Siting Primary Care Clinics To Meet Daytime And After-hours Objectives

by

John Charles Campbell

Submitted in partial fulfilment of the requirements for the degree of
Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
June 2021

Table of Contents

List of Tables

List of Figures

Abstract

Location science is used to determine the optimal geographical placement of primary care resources with operations research models. In determining the optimal placement, we account for the objectives of both patients and physicians. Patients prefer to be close to clinics to ensure access and physicians typically prefer to have minimum panel sizes to ensure consistent appointments. These objectives and the methods used to address them differ between daytime and after-hours settings. Three approaches are considered to address both time settings: independent, sequential, and simultaneous. The independent approach is based on the p-Median problem, and the other two approaches use modified forms of the p-Median. The models are generalized and applied to census data from Nova Scotia. Three case studies are examined using Canadian census data from Halifax, Cape Breton, and modified data in Cape Breton. The regular-to-after-hours approach is found to most frequently be the worst, while the simultaneous approach yields the best results while considering facility-sharing constraints.

Acknowledgments

I would like to thank my supervisors, Dr. Peter VanBerkel and Dr. Majid Taghavi, for their guidance, encouragement, and patience in helping me develop this work. Their wisdom is invaluable, and I will forever appreciate the generosity with which they committed time and effort to mentoring me. Despite the disruption to in-person collaboration I never felt distant when we met, and I greatly valued the collegiality and friendliness of our discussions.

I would also like to thank my committee members, Dr. Ronald Pelot and Dr. Melanie Reuter-Oppermann, for their care in reviewing my thesis. I had the pleasure of their supervision as a teaching assistant and research assistant during my studies, and I am thoroughly grateful for their tutelage and advice in these roles.

I thank the sponsors of the Bruce and Dorothy Rossetti Engineering Research Scholarship, as well as Mitacs for the Globalink Research Award. These awards allowed me to focus on my studies and collaborate on international projects. I am also thankful once again to my supervisors for funding my research.

Finally, I would like to thank my friends, family, colleagues, and fellow graduate students for their inspiration and support. In particular, the *Stochastic Study Squad* (or *the unINFORMedS*) deserve commendation for their camaraderie and valour in the thick of academic doldrums during the apocalypse. Godspeed, you magnificent bunch.

Jack

Chapter 1

Introduction

In most countries, primary healthcare is provided by general practitioners (GPs). Primary healthcare is the first step of treatment for most non-emergency health issues, and it is important for patients to be able to easily access primary health services. Access to health services can be measured in many ways (e.g., geographically, financially, or timeliness), but when planning health resource locations, distance between patients and services is a typical measure of access [2]. Normally patients have greater difficulty accessing health services as distance increases, thus it is valuable to patients to reduce distance to services. Distance-based access can be optimized by modelling different objectives, such as minimising the average distance between patients and services, minimising the maximum distance travelled by any patient, or maximising the percentage of patients within an acceptable distance from the nearest service. Optimised results from models can be used to compare to real-world systems or policies, or to set goals for system improvement.

It is possible to predict locations where GPs or primary care practices (PCPs) will most effectively improve patient access based on these measures. In regions where there is a shortage of GPs, this represents an opportunity to determine where recruited physicians will be most effective at meeting need for services. It is also a common problem in many countries that certain types of areas such as rural or urban hot spots are underserved per capita by primary care services [3]. In these cases, it is also valuable to determine optimal placement of existing or additional GPs to promote balanced access.

While meeting demands for patient service is crucial, it is also important to consider physician preferences for practice location, hours of service, and style of practice and the trade-offs with patient preferences for access [4]. For example, physicians typically prefer a large panel size to ensure consistent work, but not so large as to be overwhelming; however, patients may prefer to be in smaller panels so their GPs have more frequent appointments. GPs may also prefer a collegial environment, which is obtained by working in PCPs with other physicians rather than independently. This practice style has the potential to improve access to patients through shared delivery of services or coverage of after-hours clinics.

In many areas, primary care services are provided differently during daytime and nighttime hours. Typically, when night services are offered by PCPs, physicians will pool panels at night to allow them breaks from working while covering for each other [5, 6]. While this means a greater panel size, fewer physicians, and greater distances for patients, this is typically acceptable since the expectation of services during night is lower and fewer patients use them. While there are optimization models for primary care service locations, none are known which account for dual services such as the day and night primary care system. Usually, only day services are considered, or day and night hours are considered separately. By accounting for both service periods within the same model, trade-offs between patient and provider preferences can be considered and improvements to patient and physician satisfaction can be obtained.

The overarching research question to be informed by this thesis is "How do we optimally assign GP location so as to maximally satisfy patient and physician needs during both day and night services?". Specifically, this thesis will inform a study of primary care location considerations in Nova Scotia, Canada. Nova Scotia faces a shortage of family physicians and needs to recruit more (512 between 2016 and 2025 according to provincial plan) [7, 8]. It is known that some areas are more poorly served by physicians (such as rural areas, urban hot spots) [3]. Location of health services such as PCPs has an impact on patients, requiring travel time and expense and potentially limiting access to services, and it is possible to improve patient access by planning practice locations. By planning practice location in an optimized fashion, gains can be made to patient access and related health service outcomes.

Chapter 2

Literature Review

A key reason for reviewing the literature is to inform the methodological approach to this research, to compare the appropriateness of different approaches, and to reduce selection bias introduced from favouring certain approaches.

Metrics of importance to Nova Scotians will be considered. These may include access, average/maximum travel time, coverage, panel size, equity, and population participation in health services [9]. The literature review will support the identification of measurements appropriate for the selected metrics.

The review will be organized in this chapter as follows:

- Search Strategy
- Paper Selection
- Review of highlighted operations research papers
- Review of relevant healthcare literature
- Synthesis of review findings

2.1 Search Strategy

Relevant studies were identified by searching the following online databases: Google Scholar, Web of Science, and PubMed. Search terms included "primary health care", "primary care", "access to health care", "primary care location planning", and "matching patient physician preferences". Forward and backward searches from identified papers were also conducted. Relevant grey literature, including reports or other documents have been included. A jurisdictional scan was performed by appending "Nova Scotia" to the search terms, as well as searching through provincial health websites. Studies published in languages other than English were not included. The timeframe for the search began March 1994, as this corresponds to the publication of a paper that has been used to broadly define access [10]. Titles and abstracts of searched papers were reviewed for relevance. Selected full-text papers were retrieved and reviewed for inclusion.

2.2 Selected Papers

The search yielded approximately 850 titles/abstracts. Of these, 30 were selected for full-text review and, of these, 14 were determined to be directly relevant to the purpose of the literature review. The search strategy is shown in the PRISMA flow diagram in Figure 2.1. This diagram is adapted from the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) method of systematic review [11], and displays the process by which papers were selected for review. A summary of included studies by classification is presented in Table 2.1, while the detailed classifications of the included studies are shown in Table A.1. Studies were classified as belonging to daytime, after-hours, location science, operations research (OR), and healthcare settings. The classifications refer to the relevancy of the papers to the topic. For example, papers classified as daytime are relevant to normal operations (typically during the day), while after-hours papers refer to papers examining another time setting. Location science papers were related to distance measurement, geography, or mathematical concepts that could relate to spatial access. Operations research papers contained OR modelling, and healthcare papers specifically had a component relating to healthcare services or planning. The full classification of each paper can be found in Table A.1.

Table 2.1: Brief Summary of Studies

Study Type/Field	**Location Science**	**Operational Research**	**Healthcare**
Daytime	9	9	8
After-hours	1	1	2
Both	1	1	1

From Table 2.1, it appears that few of the selected articles pertain to the after-hours setting. This suggests a paucity of information in this area. This is not a

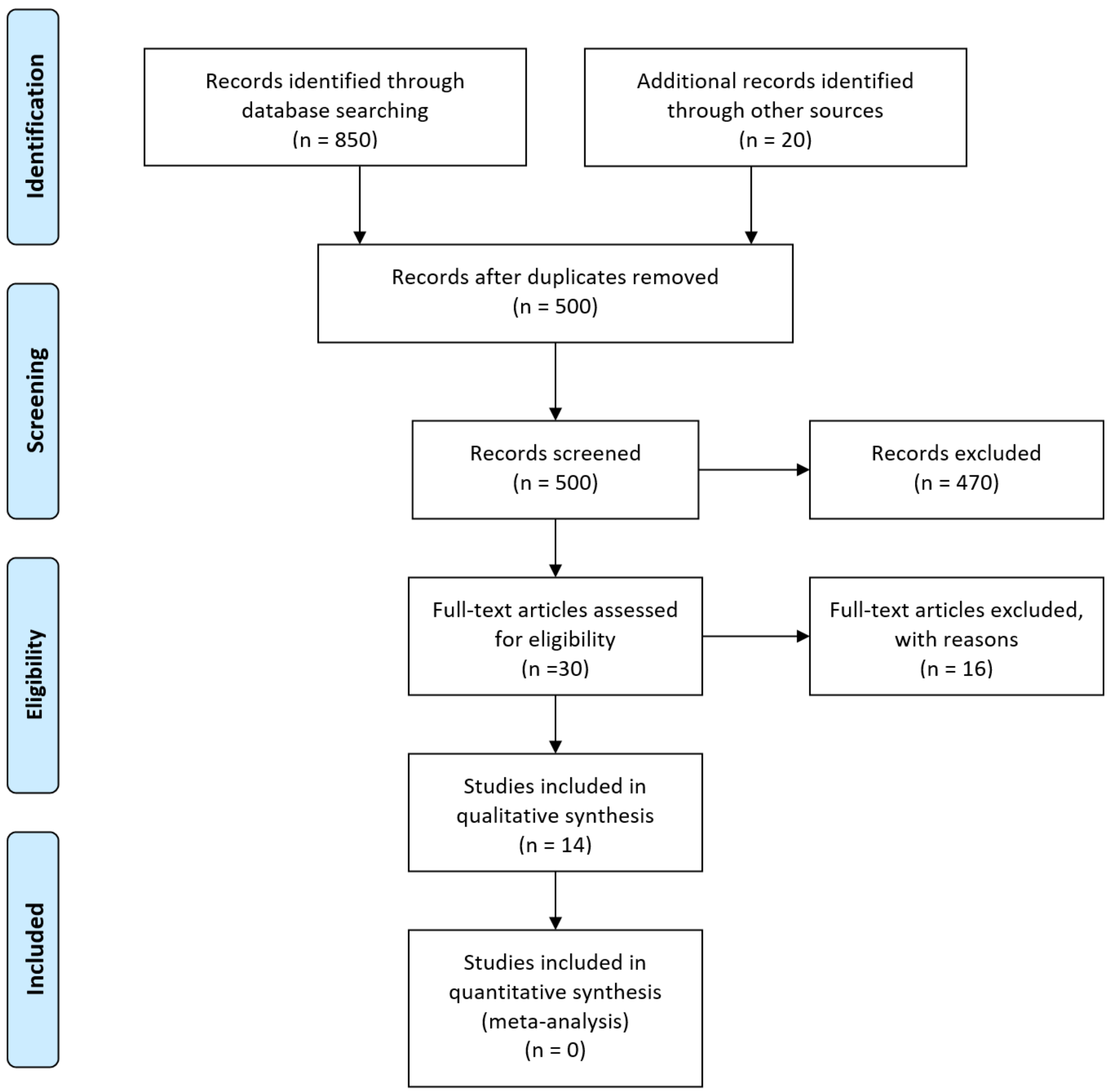


Figure 2.1: PRISMA Flow Diagram [1]

surprising result, as it was difficult to find articles on this topic specifically; the only selected article to address after-hours care from a location science or an operations research perspective is a working paper which prompted this literature review [12].

2.3 Highlighted Operations Research Papers

The following papers provide key insights into location planning of primary care practices using OR models and are discussed in greater detail.

Operations research meets need related planning: Approaches for locating general practitioner's practices

A similar problem has been tackled in 2019 using a model based on the maximum covering location problem (MCLP) to determine efficient GP placements in 21 municipalities within the state of Baden-Württemberg, Germany [13]. The study used a model intended to locate primary facilities according to three different objectives:

- Minimise travel time for all patients
- Maximise demand covered
- Minimise the maximum travel time for patients

The initial model was as follows:

$$\max \sum_{i \in I} e_i y_i$$

Subject to:

$$y_i \leq \sum_{j \in N_i} x_j \qquad \forall i \in I$$

$$\sum_{j \in J} x_j = n$$

$$x_j \in \{0, 1\} \qquad \forall j \in J$$

$$x_i \in \{0, 1\} \qquad \forall i \in I$$

where e_i is the weight of district i (which in this case represents the demand measured as population), x_j is a binary decision variable indicating whether a facility is located at j, and y_i is a binary decision variable representing whether demand at location i is covered. The objective function maximises demand coverage, and the constraints ensure that demand can only be covered if a facility is suitably placed to satisfy it, and that n facilities must be placed. This model locates n practices and determines the number of regions covered by a maximum driving time (region coverage indicates that all inhabitants can reach a practice within the maximum time). Possible facility locations that may cover the demand at node i are included in the set $N_i = \{j \mid r_{ij} \leq r_{max}\}$. To determine the closest new practice for each municipality an additional decision variable, z_{ij}, added. z_{ij} is a binary variable that determines if the demand at location i is assigned to facility location j.

