($\bar{C}_{CGH}$) stands for the expected excess revenue associated with our approximation method (column-generation heuristic).

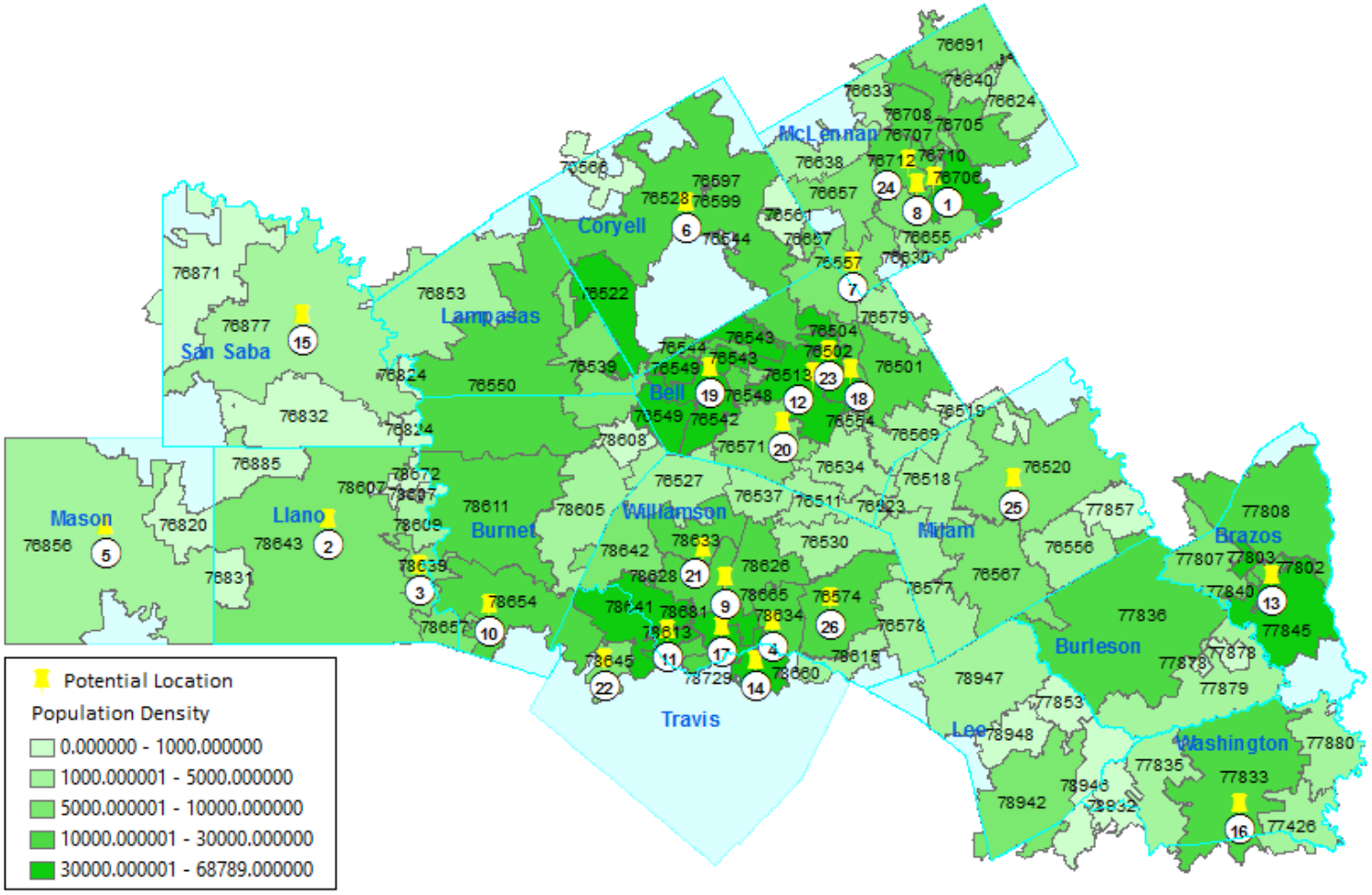

Figure 18. Mid-Texas map and potential locations

Table 36. Column-generation heuristic solutions (Mid-Texas)

Case (|*L'*|)	1 (2)		2 (2)		3 (4)		4 (6)		5 (9)	
	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)
L1	-	-	0	0	0	0	0	0	0	0
L2	-	-	-	-	-	-	0	0	0	0
L3	-	-	-	-	0	0	151	2.02E+08	0	0
L4	-	-	-	-	-	-	0	0	0	0
L5	-	-	-	-	70	9.57E+07	0	0	20	3.19E+07
L6	0	0	0	0	0	0	0	0	0	0
L7	336	5.85E+08	0	0	0	0	0	0	0	0
L8	-	-	110	1.74E+08	0	0	0	0	0	0
L9	-	-	-	-	0	0	0	0	0	0
L10	-	-	-	-	0	0	0	0	0	0
L11	-	-	-	-	-	-	0	0	190	3.11E+08
L12	0	0	650	1.07E+09	0	0	0	0	0	0
L13	-	-	-	-	-	-	-	-	380	6.70E+08
L14	-	-	-	-	-	-	-	-	0	0
L15	-	-	0	0	0	0	0	0	0	0
L16	-	-	-	-	-	-	-	-	0	0
L17	-	-	-	-	-	-	605	9.73E+08	380	5.99E+08
L18	0	0	0	0	0	0	0	0	157	2.26E+08
L19	-	-	0	0	0	0	219	3.41E+08	565	9.40E+08
L20	52	6.09E+07	0	0	0	0	0	0	0	0
L21	-	-	-	-	590	9.91E+08	70	1.01E+08	0	0
L22	-	-	-	-	-	-	0	0	222	3.48E+08
L23	0	0	0	0	0	0	0	0	0	0
L24	-	-	-	-	525	8.92E+08	110	1.69E+08	330	5.78E+08
L25	-	-	0	0	50	3.63E+07	665	1.13E+09	0	0
L26	-	-	-	-	-	-	0	0	103	1.33E+08
Total	388	6.45E+08	760	1.25E+09	1315	2.02E+09	1820	2.92E+09	2347	3.84E+09
Run time (sec)	65.574		389.391		680.66		1150.796		3226.691	