Modifications were made to this model to add capacity (by adding decision variables for doctors per practice and patients per doctor). This second model assumes that patients spread out among GPs placed within their district ("patients that choose a different GP more or less equal out over the district"), assumes that all patients must be covered, and weights the distances traveled by patients. The model also assumed a complete replacement of any existing practices.

A third version of the model includes existing practices and modeled changing numbers of inhabitants until 2023 (in the years 2017, 2020, and 2023), and determines where new practices should be located, if existing practices should be kept open, and how many GPs are needed.

The first model found results for minimising average travel time, minimising maximum travel time, and maximising coverage for varying numbers of health center locations. The second model predicted the number of GPs needed for each location, as well as the average panel size and patient travel times. The third model determines the number of GPs that should be gained in 2017, 2020, 2023 and where to place them. Additionally, another version of the model limited new GPs to the most attractive regions, and determined where to best place them within these restrictions.

The study concluded that more information may be needed to predict patient behaviour for choosing GPs, and how patients travel to their GP (as well as how to incorporate transportation preferences).

Matching patient and physician preferences in designing a primary care facility network

This 2012 study designed a model for locating family health centres (FHCs) by considering both patient and physician needs from a planning perspective, then implemented the model with data from Sakarya, a small Turkish province [4]. The objective was to maximise geographical access of the population to primary health care, measured by percentage of population within a maximum travel distance of a facility and average distance to a facility. The physician preferences considered were income and workload, equity, and professional support and collegial work. The patient preferences considered were average distance to facilities, practice size (where smaller is preferred), nearest assigned facility, and equity.

An integer programming model is used. Parameters of this model were adjusted to accommodate different scenarios based on compromises in patient and GP preference metrics. The initial model used parameters chosen by city health officers in Sakarya and focused on coverage and participation objectives (referred to as base-c and base-p). Participation is measured through likelihood of a patient visiting a family physician, with decreasing probability as distance increases (up to a maximum distance). The base-c model found that maximum coverage was close to 97%, with decreasing marginal benefit as the number of facilities increased. Increasing physicians per FHC increased average travel distance and reduced patient participation, and physician equity was more easily maintained when the number of physicians increased with the number of facilities. The results also suggested that optimal solutions typically used the same locations, indicating that opening facilities in these locations and adding new ones as necessary may be a practical approach. The base-p model found that coverage decreased by around 3% from the base-coverage model, but expected participation increased by over 9%. It also found that average travel distance was always lower than the base-c results, almost 15% on average.

The base model was then modified (physician-c/p) to favour physician preferences by setting the minimum physicians per facility to 2 and increasing the lower bound on panel size from 2000 to both 2500 and 3000. The physician-p model was found to have the best results for coverage while satisfying physician preferences, and results of both models indicated that decreases in accessibility were not substantial compared to the base models.

To emphasize patient concerns, the base model was altered (patient-c/p) by reducing maximum travel distance from 3 km to 2 km and adding a constraint to allocate patients to the nearest facility. This model found better average patient travel distances, and a much higher participation rate in the patient-p scenario indicates that expected participation is a better measure of patient satisfaction. The model also found that increasing the number of facilities was always beneficial, while raising the number of physicians may be detrimental.

Finally, a model incorporating both patient and physician preferences (all-c/p) was implemented by setting the minimum panel size to 3000, minimum physicians per facility to 2, maximum travel distance to 2 km, and adding the nearest facility allocation constraint. The base-c model was found to be dominated by all other scenarios, suggesting the current system could be improved. The best performing scenarios were found to be base-p, physician-p, all-c, and all-p.

The article concludes that it is important to optimally serve the population while maintaining physician satisfaction, but there are trade-offs between patient access and physician satisfaction when resources are constrained. Increasing the number of facilities improved patient coverage, but sometimes also increased the distance travelled by some patients. Having more physicians increased the difficulty of maintaining a minimum panel size and decreased access measures. Different directions can be taken with this study, such as incorporating the effects of distance on visit frequency or investigating more specific preferences of patients and physicians. The authors suggest that investigation of the observations made in the study using alternative participation measures may yield different results.

2.4 Healthcare Literature Scan

The selected healthcare studies are summarised for important findings on the state of primary care as it pertains to operations research and access.

2.4.1 Healthcare Literature Descriptions

The survey by Ahmadi-Javid et al. [14] identified multi-period models as a key area for future research for locating primary care facilities. A multi-period model could be used similarly to a model that incorporates regular and after-hours care. Additionally, models that centralize location decisions to share resources were identified as an important area. These types of decisions could be used to determine how to allocate after-hours services among primary care facilities.

According to Crighton et al. [6], 62% of Canadian family physicians reported providing some level of after-hours care. This varied significantly by province, with Nova Scotia at 77.4%, Quebec at a low of 34.3%, and Alberta at a high of 88.4%. Provision of after-hours service was not affected by rural/urban setting, physician satisfaction, or whether physicians operated in group or solo practices.

A model by Graber-Naidich et al [15] is constructed based on regulator needs to optimize a multi-objective network of different types of primary care facilities. A location-allocation model is used, where facilities are initially treated as uncapacitated then assigned resources depending on quantity and type-mix allocated to each facility. The model weights three objectives, overall travel distance, operating costs of facilities, and non-appropriate-service (NAS). NAS refers to the quality and suitability of the services provided to patients.

A case study of the model in Kingston, Ontario examines model results when trading off cost and NAS. The model is shown to be capable of evaluating the effects of different policies, and can be used for planning for future scenarios.

A study by Guagliardo et al. [2] found that the majority of knowledge about the population health impact of geographic distribution of services is focused on hospitals, specialty services, and rural health services. Relatively little is known about the impact of the location of health services in urban settings.

Guagliardo et al. [2] also recognized spatial distance to healthcare providers as a barrier to healthcare access, and the probability that health services are utilized decreases with distance between patient and provider. The exact impacts of spatial accessibility on population health have not been measured but quantifying them would allow for development of policies to improve public health.

Güneş et al. [4] developed a multicriteria optimisation model that is used to maximise two objectives, coverage of patients within a maximum distance from a care provider and patient use of health services. This model accommodates patient and physician preferences, and demonstrates that this multicriteria model can improve current policies.

Güneş & Nickel [9] reported that facility location problems (FLPs) are operations research models where facilities are placed so as to optimise an objective. Typical objectives in healthcare FLPs are minimal access cost (typically measured in time or distance), maximal population within a specified distance of health services, and maximal equity of access to health services. The latter objective is typically difficult to define or measure.

The P-median problem is commonly used when deciding how to locate healthcare facilities, and it minimises the total cost of patient access to services at these facilities (often by minimising the total distance between patient and clinic).

According to Morgan & Graber-Naidich [3], health professionals are disproportionately concentrated in urban areas in many countries (including Canada), leading to insufficient health resources. Rural Canadians are older, poorer, sicker, less educated, and less healthy compared to urban Canadians, resulting in a greater per-capita demand for care.

There are fewer doctors per patient in rural areas than urban areas in Canada, causing a care gap between patients in these areas. When the care gap increases, the attractiveness of rural settings to physicians decreases due to the increased workload for each rural doctor. This creates a positive reinforcement loop, where rural settings will tend to have fewer doctors per patient over time.

Government efforts so far have slightly improved the problem, but not fixed it. Educational policies focused on incentivizing doctors to complete their training in deficient areas are expected to have the greatest potential for improving care gaps.

According to O'Malley et al. [5], the provision of after-hours care is important for patient outcomes, primary care physician (PCP) stress, and effective utilization of health resources. Available after-hours care was positively associated with safe and timely triage, patient accessibility and satisfaction, quality of care, PCP burnout, and financial stability.

Five models of after-hours provision for PCPs were considered with increasing integration into a network of after-hours care, ranging from solo operation (same PCP during regular and after-hours services) to dedicated referral of after-hours patients to a third party care centre. Three common factors found to be important to the success of after-hours care models were sustainable designs that met local needs, shared patient health information between after-hours and daytime care providers, and the implementation of all-hours access as part of a broader initiative to improve access and continuity of care.

In a paper by Reuter-Oppermann et al. [13], three models were used to optimally determine potential future locations for GP practices for three objectives: minimal driving time for all patients, maximal patient demand coverage, and minimal distance between the furthest patients and their GP. These match the three objectives for FLPs identified by Güneş and Nickel [9]. The maximum coverage location problem (MCLP) is used to determine GP practice locations, and modified to also determine the number of GPs needed at each practice and to determine locations for multiple time periods.

It was assumed that patients always choose the nearest GP, but more research is needed to know how patients choose their GP and travel to their GP. This is particularly important in urban settings, where multiple modes of travel are common.

In Reuter-Oppermann et al (2017) [16], an operations research model and a geographic information system (GIS) are used together to create a decision support system for optimally locating GP practices. The integration of these two systems can allow for great flexibility in the approach and presentation of optimal location strategies.

Two approaches are proposed by Reuter-Oppermann et al. [12] to use operations research models to create daytime and nighttime primary care networks. The methods optimize for a daytime setting then sequentially create a nighttime strategy and vice-versa. In both methods, it is assumed that physicians have larger panel sizes and patients may travel greater distances for service at night. Both methods satisfy patient requirements as constraints of the model before addressing physician preferences.

2.4.2 Key Healthcare Findings

The findings from this review reveal many important qualities of the state of primary care, after-hours services, and applications of operations research and location science to health services planning in primary care settings. Of particular interest to the project that prompted this review, access to primary care services is a critical component of population health and is directly connected to the geographical barriers between patients and providers [2]. This concept can be referred to as spatial accessibility, and typical objectives that can be targeted to improve spatial accessibility are the costs of access (often measured in distance between a patient and provider), the proportion of patients that have coverage (within a maximum distance or travel time from a healthcare provider), and equity of access among patients with comparable need for care [9]. Since these objectives are all important when evaluating strategies to improve spatial accessibility, multicriteria models that can accommodate all of them are popular. Multicriteria models also allow for great flexibility in addressing the sometimes-conflicting interests of the multiple groups of stakeholders that are inevitably involved in healthcare planning. While further study is needed to quantify the effects that spatial accessibility has on population health [2], it is clear that operations research and location science methods are well suited to tackling health services planning with the goal of improving spatial accessibility.

Several studies have implemented different operations research models with the goal of reducing geographic barriers between patients and healthcare. Many trends have been identified in these studies, such as the worsening care gap between rural and urban primary care, the lack of information in this area, and the need to trade off stakeholder preferences when locating healthcare facilities [3, 15, 4]. A common result in these studies is that OR models are able to produce primary care networks with optimal geographic placement of healthcare facilities [9, 14, 15, 4, 13]. Additionally, location science has been demonstrated to allow for the integration of geospatial information systems and flexibility of model inputs and presentation [16]. In some cases, the authors also use case studies of their models applied to real-world healthcare networks to demonstrate that the models can suggest substantial improvements to spatial accessibility and support policy changes that may be effective at implementing these improvements in real healthcare systems [3, 15, 4, 13, 16].

Important gaps in the literature reveal that after-hours services are, to date, an understudied area. The provision of after-hours primary care is associated with better patient outcomes and lower costs of care [5]. While most Canadian family physicians reported providing some level of after-hours care, this was highly dependent on region, ranging between 34.3% in Quebec to 88.4% in Alberta [6]. The authors noted that communities tended to exhibit the same traits in their provision of after-hours care, and that this may be due to a "herd effect" of physicians practicing in the same patterns as their colleagues. Despite the importance of after-hours primary care and the prevalence among Canadian physicians, no studies were found that incorporated after-hours care into a healthcare facility location problem other than the working paper associated with this review [12].

2.5 Review Conclusions

Accessible primary care is an essential component of maintaining a healthy population. After-hours services are important for providing equitable access to care and meeting patients' needs, however, they represent an understudied part of primary care and are typically not considered in healthcare facility location planning models. An OR model that optimally locates GP practices in both day and night settings is an unresearched area that will provide important information for health services planning.

It may be appropriate to substitute other types of "every day" facilities (e.g. mechanic shops, grocery stores, etc.) for primary care centres in terms of individual behaviour when comparing service distribution models. This may allow for models that have been used in other applications to be adapted for healthcare facility location, expediting the model development process. Further review to determine if these sorts of models have been compared to health services models is recommended.

Another recommended area for further research is whether patient behaviour has a significant impact on these models, and how to incorporate them if so. It is frequently assumed that patients travel to the nearest primary care centre, but some of the studies in this review have suggested models of patient behaviour that allow for differences in GP choice, mode of transportation, or even stochastic demand may be appropriate [9, 15, 2, 13].

With these considerations, location planning models have much to offer in terms of planning primary health care services to address a pressing health system problem, namely, improving access to primary health care in Nova Scotia. The remainder of this thesis will examine an application of this type of model to locating day and after-hours service in primary care.

Chapter 3

Methods

This chapter describes the methodology used to create the models in this thesis. To address the issue of optimally locating health resources, operations research provides robust mathematical models which consider different objectives and constraints. Broadly, the objective is to ensure that patients are as close to primary care as possible and the main constraint is the number of healthcare facilities available. A model which reflects these characteristics is the p-median problem.

3.1 p-Median Problem

The p-Median Problem is an operations research model which locates p facilities among a set of specified locations I so as to minimize the average distance between demand and the nearest facility [17]. The elements of the p-median problem are described in Table 3.1, followed by the objective function and constraints in (3.1) - (3.6). The purpose of these equations are described in Table 3.2.