Table 37. Approximation method solution (mid-Texas)

Case (|*L'*|)	1 (5)		2 (6)		3 (7)		4 (9)		5 (9)	
	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)	# PCP yrs	c_bar ($)
L1	-	-	0	0	0	0	520	9.17E+08	520	9.17E+08
L2	-	-	-	-	-	-	0	0	0	0
L3	-	-	-	-	0	0	0	0	0	0
L4	-	-	-	-	-	-	0	0	0	0
L5	-	-	-	-	460	6.84E+08	18	1.50E+07	18	1.50E+07
L6	164	2.01E+08	383	6.16E+08	0	0	0	0	0	0
L7	40	2.21E+07	0	0	10	1.17E+07	0	0	0	0
L8	-	-	110	1.74E+08	65	5.68E+07	0	0	0	0
L9	-	-	-	-	0	0	0	0	102	1.56E+08
L10	-	-	-	-	0	0	0	0	0	0
L11	-	-	-	-	-	-	110	1.63E+08	190	3.12E+08
L12	90	1.40E+08	0	0	0	0	0	0	0	0
L13	-	-	-	-	-	-	-	-	401	6.72E+08
L14	-	-	-	-	-	-	-	-	0	0
L15	-	-	0	0	0	0	0	0	0	0
L16	-	-	-	-	-	-	-	-	0	0
L17	-	-	-	-	-	-	384	6.24E+08	465	7.08E+08
L18	50	3.84E+07	66	8.11E+07	0	0	0	0	0	0
L19	-	-	166	2.04E+08	0	0	328	5.00E+08	328	5.14E+08
L20	0	0	20	2.53E+07	336	5.64E+08	0	0	0	0
L21	-	-	-	-	0	0	60	6.34E+07	0	0
L22	-	-	-	-	-	-	197	3.12E+08	197	3.13E+08
L23	96	1.08E+08	0	0	131	1.76E+08	63	5.39E+07	95	1.08E+08
L24	-	-	-	-	218	3.04E+08	0	0	0	0
L25	-	-	20	2.21E+07	9	2.86E+06	120	1.66E+08	0	0
L26	-	-	-	-	-	-	0	0	0	0
Total	440	5.09E+08	765	1.12E+09	1229	1.80E+09	1800	2.82E+09	2316	3.71E+09
Run time (sec)	26.823		72.784		112.213		164.396		289.504	
# PCP years gap	13.4%		0.66%		-0.49%		-1.1%		-1.3%	
c_bar gap	21.1%		10.0%		10.8%		3.6%		3.2%	