Table 3.1: p-Median Notation

Sets:	
I	Set of all nodes where facilities may be placed, indexed by i.
J	Set of all nodes where demand is located, indexed by j.
Parameters:	
p	The number of facilities to locate.
d_j	The demand at node j.
c_{ij}	The unit cost of supplying demand at node j from a facility at node i.
Decision Variables:	
x_{ij}	The fraction of demand at node j satisfied by facility i.
y_i	A binary variable, 1 if a facility is located at site i and 0 otherwise.

Objective function:

$$\min \sum_{i \in I} \sum_{j \in J} d_j c_{ij} x_{ij} \tag{3.1}$$

Subject to:

$$\sum_{i \in I} x_{ij} = 1 \qquad \forall j \in J \tag{3.2}$$

$$\sum_{i \in I} y_i = p \tag{3.3}$$

$$x_{ij} - y_i \leq 0 \qquad \forall i \in I; j \in J \tag{3.4}$$

$$y_i \in \{0, 1\} \qquad \forall i \in I \tag{3.5}$$

$$x_{ij} \geq 0 \qquad \forall i \in I; j \in J \tag{3.6}$$

Table 3.2: p-Median Problem Definitions

Objective Function:

3.1 The objective function minimizes the sum of demand at each node d_j, multiplied by the cost of supplying that demand from node i (c_{ij}), multiplied by the proportion of demand allocated from j to i (x_{ij}) for all nodes i and j. Overall, this minimizes the overall cost of supplying demand with the given facilities.

Constraints:

3.2 Ensure that the demand at each node j is satisfied by located facilities.
3.3 Ensure that exactly p facilities are located.
3.4 Only allow demand to be allocated to nodes with facilities.
3.5 Locations must be binary, either 0 (closed) or 1 (open).
3.6 Demand allocation must be non-negative.

3.2 Multiple Time Setting p-Median Problem

The p-Median problem is typically used in a one-time setting. However, the problem under consideration covers multiple related time settings, regular and after-hours primary care. Regular care is typically conducted during the daytime during normal working hours. After-hours care is delivered outside of regular hours, usually at night and very early or late in the day. These time settings are mutually exclusive and primary care facilities may operate with different properties in each time setting. Thus, the p-Median problem is altered to address multiple time settings in one OR model. This allows decisions to be made while considering the costs in both time settings simultaneously. The notation used for this model is detailed in Table 3.3.

In this modification of the p-Median problem, a primary and secondary set of facilities are simultaneously located to minimize distance to demand in a primary and secondary time setting. In this case, the primary facilities are located in the regular time setting, and the secondary facilities are located in the after-hours time setting. A constraint added to the standard p-Median problem to create the modified problem is that the number of secondary facilities must be less than or equal to the number of primary facilities. This constraint is founded on the assumption that there are no more primary care practices open during the after-hours time setting than during the regular time setting.

This limitation also requires a second constraint. Secondary facilities must share locations with primary facilities. This satisfies the assumption that the same facilities are used for regular and after-hours care, which is expected to be more cost effective (and typical of real-world settings) than opening facilities separately for each setting. Together these constraints ensure that after-hours facilities must share locations with regular facilities.

Table 3.3: Multiple Time Setting p-Median Notation

Sets:

I	Set of all nodes where facilities may be placed, indexed by i.
J	Set of all nodes where demand is located, indexed by j.

Parameters:

p	The number of primary facilities to locate.
p'	The number of secondary facilities to locate.
d_j	The primary demand at node j.
d'_j	The secondary demand at node j.
c_{ij}	The unit cost of supplying demand at node j from a facility at node i.
W	The discount rate applied to the secondary setting. This can be thought of as relative cost; for example $W = 0.5$ suggests that cost incurred in the secondary setting is considered half as important as cost in the primary setting.

Decision Variables:

x_{ij}	The fraction of primary demand at node j satisfied by facility i.
x'_{ij}	The fraction of secondary demand at node j satisfied by facility i.
y_i	A binary variable, 1 if a facility is located at site i and 0 otherwise.
y'_i	A binary variable, 1 if a secondary facility is located at site i and 0 otherwise.

Objective function:

$$\min \sum_{i \in I} \sum_{j \in J} d_j c_{ij} x_{ij} + W d'_j c_{ij} x'_{ij} \tag{3.7}$$

Subject to:

$$\sum_{i \in I} x_{ij} = 1 \qquad \forall j \in J \tag{3.8}$$

$$\sum_{i \in I} x'_{ij} = 1 \qquad \forall j \in J \tag{3.9}$$

$$\sum_{i \in I} y_i = p \tag{3.10}$$

$$\sum_{i \in I} y'_i = p' \tag{3.11}$$

$$x_{ij} - y_i \leq 0 \qquad \forall i \in I; j \in J \tag{3.12}$$

$$x'_{ij} - y'_i \leq 0 \qquad \forall i \in I; j \in J \tag{3.13}$$

$$y'_i - y_i \leq 0 \qquad \forall i \in I \tag{3.14}$$

$$p - p' \geq 0 \tag{3.15}$$

$$y_i \in \{0, 1\} \qquad \forall i \in I \tag{3.16}$$

$$y'_i \in \{0, 1\} \qquad \forall i \in I \tag{3.17}$$

$$x_{ij} \geq 0 \qquad \forall i \in I; j \in J \tag{3.18}$$

$$x'_{ij} \geq 0 \qquad \forall i \in I; j \in J \tag{3.19}$$

In the multiple time setting p-median notation, an apostrophe (’) is used to denote variables relating to the secondary setting. Note that (3.8), (3.10), (3.12), (3.16), and (3.18) are identical to (3.2), (3.3), (3.4), (3.5), and (3.6). Additionally, each of these constraints have been duplicated ((3.9), (3.11), (3.13), (3.17), (3.19) respectively) to constrain the secondary facilities identically to the primary facilities. (3.14) has also been added to restrict the sites of secondary facilities to be shared with primary facilities, and (3.15) has been added so that the number of secondary facilities must be no greater than the number of primary facilities.

A weight variable, W, has also been added to the after-hours term in the objective function. This variable allows for a linear discount rate to be applied to adjust the relative cost of travel between time settings. For example, if policy dictates that after-hours travel is half as costly for patients, $W = 0.5$ incorporates this preference. If travel in either time setting is considered equally costly, $W = 1$ assigns equal weight to travel in both settings. Increasing W is expected to bias facility locations toward after-hours demand locations if demand differs between the regular and after-hours settings. This variable is proposed for the purpose of adjusting relative cost between time settings, but sensitivity analysis of this variable is not within the scope of this thesis.

3.3 Solution Approaches

To address both the regular and after-hours needs of patients, both time settings must be considered by the model. Three approaches for addressing both settings are proposed in this section. The first solves both time settings independently which gives the best case scenario for each. The second approach solves the time settings sequentially and locations selected in one setting are then required in the other time setting. This approach is expected to be more reflective of how the decision is made in practice. The final approach simultaneously solves both time settings to find the best compromise solution. This approach is expected to be the best method for optimally locating facilities in both settings while also considering facility use constraints.

3.3.1 Independent

In this approach, each time setting is considered independently. The standard p-median problem is used to solve the regular and after-hours time settings with no shared constraints or resources. This approach is less cost effective because it places facilities for use in either regular or after-hours care rather than sharing facilities between settings. However, this approach is expected to provide a lower bound for each time setting. Since this approach is the least constrained, it should provide the lowest overall cost.

3.3.2 Sequential

For this approach, the standard p-Median problem is used to solve for one time setting, and the solution for that time setting constrains the other time setting. This is more robust than the independent approach, since the results of one time setting are used to constrain the other. This is expected to more closely model real-world use of primary care facilities.

The sequential approach differs depending on which time setting is solved first. If the regular setting is solved first, the set of regular hour facility locations becomes the set of possible locations chosen by the p-median model for the after-hours p-Median problem. For example, consider a situation where three facilities are to be placed during the regular setting and two are to be located after-hours and the nodes $I \in [1, 2, 3, ..., n]$ are possible facility locations. If the model locates facilities at nodes 1, 2, and 3 in the regular setting then the after-hours setting facilities must be placed among nodes 1, 2, and 3 only.

The process for the sequential model when the regular setting is solved first (abbreviated as "RtA") is as follows:

- Solve standard p-median problem for the regular setting.
- Create a set (I') of nodes from the solution containing nodes where regular time setting facilities are located.
- Solve a p-median problem for the after-hours using I' instead of I as the set of possible facility locations.

If the after-hours problem is solved first, the nodes selected for after-hours facilities must be included in the regular setting solution. For example, consider a situation where one facility is located for the after-hours setting and three are located in the regular setting. If the model locates a facility at node 2 in the after-hours setting then a facility must also be set at node 2 in the regular setting solution. This is enforced by a constraint (shown by (3.20)) added to the regular hours problem so that regular hours facilities must be located at nodes where there are after-hours facilities. Note that this constraint requires that there must be at least as many regular facilities as after-hours facilities.

$$y_i - y'_i \geq 0 \qquad \forall i \in I \tag{3.20}$$

The process for the sequential model when the after-hours setting is solved first (abbreviated as "AtR") is as follows:

- Solve standard p-median problem for the after-hours setting.
- Solve a p-median problem for the regular setting with additional constraint (3.20).

3.3.3 Simultaneous

This approach is modelled by the multiple time setting p-Median problem (3.7 - 3.19). This method considers both time settings simultaneously, rather than one before the other. It is expected that this will provide a better overall result than the sequential method, since all possible location choices are considered in one problem rather than being constrained by a previous problem which is "blind" to the overall objective. For example, a desirable solution in one time setting may place facilities away from demand in the other time setting because the model is unaware of the details of the other time setting. By considering both time settings at the same time, this issue is avoided.

3.4 Model Formulation

The model was written in Python 3.7.3 using Jupyter Notebook. The linear program (LP) was implemented using the PuLP LP modeller, and a Python class was written with PuLP to formulate and solve p-Median problems when given the problem parameters. Once this class was confirmed to solve standard p-Median problems correctly, additional elements were added to the class to allow it to solve each of the solution approaches.

The class written to handle p-Median problems, `pmed`, will be described in this section. A class is a object-oriented programming template for new objects. In this case `pmed` is a template for p-Median problems. The class described in this section is used to format problem information, formulate a p-Median problem from the information, pass the problem to a linear program solver (such as Coin-or branch and cut (CBC), an open source solver, or Gurobi, a commercial solver), record how long each process takes, and report the results. This section will proceed through the functions used in `pmed`.

The class is initialized by providing a dictionary containing the number of facilities (p) to be placed and a cost matrix ($c_{ij}\ \forall\ i, j$) specifying the cost of assigning demand between any pair of nodes. This is the fundamental information required for a p-Median problem, and is thus included to define a new problem. These two pieces of information are retained as class attributes when creating a new `pmed` object as shown in Listing 3.1.

```
#pmed class definition
class pmed:

    def __init__(self, data):
        self.facilities = data['p']
        self.csts = data['c']
```

Listing 3.1: pmed Instantiation Function (`__init__()`)

The next function in the `pmed` class is `defprob()`. This function takes the attributes defined by the instantiation function (p and the cost matrix), and has an optional argument for demand. By default demand is assumed to be 1 at each node, but an ordered list of demand for each node may be supplied to `defprob()`. `defprob()` uses this information to define the objective function and each constraint for the problem.

```
def defprob(self, d = None):

    #Start timing the function
    self.tstart = time.time()

    #Assign names to class attributes
    p = self.facilities
    costs = self.csts

    #If demand is unspecified, assign demand = 1 to each node in a
   list
```

```
    if d == None:
        d = [1 for i in range(len(costs))]

    #Create list of nodes
    sites = [x for x in range(len(costs))]

    #Create a list of pairs for each pair of nodes
    Routes = [(i,j) for i in sites for j in sites]

    #Define decision variables x_ij and y_i
    route_vars = LpVariable.dicts("Route",(sites,sites),0,None)
    site_vars = LpVariable.dicts("Site",(sites),0,1,LpInteger)

    #Create a minimization linear program
    self.prob = LpProblem("pmedprob",LpMinimize)

    #Define objective Function
    self.prob += lpSum([costs[i][j]*route_vars[i][j]*d[j] for (i,j) in Routes]), "Sum of Costs"

    #Constraints
    for j in sites:
        self.prob += lpSum([route_vars[i][j] for (i) in sites]) == 1

    self.prob += lpSum([site_vars[i] for (i) in sites]) == p

    for j in sites:
        for i in sites:
            self.prob += route_vars[j][i] - site_vars[j] <= 0

    #Record time to formulate problem
    self.tdef = time.time() - self.tstart
```

Listing 3.2: pmed Problem Definition Function (`defprob()`)

Once the problem has been formulated with `defprob()`, it is ready to be solved. This is straightforward since PuLP has a built-in function to call linear program solvers to solve formulated problems. Thus, the `solveprob()` function of `pmed` calls PuLP's `solve()` function and specifies which solver to call if one is given, or calls the default solver (CBC) if no solver is specified.

```
def solveprob(self, solver = None):

    #Start timing the solution
    self.tstart = time.time()

    #If no solver is specified, PuLP's solve() defaults to using CBC
    if solver == None:
        self.prob.solve()

    #Other solvers may be called as an argument in PuLP's solve()
    function
    else:
        self.prob.solve(solver)

    #Finish timing the solution
    self.tsolve = time.time() - self.tstart
```

Listing 3.3: pmed Solution Function (`solveprob()`)

The remaining functions in `pmed` are for reporting the total cost of the problem, the decision variable values, and the times taken to define and solve the problem.

```
def soln(self):
    #Returns the solution cost
    return value(self.prob.objective)

def fullsoln(self):
    #Prints the optimality status of the problem, followed by every
    nonzero decision variable and the solution cost
    print("Status:", LpStatus[self.prob.status])
    for v in self.prob.variables():
        if v.varValue != 0:
            print(v.name, "=", v.varValue)
    print("Total Cost = ", value(self.prob.objective))

def times(self):
    #Returns time to define and solve the problem in seconds
    return self.tdef, self.tsolve
```

Listing 3.4: Solution and Timing Report Functions

Here is a demonstration of using `pmed` to solve a toy problem. Consider three nodes located at the vertices of a 3, 4, 5 Pythagorean triangle. These nodes are assigned demands of 5, 50, and 100, and 2 facilities are placed among these nodes. A diagram of the problem is shown in Figure 3.1.

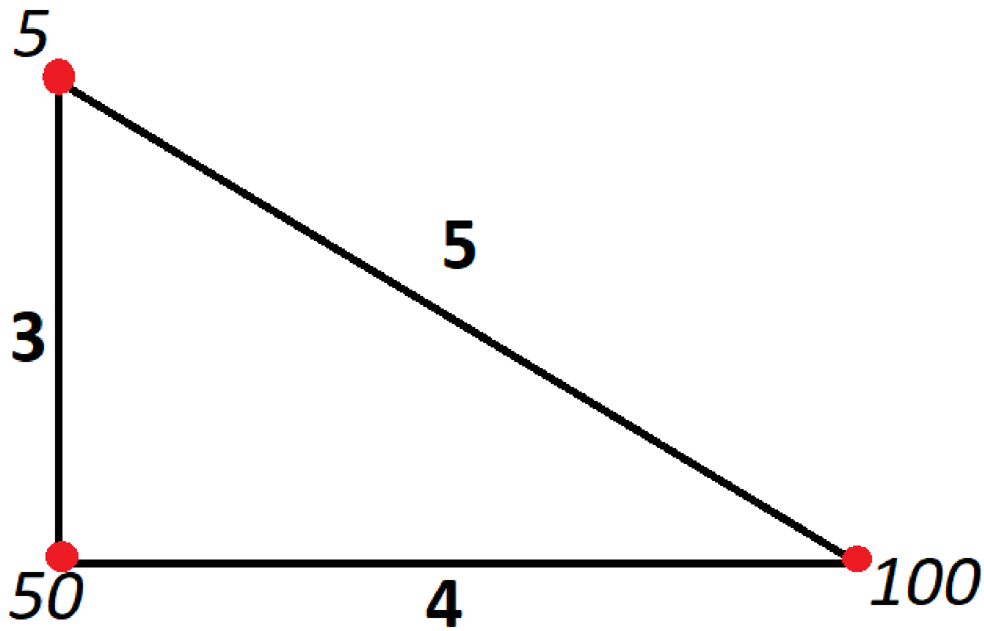


Figure 3.1: p-Median Toy Problem

```
#Set the number of facilities to 2
p = 2

#Define the distances between each node
        # 0 1 2
costs = [[0,3,5], #0
         [3,0,4], #1
         [5,4,0]] #2

#Define the demand at each node
    #0  1  2
d = [5,50,100]

```

```
14 #Put the information into a dictionary before creating a new problem
15 info = {'p':p, 'c':costs}
16
17 #Specify Gurobi as the solver to be used
18 solver = pulp.GUROBI_CMD()
19
20 #Create the problem
21 newprob = pmed(info)
22 #Define the problem
23 newprob.defprob(d)
24 #Solve the problem
25 newprob.solveprob(solver)
26
27 #Print the solution cost, problem status, and decision variable
     values
28 print("Cost: " + str(newprob.soln()))
29 newprob.fullsoln()
30
31 #Print the times taken to formulate and solve the problem in seconds
32 t = newprob.times()
33 print("Formulation time: " +str(t[0]))
34 print("Solution time: " +str(t[1]))
35
36 #The above returns the following when run:
37 Cost: 15.0
38 Status: Optimal
39 Route_1_0 = 1.0
40 Route_1_1 = 1.0
41 Route_2_2 = 1.0
42 Site_1 = 1.0
43 Site_2 = 1.0
44 Total Cost =  15.0
45 Formulation time: 0.0009970664978027344
46 Solution time: 0.22934317588806152
```

Listing 3.5: `pmed` Use Example

The output indicates that the optimal cost of the solution is 15 (line 44 of Listing 3.5), that facilities were located at sites 1 and 2 (lines 42 and 43), and that demand was allocated from node 0 to node 1, from node 1 to node 1, and from node 2 to node

2 (lines 39, 40, and 41).

3.5 Model Verification

To test this class, a set of previously solved p-Median problems from J. E. Beasley's OR-Library were solved using the code and the results were compared to the reference solutions [18]. There are 40 uncapacitated p-Median problems in the OR-Library with increasing complexity; the first problem contains 100 nodes with 200 connections and places 5 facilities, and the last problem contains 900 nodes, 19,200 connections, and places 90 facilities. The solution to each problem is also available for comparison in the OR-Library.

The problems are available as text files with the following format: the first line lists the number of vertices, the number of edges, and the number of facilities to place (p), and the remaining lines list the vertices for each edge and the cost of the edge. In the language used to describe p-Median problems in this thesis, vertices are nodes where facilities may be placed and edges are the costs between each node. Thus, the problem files may be considered to provide a hypothetical list of possible facility locations and the distances between locations which are directly connected. Demand is assumed to share the same set of nodes as possible facility locations, and demand at each node is 1. A sample of the first problem, found in `pmed1.txt` on the OR-Library website, is displayed in Table 3.4. This table shows that problem 1 contains 100 nodes, specifies 200 of the connections between these nodes, and places five facilities among these nodes. The second row indicates that the cost between nodes 1 and 2 is 30 (the units of cost are unspecified, but these could be kilometres for example). The third row indicates that the cost between nodes nodes 2 and 3 is 46. There are 198 more lines with similar information for different nodes.

Table 3.4: OR-Library p-Median Problem 1 Sample

100	200	5
1	2	30
2	3	46
3	4	1
...	...	...

The information given by the OR-Library files is not enough to formulate a p-Median problem without some processing. This is because the given data do not explicitly specify the distance between all nodes, which is required to solve the p-Median problem. To illustrate, consider an example: a file specifies the distances between locations A and B and between locations B and C, but not between A and C. It is possible to travel from A to C through B, but the cost of travelling from A to C is not explicitly given. Also, if multiple paths exist between A and C it is not immediately clear which path is the shortest. For the p-Median problem to be solved it is necessary to know the least possible cost between each pair of nodes. As Beasley notes, it is necessary to employ Floyd's algorithm before creating a p-Median problem from any file. Floyd's algorithm (also known as the Floyd-Warshall algorithm) is a process that can take a set of connected nodes, such as given by the OR-Library problems, and produce a complete matrix of the shortest distance between any pair of nodes [19]. For problem 1, this algorithm would produce a 100 by 100 symmetrical matrix containing the distances from each node to any other node. This was implemented using the `floyd_warshall` function in Python's `scipy.sparse` library. A function written to do this, `getdata()`, is shown in Listing 3.6.

```
def getdata(filename):
    prb = []
    inf = 9999

    #Parse the file into a series of strings
    with open(filename, 'r') as f:
        data = f.read().splitlines()

    #Create a matrix of integer values from the strings
    for line in data:
        prb.append(list(map(int, line.strip().split(' '))))

    #Read the number of nodes to expect
    nodes = prb[0][0]

    #Create a "maximum distance" matrix
    dist = [[inf for x in range(nodes)] for y in range(nodes)]

```

```
    #Replace maximum values with specified values
    for edge in prb[1:]:
        dist[edge[0] - 1][edge[1] - 1] = edge[2]
        dist[edge[1] - 1][edge[0] - 1] = edge[2]

    #Set self-travel distance to 0
    for i in range(nodes):
        dist[i][i] = 0

    #Create a matrix of shortest possible distances
    graph = csr_matrix(dist)
    dist_matrix = floyd_warshall(csgraph=graph, directed=False)

    #Read the number of facilities to place
    p = prb[0][2]

    return {'p':p, 'c':dist_matrix}
```

Listing 3.6: Problem Formatting Function (`getdata()`)

The formatting function takes a problem file location, parses the file, and returns two objects: p, an integer number of facilities to be placed in the problem, and c, a cost matrix indicating the cost between any two nodes (described in the OR model as $c_{ij}\ \forall\ i, j$). These outputs were put into the p-Median class and solved for each problem in the OR-Library. The results using the code described herein exactly matched the known solutions.

Once the basic `pmed` class was verified, the changes needed to implement the sequential and simultaneous approaches were gradually added and reviewed to ensure they accurately represented the mathematical formulations of the problems. The model returned similar results with these additions, and simple changes to the model's inputs yielded expected results.

3.6 Solver Selection

The linear models were solved using Gurobi, a commercial optimization solver. The default solver used by PuLP is CBC, an open-source solver which was initially used to test that the model functioned correctly. However, CBC proved to be too slow to solve all of the verification problems so Gurobi was used instead. For comparison, the first 24 OR-Library problems were solved and timed using each solver. Only the time taken for each solver to run was included in the comparison, time to format the data and formulate the problem were excluded. CBC took 52.6 minutes in total, and Gurobi took 4.3 minutes. The time taken for CBC to solve problems grew significantly as the problem complexity increased, and not all problems could be solved within a day of running when tested. Thus, the problems solved within an hour were deemed to be an acceptable cutoff for testing the difference in speed since Gurobi outperformed CBC by an order of magnitude.

Chapter 4

Results

To test how the models perform with real-world data, three case studies were completed using 2011 Canadian census data in the province of Nova Scotia [20]. The models were tested in the most densely populated region of Nova Scotia, Halifax, as well as a more rural area, Cape Breton. Lastly, the models were tested on a hypothetical setting where Cape Breton data were altered to create a discrepancy between regular and after-hours demand.

4.1 Data

The parameters for nodes, unit supply costs between nodes, and demand at each node were derived from the census data. Census dissemination areas (DAs) were used as the nodes of the model. Dissemination areas are the smallest geographical unit published by Statistics Canada. They are stable geographical areas typically populated by 400 to 700 people. The centroid of each DA was computed and used for possible facility locations. The population of each DA was used as the demand at each node for both time settings (except for the modified case study). The travel distance between nodes was considered the cost between nodes. The distance between nodes was the shortest distance between every pair of nodes in the Nova Scotia road network. Distance calculations between each DA are provided by McNamara [21]. Note that this method assumes travel distance within one node to be zero; for example, if a primary care practice is located at node i, patients within node i are assumed to travel a distance of 0 to this practice.

The dissemination areas selected to represent Halifax and Cape Breton were those contained within the Halifax Census Metropolitan Area (CMA), and the Cape Breton Census Agglomeration (CA). CMAs and CAs are geographical regions used by the census to indicate which DAs belong to a larger community. They are defined by regions where one or more adjacent municipalities are centred on a population core.

For a CA this means a core consisting of at least 10,000 people. A CMA must have a total population of more than 100,000 with at least 50,000 living within the core. For example, a CMA could have a core of 75,000 residents with 30,000 people living in surrounding municipalities. Images of these dissemination areas within Nova Scotia have been created using ArcMap, shown in Figure 4.1.

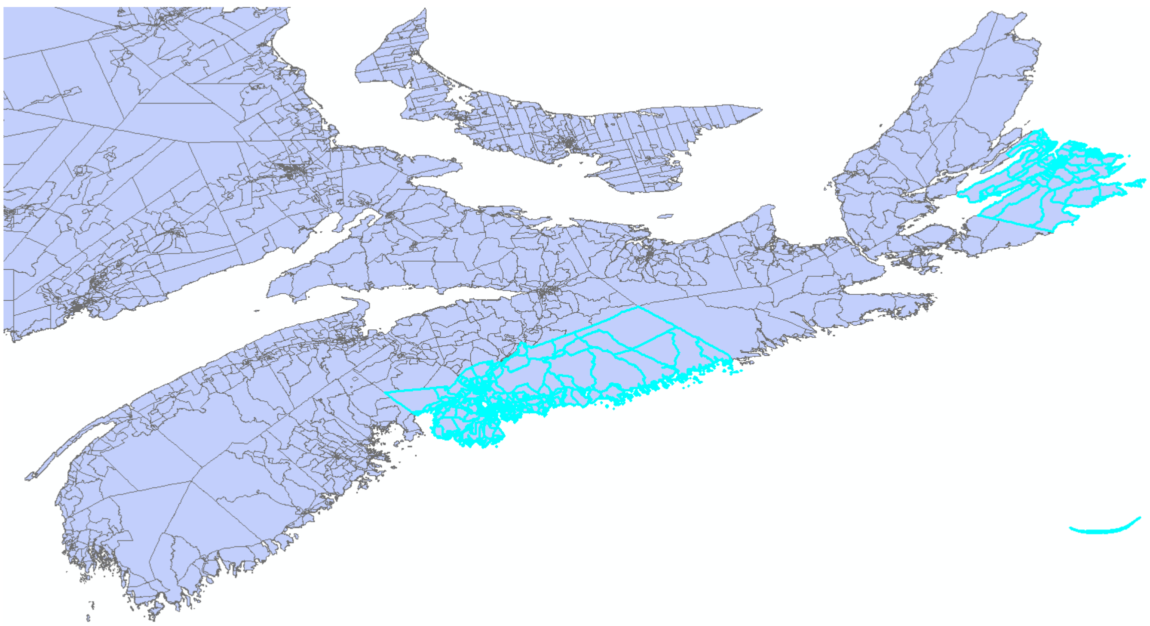


Figure 4.1: Nova Scotia with Halifax (center) and Cape Breton (upper-right) Dissemination Areas Highlighted

By using population to model demand in both settings, demand is considered to be identical during both the day and after-hours. To test the effect of geographical differences in demand between regular and after-hours settings, a hypothetical setting and data set were created by modifying the census data. This data set was used to simulate a scenario where a large portion of people moved from urban areas to rural areas after-hours. This may mimic a large population's commute to urban areas from rural areas for work. A cluster of DAs around Sydney (the most densely populated part of Cape Breton) was chosen to represent an urban area, and 25% of the population in these DAs was distributed proportionally among the remaining DAs in the after-hours to create the hypothetical setting. This was done by rounding the population in urban DAs to 75%, then increasing each remaining DA proportionately so that the total number of people remained the same.