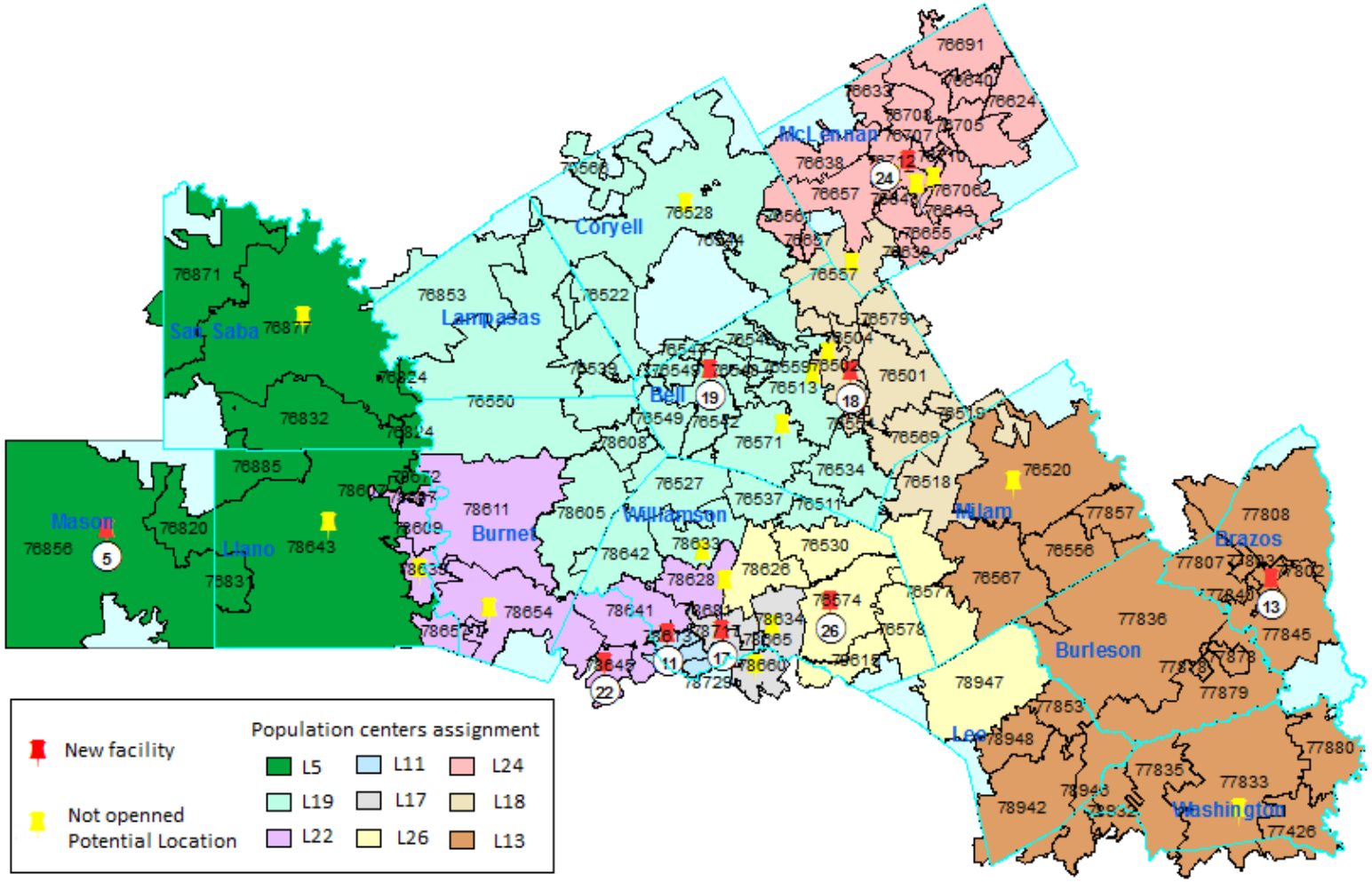

Figure 19. Column-generation heuristic solution for mid-Texas case study

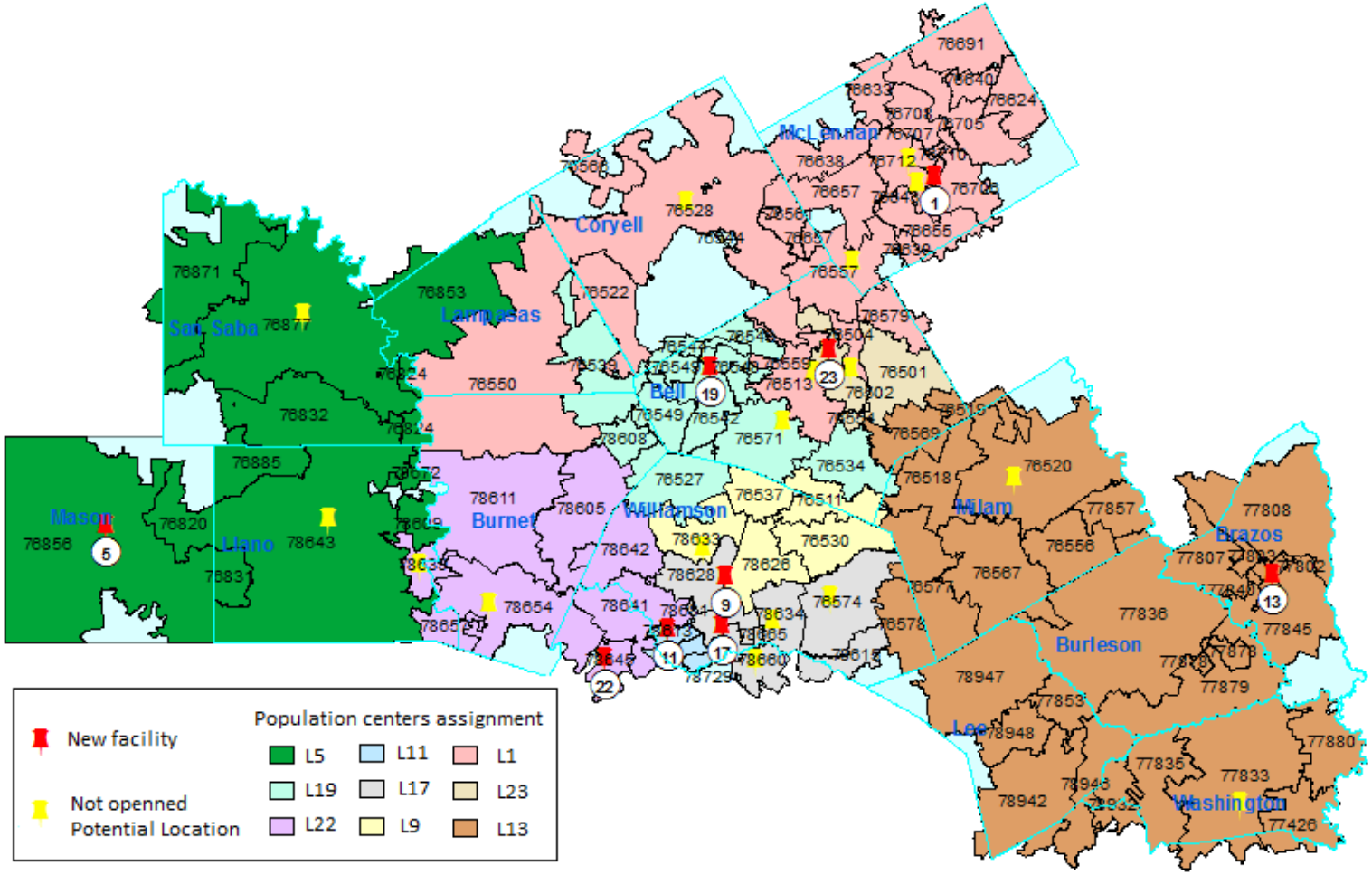

Figure 20. Approximation method solution for mid-Texas case study

Table 38. Bounding by our approximation method in case 5

q	**1**			**2**			**3**			**4**			**5**			
	N_q	Gap1	Gap2	N_q	Gap1	Gap2	N_q	Gap1	Gap2	N_q	Gap1	Gap2	N_q	Gap1	Gap2	Gap
L1	10	15.8%	0.6%	18	8.4%	0.9%	28	8.5%	0.2%	43	41.1%	6.4%	-	-	-	6.4%
L5	3	25.6%	1.9%	11	29.0%	2.1%	18	14.1%	1.0%	31	11.2%	0.9%	-	-	-	2.1%
L9	9	14.7%	0.7%	21	12.9%	1.5%	37	78.4%	7.5%	-	-	-	-	-	-	7.5%
L11	5	10.2%	0.6%	12	10.0%	0.7%	22	43.6%	6.6%	38	91.3%	6.6%	-	-	-	6.6%
L13	4	14.8%	0.6%	9	14.3%	0.3%	17	11.7%	0.4%	26	12.9%	1.1%	33	14.2%	0.5%	1.1%
L17	5	10.9%	0.6%	12	15.2%	1.8%	23	41.0%	6.2%	34	57.7%	8.4%	-	-	-	8.4%
L19	8	26.0%	1.0%	21	8.7%	0.3%	37	35.2%	5.3%	-	-	-	-	-	-	5.3%
L22	3	24.9%	1.8%	11	16.3%	0.9%	23	24.0%	3.5%	39	92.0%	6.4%	-	-	-	6.4%
L23	7	17.7%	0.9%	22	15.3%	0.7%	40	20.4%	2.8%	-	-	-	-	-	-	2.8%

CHAPTER VI

CONCLUSIONS AND FUTURE RESEARCH

This chapter comprises two sections. The first section draws conclusions on our DSCR research and presents further research opportunities. The second section draws conclusions on our research in healthcare facility location and capacity configuration under stochastic demand (i.e., SSHFCP and SMHFCP) and presents relevant further research.

6.1 DSCR: conclusions and future research

The first part of this dissertation contributes by formulating two models of the DSCR problem, one a traditional integer program and one based on a network structure. Tests promote intuitive interpretation of these models, assess their solvability, and identify the sensitivity of run time to primary parameters. This research also contributes because it is applicable to a number of DSCR problems in both of public and private sectors.

Our first model, DSCR-T, establishes a framework that has the flexibility to deal with many practical aspects of dynamic supply chain reconfiguration, providing a unique capability to reconfigure within a multi-period, multi-product, multi-echelon supply chain network to meet a time varying demand and/or cost structure. Thus, the model provides the adaptability needed in the competitive modern business environment. The model incorporates practical features that have not been taken into account in prior models, including budget constraints, single sourcing, inventory, backordering, outsourcing, and limits on the numbers of capacity expansions and contractions.