4.2 Nova Scotia Case Studies

For each geographical test setting, the four solution approaches were run for differing numbers of regular and after-hours facilities. The total cost of each scenario (which can be expressed as the sum of all distance between each person and their primary care in kilometres) was recorded, along with the individual costs of the regular and after-hours settings. Note that

4.2.1 Halifax

In the Halifax setting, there are 594 DAs and a total demand of 390,096 people. Figure 4.2 displays the geography of the included DAs.

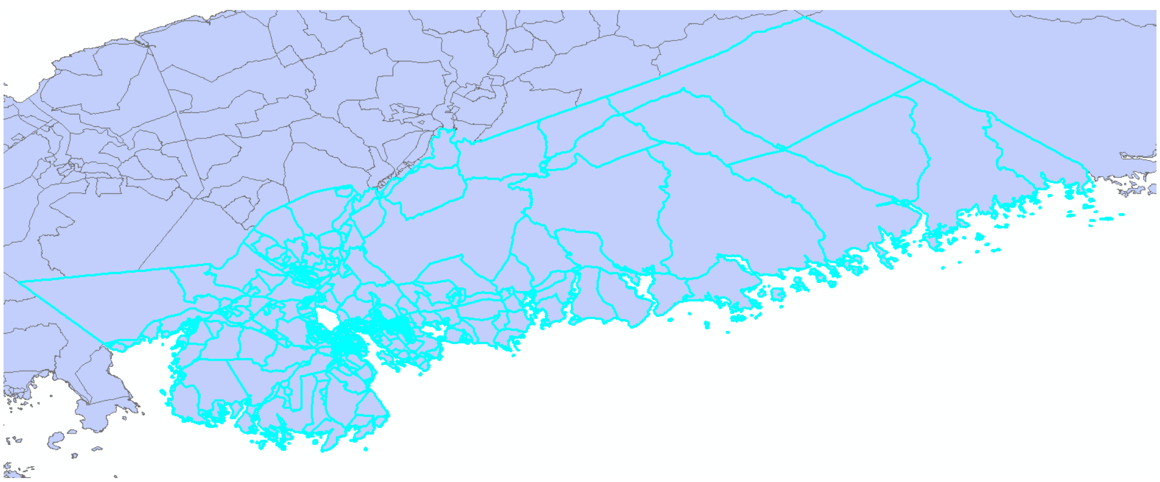


Figure 4.2: Halifax Dissemination Areas

Tests were run for increments of 5, 20, 40, 60, and 200 regular facilities with after-hours facilities increased in fifths. For example, the 60 regular facilities scenario was considered with 12, 24, 36, 48, and 60 after-hours facilities. In total, $5 \times 5 = 25$ facility placement scenarios were tested for each solution approach.

Table 4.1 contains results for each solution approach (independent, sequential, and simultaneous), and time periods for one scenario in the Halifax setting. There are two rows for the sequential approach, since this approach differs between regular to after-hours (AtR) and after-hours to regular (RtA). In the scenario used for this example, sixty facilities are located for the regular setting and twelve are located for the after-hours setting.

Table 4.1: Halifax Total Cost (km)

	$p = 60,\, p' = 12$		
Approach	Day-Time	After-Hours	Total
Independent	327,904	475,128	803,032
Sequential (RtA)	327,904	482,107	810,011
Sequential (AtR)	328,352	475,128	803,480
Simultaneous			803,277

By dividing the results in Table 4.1 by the total demand the average distance travelled by each person is calculated, as shown in Table 4.2.

Table 4.2: Halifax Average Cost Per Person (km)

	$p = 60,\, p' = 12$		
Approach	Regular	After-Hours	Total
Independent	0.8365	1.2121	2.0486
Sequential (RtA)	0.8365	1.2299	2.0664
Sequential (AtR)	0.8376	1.2121	2.0497
Simultaneous			2.0492

The lower the results are, the better the solution is. This is because a lower overall cost means that the demand (people in this case) need to travel less distance on average. Note that these results are based on the assumption that distance is the main metric of consideration, and economic costs are not a factor in the model outside of constraints that cause facilities to be shared across time settings. It is expected that as p and p' increase the cost will go down. This is because locating more facilities allows facilities to be more spread out and closer to demand on average. This can be easily seen if we plot the cost in the independent setting against the number of facilities, shown in Figure 4.3. This applies to both settings and each approach. If facilities are added the cost is expected to decrease.

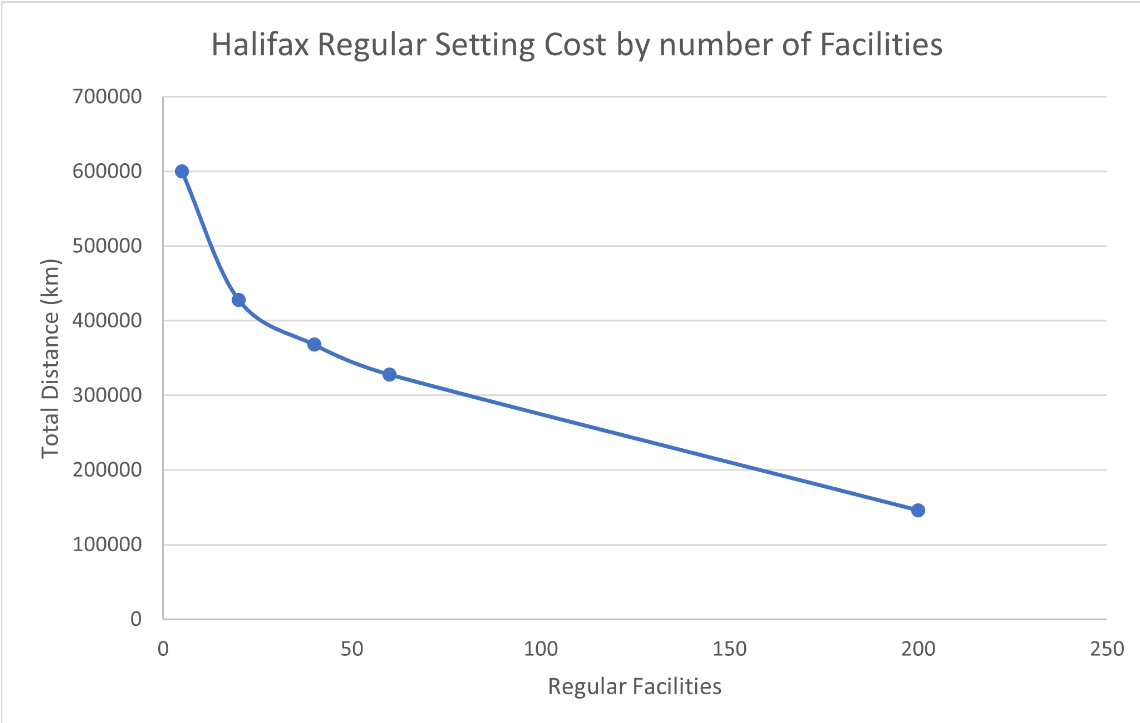


Figure 4.3: Halifax Regular Setting Independent Approach Cost by number of Facilities

In many cases, multiple approaches result in the same overall cost. When there are differences in cost they tend to be relatively small. However, the approaches tend to perform relative to each other as would be expected by how constrained they are. The least constrained approach, independent, has the lowest costs, while the sequential approaches have higher costs than the simultaneous approach. A sample is shown in Table 4.3, where the independent approach is treated as the base case to calculate the difference in other cases.

Table 4.3: Halifax Total Cost (km per person)

	$p = 60$, $p' = 36$		
Approach	Total	Difference	Percent Difference
Independent	705,368	0	0%
Sequential (RtA)	706,736	1,367	0.19%
Sequential (AtR)	705,633	265	0.04%
Simultaneous	705,578	210	0.03%

These results suggest that the approaches perform as expected relative to the degree to which they are each constrained.

Approach Results

The total costs of the independent approach for different values of p' and p are shown in Figure 4.4. Each series on the chart represent the ratio of after-hours facilities p' to the quantity of regular facilities p. For example, each point of the blue series labelled 0.2 represents a scenario where one fifth of the regular facilities are placed during the after-hours.

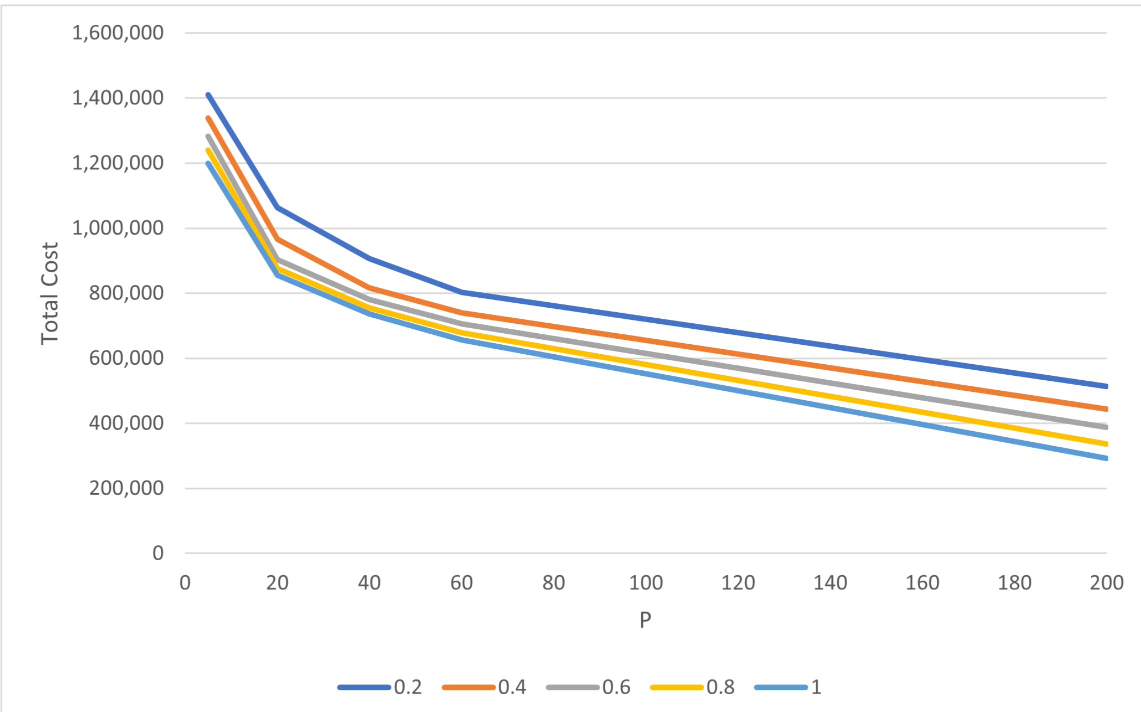


Figure 4.4: Minimum cost of p facilities for different ratios of p' (specified in legend). Independent approach in the Halifax setting is shown.

It is expected that the total cost is reduced as facilities are added. This is because greater numbers of facilities allow for solutions that reduce the average distance between demand and facilities, and this is demonstrated in Figure 4.4. It is also evident that increasing the ratio of p' to p decreases total cost. In the case of the independent approach, this simply means adding more facilities to the after-hours p-median problem, so a reduced cost is expected. Finally, it appears from the graph that the addition of more facilities has a decreasing marginal reduction in cost. That is, the benefit from each new facility decreases as the number of facilities being placed increases.

The trends of decreasing cost and decreasing marginal improvement in cost are apparent for each approach, and plotting the results from each approach produces very similar graphs to Figure 4.4. This indicates that these effects tend to be present for any of the tested approaches when locating facilities.

Approaches Comparison

Each approach broadly exhibits similar behaviour as the number of facilities to be located increases, but it is also appropriate to compare how each approach performs relative to each other. In most scenarios, the independent approach has the lowest cost, and the simultaneous approach a lower bound for the sequential approaches. This is an expected result, as approaches that are more constrained tend to have higher costs. This can be seen in Figure 4.5, where each series shows the percent increase in costs compared to the independent approach for five different scenarios. This is calculated with the formula (approach cost - independent cost) / independent cost.

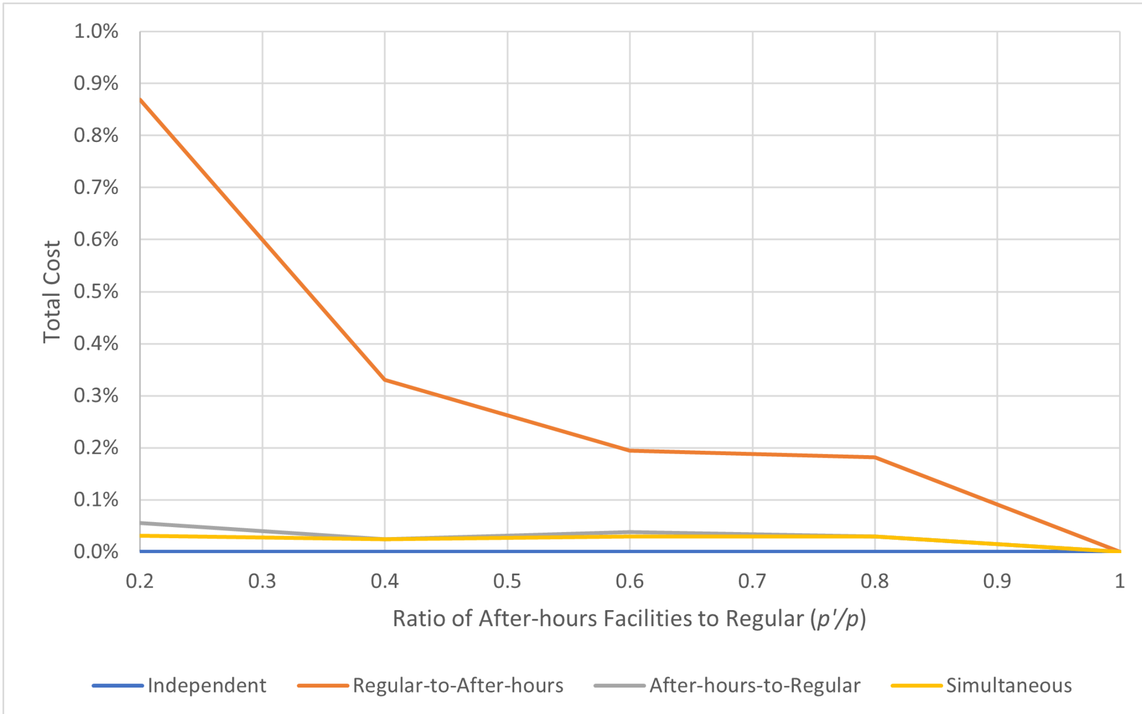
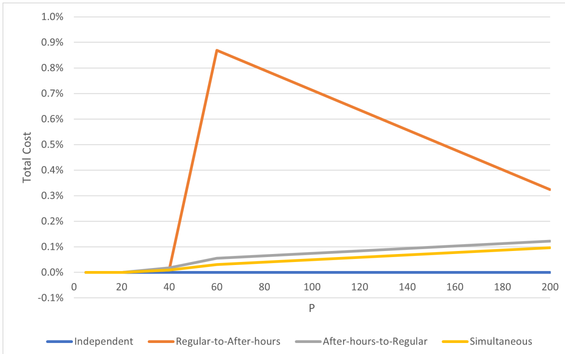


Figure 4.5: Percent cost increase of each approach compared to independent ($p = 60$)

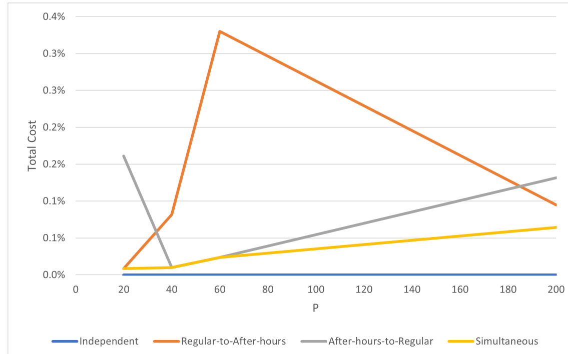
One trend that is apparent in Figure 4.5 is that each approach converges to the same cost as the number of after-hours facilities approaches the number of regular facilities. This makes sense when considering the data used for this case study. The demand at each node is considered to be the same during both the regular and after-hours time settings, so the differentiating factor between problems in each setting is the number of facilities being placed. If the number of facilities being placed is identical, then the problems being solved will be identical for this case data. This is not a phenomenon that would be expected if demand differs between time periods, and this is not expected to be observed in the third case study.

The phenomena of the independent approach providing the best cost followed by the simultaneous approach is consistent across each scenario. In some cases the RtA approach yields lower costs than the AtR, but in most cases the AtR approach is better. However, both are consistently worse than the simultaneous approach.

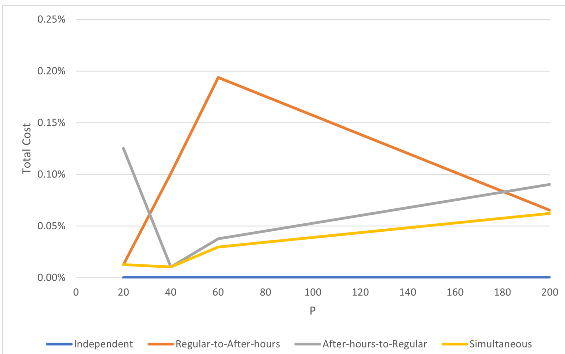
Some of the same trends are apparent when plotting the number of facilities on the X axis for static ratios of p' to p. In Figure 4.6 it is apparent that the simultaneous approach provides a lower bound to the sequential approaches.



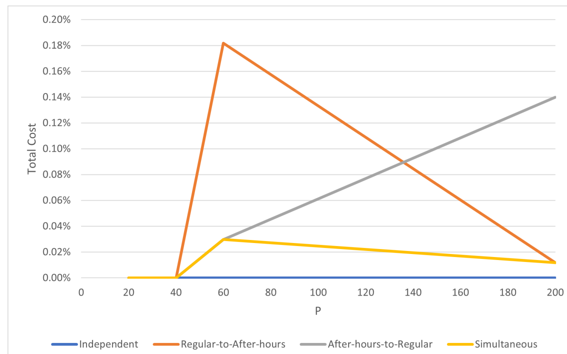
(a) $p'/p = 0.2$



(b) $p'/p = 0.4$



(c) $p'/p = 0.6$



(d) $p'/p = 0.8$

Figure 4.6: Percent cost increase of Halifax approaches compared to independent

These charts also display the discretized nature of the problem, where the solution for any approach can jump sharply between points. While solution approaches always follow the pattern of independent-simultaneous-sequential in terms of the lowest cost, whether the RtA or AtR solution is better depends on the scenario. It can be seen that for many scenarios the RtA sequential approach provides a worse solution than the AtR approach, but this is not always the case.

It is possible to explain some of the trends that cause differences between sequential approaches. As more nighttime facilities are added, the AtR solution becomes more constrained (since each after-hours facility becomes a constraint for the regular solution). This is easy to see for $p'/p = 0.4$ in Figure 4.6b (see full size in Figure C.2) by following the increasing grey line. Conversely, the RtA solution (shown by the orange line) improves as more facilities are added since every regular facility in this approach provides more possible solutions for the after-hours problem. Looking at the right side, the case seems to be that AtR tends to get worse and RtA gets better after the middle. This is logical since they are respectively becoming more and less constrained.

On the left side, it is sometimes the case that the AtR is the worse sequential approach. It may be the case that when there are enough daytime facilities located they also tend to include the best nodes for after-hours facilities as well (since the "best" nodes for demand satisfaction are the same in each setting for this test). However, picking very few after-hours facilities may constrain the regular problem to a bad solution. This can be seen for $p = 20$ when $p'/p = 0.4, 0.6$ (Figures 4.6b, 4.6c, see full size in Figures C.2, C.3). Though this is not always the case as is shown by the other charts in Figure 4.6, in which the AtR approach yields the same solution as the simultaneous approach. This is likely due to a discretization effect. Since the solutions involve discrete parameters and decision variables (the number of facilities, p, and the locations of those facilities, y_i), the solution costs are discontinuous. Since the differences in cost are relatively small between approaches and this effect is not present when there are many and few after-hours facilities, it is likely that the small number of facilities causes the solutions to "jump" in cost.

Initial condition scenarios ($p = 5$) include difficult behaviour to analyze, and have been excluded. The differences between approaches for $p = 5$ are quite significant, and they are difficult to categorize into trends. This is likely due to a pronounced discretization effect, which becomes more prominent as the number of facilities being placed is reduced. Since solutions are likely to differ greatly for adjacent, small values of p and p', attempts to analyze solution differences would not likely produce more meaningful results. If p and p' were tested in smaller iterations (e.g. for each integer less than 20), this may provide results with more apparent phenomena.

4.2.2 Cape Breton

Cape Breton is a less densely populated area of Nova Scotia than Halifax. The Cape Breton census agglomeration contains 101,619 people in 197 DAs, which are shown in Figure 4.7.

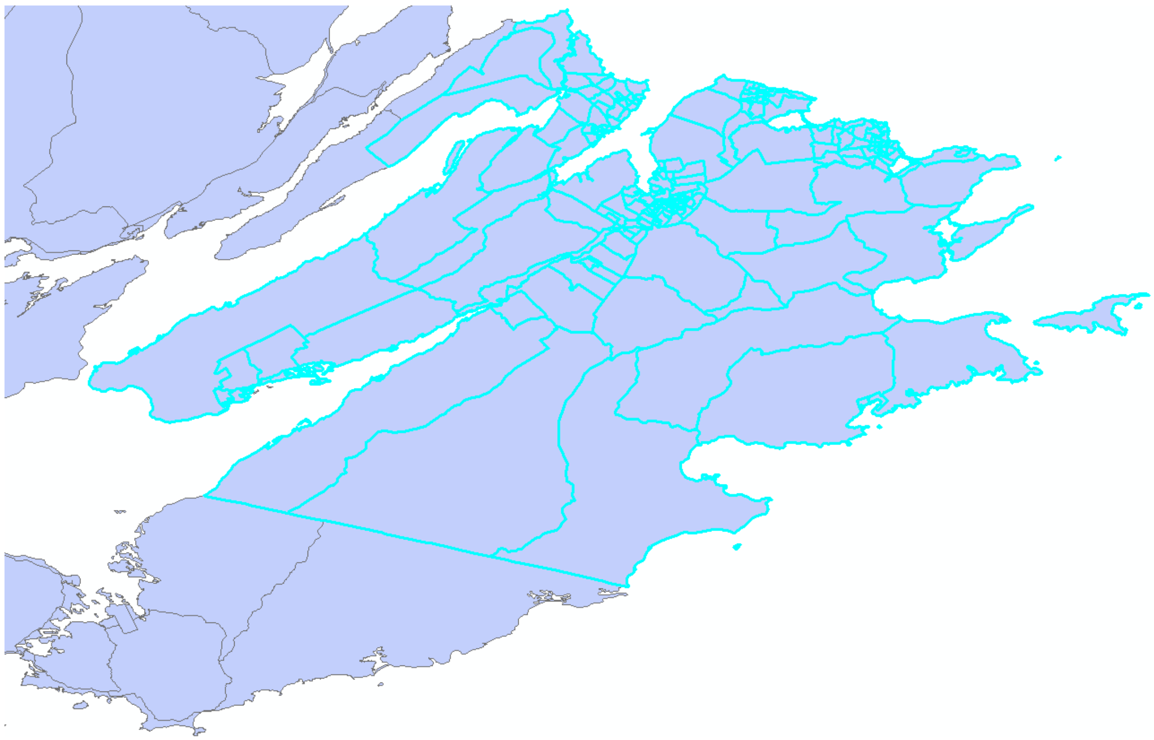


Figure 4.7: Cape Breton Dissemination Areas

Tests were run for increments of 5, 20, 40, 60, and 100 regular facilities with after-hours facilities increased in fifths.

Similar results to the Halifax tests are observed in the Cape Breton setting. This can be seen in Figure 4.8, which displays the same trends as Figure 4.4. Locating more facilities results in reduced costs, and the solution approaches differ in the same magnitude. The absolute solution costs are lower when the same number of facilities are placed, but this is expected with fewer people in the Cape Breton region.