Our second model, DSRC-N, employs a convenient network structure to model the same set of practical features. It requires fewer binary variables, continuous variables, and constraints

and less run time than DSCR-T. Interestingly, our tests show that the linear relaxations of both models give the same optimal solution in each case but that DSCR-N offers superior solvability compared to DSCR-T.

Test examples of a product life cycle and an economic downturn with recovery lend insight into model interpretation, showing that the models prescribe the same solution in each case, one that results in a capacity utilization of 0.85, a typical industrial target. Tests show that model run time is highly sensitive to $|P|$, sensitive to $|T|$, mildly sensitive to $|L|$, and relatively insensitive to N_Tot.

Future research could formulate models with improved solvability and/or more effective solution methods. This should not be unexpected, however; (A Klose & Drexel, 2005)emphasized the challenges posed by production-distribution network design, a specialization of DSCR. In particular, scalability becomes an issue as problem size increases but DSCR-N lends itself well to branch-and-price decomposition, which would result in a relatively small, resource-constrained shortest path sub-problem for each product and DC location combination, a structure that promises improved computational capability. For cases in which uncertainties (e.g., demands and costs) are significant, future research could devise an effective stochastic optimization method. The proposed DSCR models deal with domestic financial issues; however, they could easily be extended to address the international business environment by incorporating corresponding financial issues (e.g., border crossing costs, transfer prices, tariffs, income tax rates, local contents rules). Our research continues along these lines.