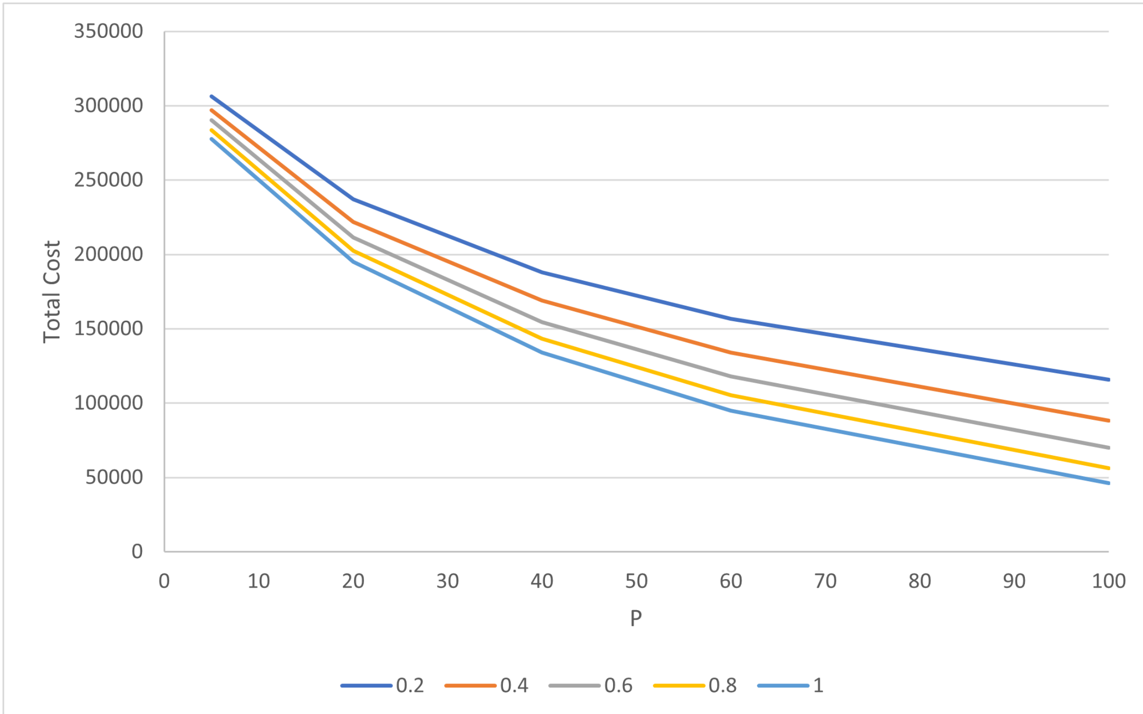
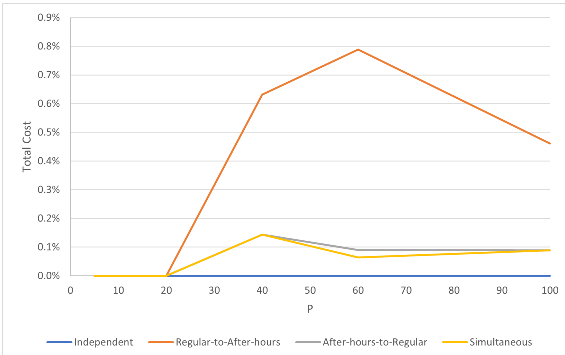


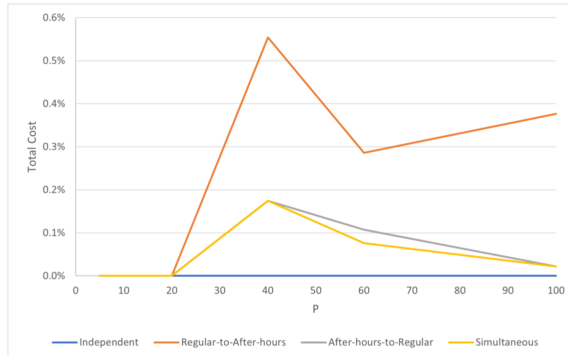
Figure 4.8: Minimum cost of p facilities for different ratios of p' (specified in legend). Independent approach in the Cape Breton setting is shown.

Approaches Comparison

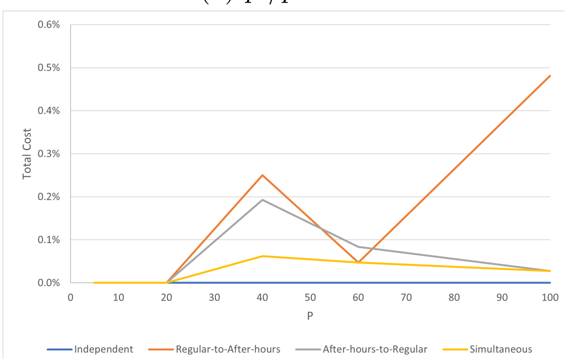
Comparing the different solution approaches for Cape Breton, many similarities to the Halifax tests can be observed in Figure 4.9. The same order of solution effectiveness matches the degree to which solutions are constrained (that is, the independent solution approach produces the lowest-cost solution, the sequential approaches produces the highest solution cost, and the simultaneous approach is in between the others).



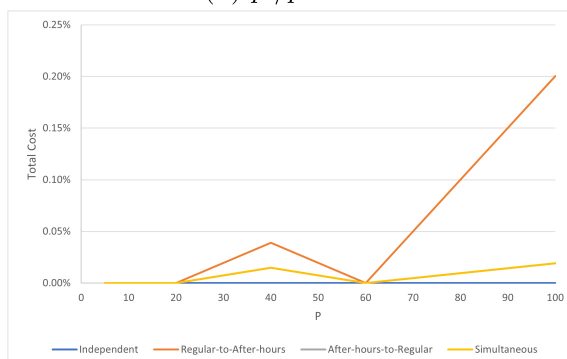
(a) $p'/p = 0.2$



(b) $p'/p = 0.4$



(c) $p'/p = 0.6$



(d) $p'/p = 0.8$

Figure 4.9: Percent cost increase of Cape Breton approaches compared to independent

The RtA solution approach most frequently has the highest cost, although there are cases where the AtR produces a higher-cost solution. This suggests that the RtA is typically the worst approach, which is consistent with the results of the Halifax tests. Another similarity to the Halifax tests is that all approaches produced the same solution for $p'/p = 1$. This is not surprising, since these tests also do not have different demand between time settings.

There is also a discretization effect that causes discontinuities between tests. For example, for each series there is a notable increase for $p = 40$; this may be due to the constrained approaches being restricted to a relatively poor solution for this particular number of facilities when compared to the independent approach. This result may be due to the geographical distribution of demand in these tests which increases the cost of facility-sharing constraints when $p = 40$.

4.2.3 Cape Breton Modified

In the modified Cape Breton setting there are still 101,619 people, and 197 DAs with the same geography. However, a set of urban DAs shown in Figure 4.10 have reduced after-hours demand by 25% , and other DAs have proportionally increased after-hours demand.

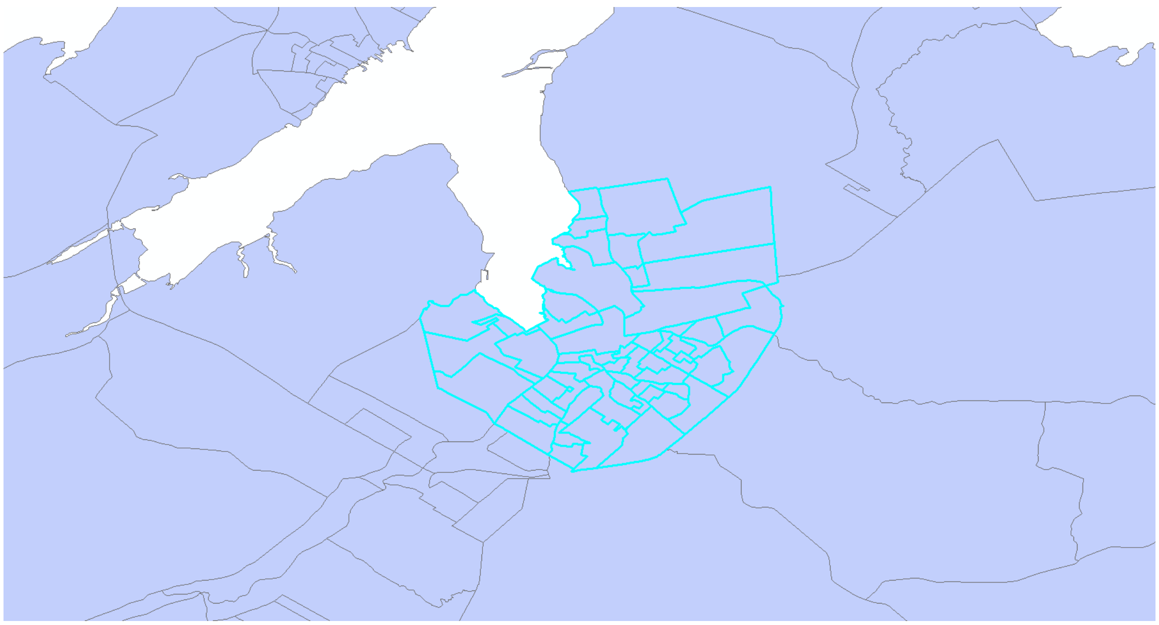


Figure 4.10: Cape Breton Urban Dissemination Areas

Plotting the test results for the independent solution approach (Figure 4.11) produces a very similar graph to the other geographical settings. However, the solution costs are lower than in the unmodified Cape Breton tests (between 4% and 15% per test, 7% on average).

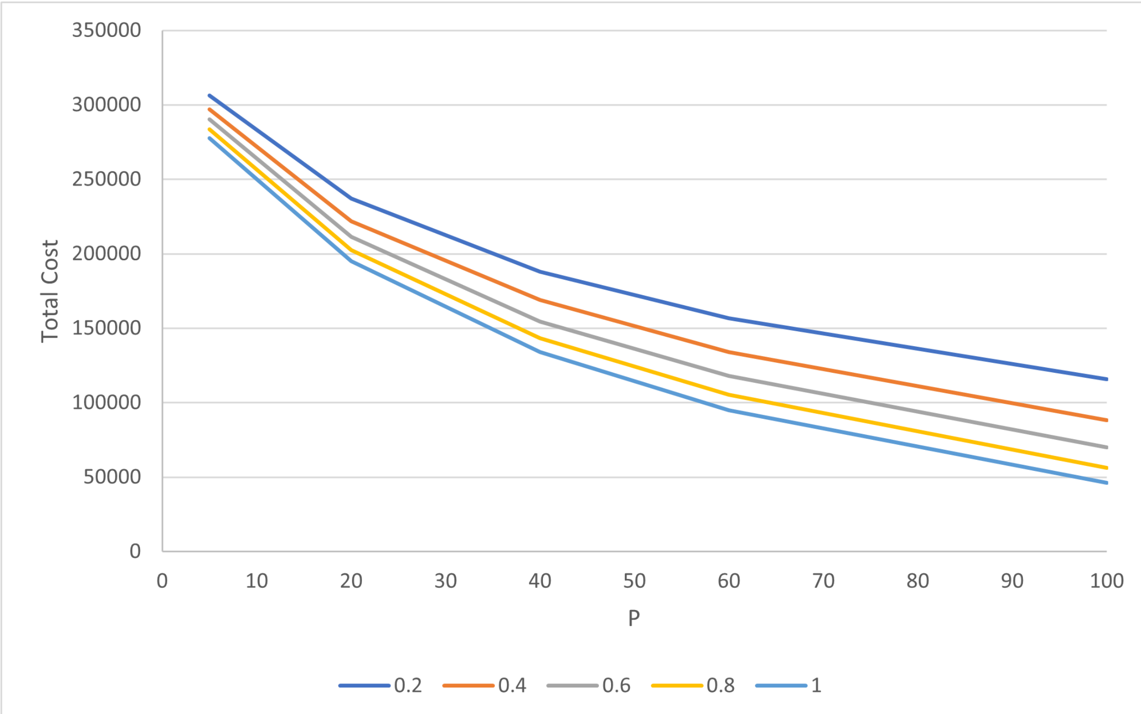
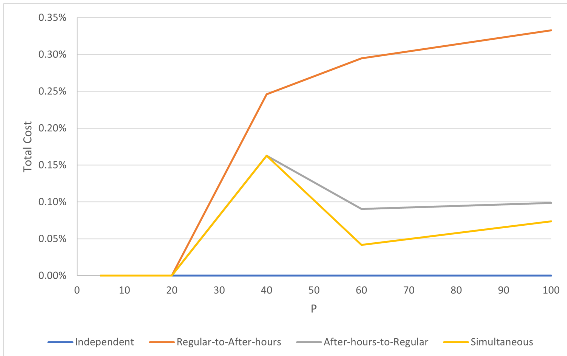


Figure 4.11: Minimum cost of p facilities for different ratios of p(specified in legend). Independent approach in the Cape Breton Modified setting is shown.

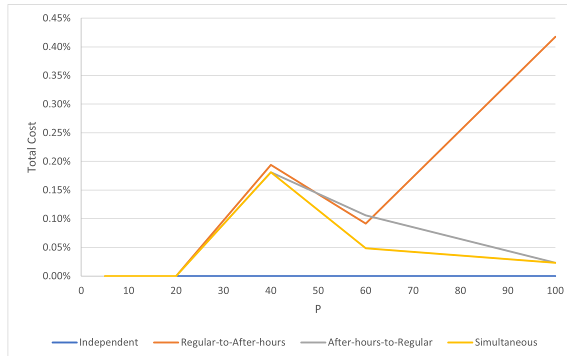
Two possible explanations for the discrepancy in solution cost are apparent. The first is that this population distribution allows for solutions with lower distances travelled on average. The second reason is that dissemination areas are defined to have a specific population range, and the model assumes that patients travel no distance to a facility within their DA. By creating DAs with differing numbers of people facilities located in higher population DAs will effectively consider more people to travel 0 distance, and DAs with fewer people will be left to travel. These two effects could account for the lower costs of the modified Cape Breton setting.