6.2 SSHFCP and SMHFCP: conclusions and future research

Our model shows that SSHFCP can be cast as a deterministic RCSPP to locate a facility and configure its capacity. We model competition among facilities using an attraction model, which

estimates the demand each facility would attract as a function of healthcare factors (physician ratings; facility capacity; the shortest-path distance, the time to travel the fastest route, and a measure of accessibility from (the centroid of) each population center to each location) and parameters that reflect the emphasis that patients place on each factor. We deal with stochastic demand by employing recourse to maximize expected excess revenue, which includes closed-form penalties for expected excess demand and expected excess capacity. The network structure allows each arc to be labeled with the expected excess revenue it provides, facilitating solution by avoiding the need to optimize a stochastic program with a nonlinear objective function. The case study demonstrates application of the model in a realistic setting and our sensitivity analyses assess the relative impacts of important factors, key parameters, variability of physician capacity, and penalty cost values, revealing insights that can be used to guide healthcare decisions.

Our SSHFCP model can represent demand in terms of fundamental parameters for which data is available from commercial sources, but it is also sufficiently flexible to accommodate the data that is available publically. The Center for Disease Control and Prevention provides the number of visits/year/100 persons for primary care services by age, not by demographic group. We have not been able to find data that give the probability that a resident would consider changing her\his PCP or the distribution of visits/year for each person who seeks treatment. The ACA requires providers to maintain extensive data bases; studies like ours show the utility of gathering data needed to support the level of quantitative analysis that our prototype model shows to be possible.

Test results validate our model to large degree. After completing our tests, we learned that one local provider had already planned to build a new primary care center at location L3, which we found (independently) to be quite robust. Also, another local provider opened a new primary care center two years ago at a site that we used as a potential location, expecting our model

to prescribe it. However, our model did not prescribe that location under any of our test scenarios. We were concerned but later learned that the provider found that the location did not attract the new demand that some decision makers had expected. In addition, our healthcare collaborators agreed with test results, providing further validation based on their extensive practical experience.

Competition factors in attraction model (36) (i.e., Γ_{lst}, d_{lp}, k_{lst}, and f_{lp}) are objective measures that are readily available. The emphasis actually placed on them by patients, as represented by parameters (i.e., $\alpha, \beta, \gamma, \delta$), could be estimated by surveying patients, should values assigned intuitively and/or by sensitivity analysis prove unacceptable. Future research could evaluate yet additional factors and seek further empirical validation. In particular, it would be of interest to characterize how a new facility cannibalizes demand from existing facilities in the same firm. Cannibalization can be favorable if it provides better service and/or frees up space needed for other purposes but detrimental if it reduces economies of scale.

Healthcare insurance and government policy play major roles in patients' selections of physicians and facilities. The selection of an insurer is an involved decision that may be affected by insurance policies that are offered by one's employer, government policy, and coverage/deductibles. Our attraction model (36) could easily be extended with an additional dimension (i.e., subscript) for insurer so that it finds the probability that a patient carrying a type of insurance in a demographic group in a population center will seek service at a particular location. We were not able to find publically available data that relates insurance carriers to healthcare services. We expect that providers who apply our model can employ proprietary historical data and/or purchased information that would allow the full capabilities of our prototypical model to be used.

SMHFCP employs a location-allocation model for stochastic, multiple healthcare facility *location* problem and a network model to solve multi-period facility *capacity configuration*

problem under stochastic patient demand. We propose a novel column-generation heuristic that deals with columns associated with stochastic objective function value, using RCSPP as assignment subproblem associated with each facility. We show that, under certain specific conditions, the RMP in our column-generation heuristic can be solved to optimality. In addition, we propose an approximation method that utilizes the average incremental expected excess revenue to approximate the expected excess revenue for each pair of population-center-to-location assignment, as well as a bounding procedure to find the gap between the expected excess revenue prescribed by our approximation method and the optimal solution.

We perform a set of computational experiments to locate primary care centers in two realistic case studies: one involves 36 census tracts in Brazos County, Texas; the other, 101 zip codes in rural mid-Texas. Our column-generation heuristic prescribes somewhat larger expected excess revenues, while our approximation method requires less run time. The difference between the solution values prescribed by our two methods reduces as the number of population centers increases. We recommend using our column-generation heuristic to solve an instance and our approximation method in conjunction to assess the quality of the solution.

Future research can deal with problems of expanded scope, for example, prescribing multiple levels of healthcare services (i.e., primary, secondary and tertiary care) in a rural health care network. Current and future capacities of existing facilities could be treated as decision variables to better represent the leader-follower decisions of competitors as the new facility changes capacity. Finally, some patients may require multiple services (e.g., diabetes often leads to heart disease) and such relationships can be incorporated as well. Our research continues along these lines. Our SMHFCP approach can also be extended for application in other stochastic facility location and capacity configuration settings, such as schools, supply chains, retail stores and restaurants.

REFERENCES

Adenso-Diaz, B, & Rodriguez, F. (1997). A simple search heuristic for the MCLP: application to the location of ambulance bases in a rural region. *Omega, 25*(2), 181-187.

Afshari, H, & Peng, Q. (2014). Challenges and Solutions for Location of Healthcare Facilities. *Ind Eng Manage, 3*(127), 2169-0316.

Aghezzaf, E. (2005). Capacity planning and warehouse location in supply chains with uncertain demands. *Journal of the Operational Research Society, 56*(4), 453-462.

Albareda-Sambola, M., Fernández, E., & Saldanha-da-Gama, F. (2011). The facility location problem with Bernoulli demands. *Omega, 39*(3), 335-345.

Balakrishnan, J. (2004). An improved algorithm for solving a multi-period facility location problem. *IIE Transactions, 36*(1), 19-22.

Ballou, R. H. (1968). Dynamic warehouse location analysis. *Journal of Marketing Research*, 271-276.

Baron, O., & Milner, J. (2010). Facility Location: A robust optimization approach. *Production and Operations Management Society*, 1-14.

Behmardi, B., & Lee, S. (2008). *Dynamic multi-commodity capacitated facility location problem in supply chain.* Paper presented at the Proceedings of the 2008 Industrial Engineering Research Conference.

Benneyan, J. C., Musdal, H., Ceyhan, M. E., Shiner, B., & Watts, B. V. (2012). Specialty care single and multi-period location–allocation models within the Veterans Health Administration. *Socio-Economic Planning Sciences, 46*(2), 136-148.

Beraldi, P., Bruni, M. E., & Conforti, D. (2004). Designing robust emergency medical service via stochastic programming. *European Journal Of Operational Research, 158*(1), 183-193.

Berman, O., Ganz, Z., & Wagner, J.M. (1994). A stochastic optimization model for planning capacity expansion in a service industry under uncertain demand. *Naval Research Logistics, 41*(4), 545-564.

Bruni, M. E., Conforti, D., Sicilia, N., & Trotta, S. (2006). A new organ transplantation location–allocation policy: a case study of Italy. *Health Care Management Science, 9*(2), 125-142.

Canel, C., Khumawala, B. M., Law, J., & Loh, A. (2001). An algorithm for the capacitated multi-commodity multi-period facility location problem. *Computers and Operations Research, 28*, 411-427.

Canel, C., & Khumawala, B. M. (1997). Multi-period international facilities location: an algorithm and application. *International Journal of Production Research, 35*(7), 1891-1910.

Carlyle, W. M., Royset, J. O., & Wood, R. K. (2008). Lagrangian relaxation and enumeration for solving constrained shortest-path problems. *Networks, 52*(4), 256-270.

Casella, G., & Berger, RG. (2001). *Statistical Inference, 2nd edition.* Australia: Duxbury Press.

Chardaire, P., Sutter, A., & Costa, M. C. (1996). Solving the dynamic facility location problem. *Networks, 28*(2), 117-124.

Cheu, R. L., Huang, Y., & Huang, B. (2008). Allocating emergency service vehicles to serve critical transportation infrastructures. *Journal of Intelligent Transportation Systems, 12*(1), 38-49.

Cho, C. J. (1998). An equity-efficiency trade-off model for the optimum location of medical care facilities. *Socio-Economic Planning Sciences, 32*(2), 99-112.

Côté, M. J., Syam, S. S., Vogel, W. B., & Cowper, D. C. (2007). A mixed integer programming model to locate traumatic brain injury treatment units in the department of veterans affairs: a case study. *Health Care Management Science, 10*(3), 253-267.

Daskin, M. S., Coullard, C. R., & Shen, Z. (2002). An inventory-location model: Formulation, solution algorithm and computational results. *Annals of Operations Research, 110*(1-4), 83-106.

Daskin, M. S., Snyder, L. V., & Berger, R. T. (2005). Facility location in supply chain design. *Logistics Systems: Design and Optimization* (pp. 39-65), Springer, New York, NY.

Daskin, M. S., & Dean, L. K. (2004). *Location of Health Care Facilities in Handbook of OR/MS in Health Care: A Handbook of Methods and Applications*, Norwell, MA: Kluwer.

De Angelis, V., Felici, G., & Impelluso, P. (2003). Integrating simulation and optimisation in health care centre management. *European Journal of Operational Research, 150*(1), 101-114.

Dumitrescu, I., & Boland, N. (2003). Improved Preprocessing, Labeling and Scaling Algorithms for the Weight-Constrained Shortest Path Problem. *Networks, 42*(3), 135-153.

Fong, C. O., & Srinivasan, V. (1981a). The multi region dynamic capacity expansion problem—Part II. *Operations Research, 29*, 800-816.

Fong, C. O., & Srinivasan, V. (1981b). The multi region dynamic capacity expansion problem—Parts I. *Operations Research, 29*, 787–799.

Galvão, R., Espejo, L. G. A., Boffey, B., & Yates, D. (2006). Load balancing and capacity constraints in a hierarchical location model. *European Journal of Operational Research, 172*(2), 631-646.

Gendreau, M., Laporte, G., & and Semet, F. (2006). The maximal expected coverage relocation problem for emergency vehicles. *Journal of the Operational Research Society, 57*(1), 22-28.

Geoffrion, A. M., & Graves, G. W. (1974). Multicommodity distribution system design by Benders decomposition. *Management Science, 20*(5), 822-844.

Griffin, P. M., Scherrer, C. R, & Swann, J. L. (2008). Optimization of community health center locations and service offerings with statistical need estimation. *IIE Transactions, 40*(9), 880-892.

Gue, K. R. (2003). A dynamic distribution model for combat logistics. *Computers & Operations Research, 30*(3), 367-381.

Gunes, E. D., & Yaman, H. (2010). Health network mergers and hospital re-planning. *Journal of the Operational Research Society, 61*(2), 275-283.

Hagerty, J. R. (2012, May 17). 3M Begins Untangling Its 'Hairballs', *Wall Street Journal*, p. 2.

Hagerty, J. R. (2012, May 22). Once Made in China: Jobs Trickle Back to U.S. Plants, *Wall Street* Journal, p. 14.

Haghani, A. (1996). Capacitated maximum covering location models: formulations and solution procedures. *Journal Of Advanced Transportation, 30*(3), 101-136.

Handler, G. Y., & Zang, I. (1980). A dual algorithm for the constrained shortest path problem. *Networks, 10*(4), 293-309.

Harper, P. R., Shahani, A. K., Gallagher, J. E., & Bowie, C. (2005). Planning health services with explicit geographical considerations: a stochastic location-allocation approach. *Omega, 33*(2), 141-152.

Hinojosa, Y., Kalcsics, J., Nickel, S., Puerto, J., & Velten, S. (2008). Dynamic supply chain design with inventory. *Computers & Operations Research, 35*(2), 373-391.

Hinojosa, Y., Puerto, J., & Fernández, F. R. (2000). A multiperiod two-echelon multicommodity capacitated plant location problem. *European Journal of Operational Research, 123*(2), 271-291.

Huff, D. L. (1963). Probabilistic analysis of shopping center trade areas. *Land Economics, 39*, 81-90.

Jacobs, D., Silan, M., & Clemson, B. (1996). An analysis of alternative locations and service areas of American red cross blood facilities. *Interfaces, 26*(3), 40-50.

Jayaraman, V., & Srivastava, R. (1995). A service logistics model for simultaneous siting of facilities and multiple levels of equipment. *Computers & Operations Research*, 191-204.

Joksch, H. C. (1966). The shortest route problem with constraints. *Journal of Mathematical Analysis and Applications, 14*(2), 191-197.

Klose, A., & Drexl, A. (2005). Facility location models for distribution system design. *European Journal of Operational Research, 162*(1), 4-29.

Knight, V. A., Harper, P. R., & Smith, L. (2012). Ambulance allocation for maximal survival with heterogeneous outcome measures. *Omega, 40*(6), 918-926.

Knight, V. A., Williams, J. E., & Reynolds, I. (2012). Modelling patient choice in healthcare systems: development and application of a discrete event simulation with agent-based decision making. *Journal of Simulation, 6*(2), 92-102.

Lee, S. B., & Luss, H. (1987). Multifacility-type capacity expansion planning: algorithms and complexities. *Operations Research, 35*(2), 249-253.

Liu, N., Ziya, S., & Kulkarni, V. G. (2010). Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. *Manufacturing & Service Operations Management, 12*(2), 347-364.

Louveaux, F. (1993). Stochastic location analysis. *Location Science, 1*, 127-154.

Lowe, T. J, & Preckel, P. V. (2004). Decision technologies for agribusiness problems: A brief review of selected literature and a call for research. *Manufacturing & Service Operations Management, 6*(3), 201-208.

Lozano, L., & Medaglia, A. L. (2013). On an exact method for the constrained shortest path problem. *Computers & Operations Research, 40*(1), 378-384.

Mahar, S., Bretthauer, K. M., & Salzarulo, P. A. (2011). Locating specialized service capacity in a multi-hospital network. *European Journal of Operational Research, 212*(3), 596-605.

Malczewski, J., & Ogryczak, W. (1990). An interactive approach to the central facility location problem: locating pediatric hospitals in Warsaw. *Geographical Analysis, 22*(3), 244-258.

Marianov, V., Rios, M., & Icaza, M. J. (2008). Facility Location for Market Capture When Users Rank Facilities by Shorter Travel and Waiting Times. *European Journal of Operational Research, 191*, 32-44.

Mehrez, A., Sinuany-Stern, Z., Arad-Geva, T., & Binyamin, S. (1996). On the implementation of quantitative facility location models: the case of a hospital in a rural region. *Journal of the Operational Research Society, 47*(5), 612-625.

Melachrinoudis, E., & Min, H. (2000). The dynamic relocation and phase-out of a hybrid, two-echelon plant/warehousing facility: A multiple objective approach. *European Journal of Operational Research, 123*(1), 1-15.

Melo, M. T., Nickel, S., & Saldanha-da-Gama, F. (2009). Facility location and supply chain management–A review. *European Journal of Operational Research, 196*(2), 401-412.

Melo, M. T., Nickel, S., & Saldanha da Gama, F. (2006). Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research, 33*(1), 181-208.

Muthuraman, K., & Lawley, M. (2008). A stochastic overbooking model for outpatient clinical scheduling with no-shows. *IIE Transactions, 40*(9), 820-837.

Nakanishi, M., & Cooper, L. G. (1974). Parameter estimation for a multiplicative competitive interaction model: least squares approach. *Journal of Marketing Research, 11*(3), 303-311.

Ndiaye, M., & Alfares, H. (2008). Modeling healthcare facility locations for moving population groups. *Operations Research, 35*(7), 2154-2161.

Owen, S. H., & Daskin, M. S. (1998). Strategic facility location: a review. *European Journal of Operations Research, 111*, 423-447.

Price, W. L., & Turcotte, M. (1986). Locating a blood-bank. *Interfaces, 16*(5), 17-26.

Rais, A., & Viana, A. (2010). Operations research in healthcare: a survey. *International Transactions in Operational Research, 18*(1), 1-31.

Rivoli, P. (2009). *The Travels of a T-Shirt in the Global Economy: An Economist Examines the Markets, Power, and Politics of World Trade*, John Wiley & Sons.

Şahin, G., Süral, H., & Meral, S. (2007). Locational analysis for regionalization of Turkish Red Crescent blood services. *Computers & operations research, 34*(3), 692-704.

Santos, L., Coutinho-Rodrigues, J., & Current, J. R. (2007). An improved solution algorithm for the constrained shortest path problem. *Transportation Research Part B: Methodological, 41*(7), 756-771.

Santoso, T., Ahmed, S., Goetschalckx, M., & Shapiro, A. (2004). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research, 167*(1), 96-115.

Schweikhart, S., & Smith-Daniels, V. (1993). Location and service mix decisions for a managed health care network. *Socio-Economic Planning Sciences, 27*(4), 289-302.

Shen, Z. M., Coullard, C., & Daskin, M. S. (2003). A joint location-inventory model. *Transportation Science, 37*(1), 40-55.

Shulman, A. (1991). An algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. *Operations Research, 39*, 423-436.

Sinuany-Stern, Z., Mehrez, A., Tal, A. G., & Shemuel, B. (1995). The location of a hospital in a rural region: the case of the negev. *Location Science, 3*(4), 255-266.

Smith-Daniels, V., Schweikhart, S., & Smith-Daniels, D. (1988). Capacity management in health care services: review and future research directions. *Decision Sciences, 19*(4), 889-919.

Smith, H. K, Harper, P. R., Potts, C. N., & Thyle, A. (2009). Planning sustainable community health schemes in rural areas of developing countries. *European Journal of Operational Research, 193*(3), 768-777.

Snyder, L. V. (2007). Facility location under uncertainty. *IIE Transactions, 38*(7), 537-554.

Syam, S. S. (2008). A multiple server location-allocation model for service system design. *Computers & Operations Research, 35*(7), 2248-2265.

Syam, S. S., & Côté, M. J. (2010). A location-allocation model for service providers with application to not-for-profit health care organizations. *Omega, 38*(3), 157-166.

Tcha, D., & Lee, B. (1984). A branch-and-bound algorithm for the multi-level uncapacitated facility location problem. *European Journal of Operational Research, 18*, 35-43.

Thanh, P. N., Bostel, N., & Péton, O. (2008). A dynamic model for facility location in the design of complex supply chains. *International Journal of Production Economics, 113*(2), 678-693.

Torres Soto, J. E. (2009). *Dynamic and robust capacitated facility location in time varying demand environments.* Ph. D. Dissertation, Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX.

Van Roy, T. J., & Erlenkotter, D. (1982). A dual-based procedure for dynamic facility location. *Management Science, 28*(10), 1091-1105.

Verter, V., & LaPierre, S. D. (2002). Location of preventive health care facilities. *Annals of Operations Research*(110), 123-132.

Wang, Q., Batta, R., Bhadury, J., & Rump, C. (2003). Budget constrained location problem with opening and closing of facilities. *Computers & Operations Research, 30*(13), 2047-2069.

Wesolowsky, G. O., & Truscott, W. G. (1975). The multiperiod location-allocation problem with relocation of facilities. *Management Science, 22*(1), 57-65.

Wilhelm, W. E. (2001) A technical review of column generation in integer programming, *Optimization and Engineering*, 2 (2), 159-200.

Wilhelm, W. E., Han, X, & Lee, C. (2013). Computational comparison of two formulations for dynamic supply chain reconfiguration with capacity expansion and contraction. *Computers & Operations Research, 40*(10), 2340–2356.

Wilhelm, W. E., Liang, D., Rao, B., Warrier, D., Zhu, X., & Bulusu, S. (2005). Design of international assembly systems and their supply chains under NAFTA. *Transportation Research Part E: Logistics and Transportation Review, 41*(6), 467-493.

Zhang, Y., Berman, O., Marcotte, P., & Verter, V. (2010). A bilevel model for preventive healthcare facility network design with congestion. *IIE Transactions, 42*(12), 865-880.

Zhang, Y., Berman, O., & Verter, V. (2012). The Impact of Client Choice on Preventive Healthcare Facility Network Design. *OR Spectrum, 34*(2), 349-370.

Zhang, Y., Berman, O., & Verter, V. (2009). Incorporating congestion in preventive healthcare facility network design. *European Journal of Operational Research, 198*(3), 922-935.

Zhu, X., & Wilhelm, W. E. (2012). A three-stage approach for the resource-constrained shortest path as a sub-problem in column generation. *Computers & Operations Research, 39*(2), 164-178.