Approaches Comparison

As with the other geographical settings, the different solution approaches may be compared within the modified Cape Breton setting. The test results are shown in Figure 4.12.

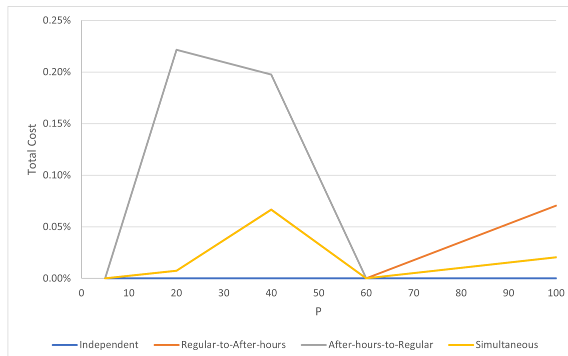


(a) $p'/p = 0.2$



(b) $p'/p = 0.4$

(c) $p'/p = 0.6$



(d) $p'/p = 0.8$

Figure 4.12: Percent cost increase of Cape Breton Modified approaches compared to independent

An expected result is that the solution approaches remain bounded by order of constraint. That is, the independent solution approach provides a lower bound, and the simultaneous approach remains between the sequential and simultaneous approaches. The same "hump" as observed in the unmodified Cape Breton tests for $p = 40$ is also present in these results, which suggests that this effect is due to a feature of the geography.

A different phenomenon in these graphs is that the sequential approaches seem to have a relationship with the number of facilities placed. It appears that for $p < 60$, as p' increases the cost of the RtA approach decreases and the AtR cost increases. This suggests that for the modified scenario, as the number of after-hours facilities increases relative to the number of regular facilities, it is more effective to consider the regular setting first. This is particularly noticeable for $p'/p = 0.6, 0.8$, as shown in Figures 4.12c, 4.12d (see full size in Figures C.11, C.12). This effect may be due to the increased demand in the rural DAs in the after-hours setting. With a large number of after-hours facilities placed first, the model likely places many of them away from the more densely-populated urban area, which then constrains the secondary problem in the regular time setting to keep many facilities away from the area with the most demand in that time setting. At $p \geq 60$, it is possible that there are enough facilities to overcome this effect. For example, 60 facilities may be enough to achieve sufficient spread across the area in the regular time setting even if 80% of those facilities have already been located according to the after-hours.

Another important point for comparison is when $p'/p = 1$, which is when the number of facilities placed is the same in the regular and after-hours settings. The results for this condition are shown in Figure 4.13.

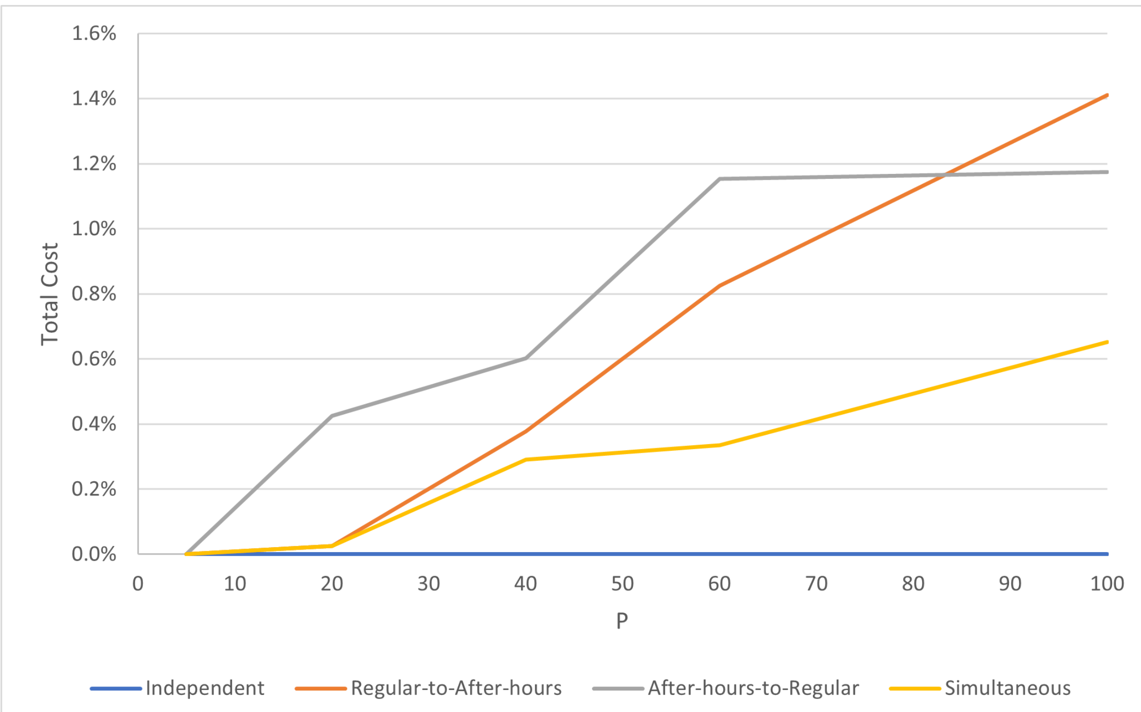


Figure 4.13: Cape Breton Modified percent cost increase of each approach compared to independent ($p'/p = 1$)

In the Halifax and unmodified Cape Breton tests, there was no difference between solution approaches when $p'/p = 1$. This was expected since demand did not differ between the regular and after-hours time settings, and thus the facility-sharing constraints would have no effect when the same number of facilities are placed in both time settings. However, Figure 4.13 demonstrates that differing demand between time settings affects the effectiveness of each solution approach.

An interesting result is that the non-independent solutions become relatively less effective as the number of facilities increases. This suggests that the constraints that cause facilities to be shared across time settings become more constraining as the number of facilities increases. While the actual solution cost decreases as the number of facilities increases, this effect may contribute to the decreasing marginal efficiency of additional facilities.

Chapter 5

Discussion

In this thesis, multiple approaches to optimally locating primary care practices were tested in three case studies. The results of these tests reinforce the position that some methods produce better results than others.

5.1 Key Findings

In both the Halifax and Cape Breton settings, the absolute distance travelled differed little (less than 1% difference) between solution approaches when using the same number of facilities. In terms of distance travelled per person, the largest difference in cost was in the magnitude of tens of meters, which may be considered marginal when looking at a population of hundreds of thousands of people.

However, this result is unsurprising when considering certain details of the study. Foremost, the demand assigned at each node was considered to be the same for both the day and after-hours. The difference between settings is the number of facilities located, and the difference between models are the constraints used to restrict facility location. When most of the model parameters are homogeneous across the time setting, it is logical that each approach results in a similar cost.

It is also necessary to consider that the region of study contains urban and rural areas with very different population densities. While the total solution cost may not have changed much on average per person between approaches, this is likely not distributed equally among demand. If a large portion of this cost discrepancy was due to lower distance travelled by people in rural areas, this could mean a much larger difference for those people since there are fewer of them than the urban population. This could mean that the model has a more pronounced effect for the rural population than is apparent when considering urban demand in the average denominator. It is also worth noting that the large portion of people who live in the dense, urban part of Halifax will never be more than a kilometer or two away from a PCP, and optimising for geospatial access is likely best applied to rural demand where the population density is lower and services are farther away.

Another important result is that each solution approach performs relatively as expected. The independent approach represents a naive scenario where both the day and after-hours facility locations can be treated independently with no cost, and this approach always had the lowest total cost. Conversely, the more constrained approaches had higher costs, with the most constrained approaches (the sequential RtA and AtR approaches) resulting in higher cost solutions. While this was expected, an interesting result was that which sequential approach was better depended on the test conditions. The regular-to-after-hours approach yielded worse results in most cases, which is noteworthy since this is the most likely real-world method for determining facility locations.

For the Halifax and Cape Breton tests, each solution approach provided the same result when the number of regular and after-hours facilities were the same. Since the demand in these tests was identically located across both time settings, it makes sense since the same solution would be identical for both settings. However, in the modified Cape Breton setting this was not the case. Since the demand differed between time settings in these tests, a difference could be observed in the efficacy of each approach (seen in Figure 4.13).

5.2 Limitations

There are some limitations of the data used in the case study that could be improved upon for more accurate modelling of primary care needs in Nova Scotia. One problem is how the cost parameter c_{ij} assumes that travel distance is an equal measure of access for all patients under all circumstances. The values of c_{ij} are based on minimum road distance from one DA centroid to another, but the time and location are factors that may change how accessible this distance really is for a patient. For example, travelling across a dense city during rush hour is likely to take much longer than travelling the same distance in a rural setting late in the evening. A measure of time travelled may be a better metric for measuring access, including the time taken for different forms of available transport in the area.

Another issue with the c_{ij} values is that demand assigned to the same node is assumed to travel no distance. That is, $c_{ii} = 0 \; \forall i \in I$. This assumption may be robust in the case of very small DAs where the distance from the centroid to any edge is negligible. However, as DA size increases the distance that a patient might have to travel within the DA also increases. A possible improvement may be to calculate the average distance within a DA from the centroid to each other point and use this value as the intranodal travel distance for each node. While the number of facilities to be placed is fixed in the p-median formulations used in this study, this would be particularly important to address in a model that considers the number of facilities as a decision variable. Since every additional facility allows demand at the same node to travel no distance when $c_{ii} = 0$, this may unrealistically incentivize adding facilities to reduce distance.

Finally, analysis for $p = 5$ was excluded from the Halifax case study. This is because the results for these tests were highly variable. It is likely that this is a discretization effect and that the low number of facilities yielded very different adjacent solutions. Sampling more cases may improve the interpretability of the results, but it is possible that this effect is unavoidable for low values of p. Testing this was excluded from the scope of this thesis.

5.3 Future Research

The case study indicates that the simultaneous approach is an effective method of addressing both the regular and after-hours time settings when considering practical facility location constraints. However, there are other considerations when addressing this problem that should be considered in future work. These points are outside of the scope of this research, but may be important to consider in the context of developing real-world policy for primary care resource allocation.

Firstly, the p-median problem was determined to be an appropriate OR model for demonstrating the merits of considering multiple time settings simultaneously. Particularly since the case studies cover populations with varying densities, the objective of the p-median to minimize all distance travelled is desirable when compared to other models that were considered. In other contexts, the p-center problem, the location set covering problem (LSCP), or the maximum covering location problem (MCLP) may be more appropriate [22, 23, 24]. Respectively the objectives of these models are to minimize the greatest distance travelled by any customer, to minimize the number of facilities needed to cover a population (coverage meaning ensuring none of the population are further than a specified distance), and to maximise the number of customers within a specific distance of a facility. These models are compatible with the simultaneous approach to treating the regular and after-hours settings, and there may be scenarios where these are more appropriate than the p-median problem. It may also be of interest to use different models in different time settings for unique objectives. For example, the p-median could be more appropriate during regular hours when there are more facilities available, but if there are only a small number of after-hours facilities the LSCP might be better able to maximise the number of people who have access to a facility if they need it. This would provide a solution where regular facilities provide an optimal solution for the average patient, and the number of people who are far from after-hours facilities would be minimized. The p-center problem may be of particular interest for improving accessibility in rural areas since it provides an equitable solution by reducing the greatest distances travelled.

Another potential adjustment to the models used for primary care facility location is to incorporate capacitated facilities. Many GPs prefer to work in practices with several other physicians, and realistically this is the case for many PCPs. Rather than placing individual facilities with no regard for capacity, a model that places facilities with multiple GPs that can handle a specified panel size would more accurately reflect the needs of physicians and patients.

It may also be appropriate to relax facility location constraints and instead add facility costs to control the number of facilities opened. While it is clear that adding more facilities will always allow for solutions with lower distances between demand and facilities, this does not account for the economic costs of additional facilities. A multicriteria decision analysis may help determine the cost-effectiveness of the number of facilities available. Since the marginal improvement in patient distance travelled decreases as the number of facilities increases, this suggests that there is a point where the gains in accessibility are outweighed by the additional facility costs. A model which replaces facility location constraints with facility costs may also provide more effective solutions for both the regular and after-hours. It is intuitive to assume that it is cost-effective for after-hours care to use the same facilities as regular care, but there may be cases where this is not the best solution. By allowing for these facilities to be separated at a higher cost, a model could potentially provide a better result than if it were constrained to sharing facilities between time settings. It may also be appropriate to examine the effects of the weighting variable, N, on the model. By decreasing or increasing N, solutions that respectively decrease distances during regular service or the after-hours should be favoured.

While census population data was used to model demand for primary care services in the case studies, this information does not necessarily provide an accurate representation of demand in different areas. Of particular importance the census data do not capture the disparity across individuals in demand for healthcare, nor is the geographical shift in population between time settings measured well by the census. Given that this information may not be easily obtained, a possible approach to testing how these factors could affect results with the available data may be to make adjustments to demand in different areas using demography. For example, an area's average demand may be increased or decreased by a factor derived from the deviation from median age or income. For demand differences across time settings, further changes could be made such as was done for the modified Cape Breton case study.

The case studies are a greenfield approach, but it may be of interest to compare solutions to real-world conditions. The difference between optimal facility locations suggested by the model and actual PCPs may reveal factors that have been unaccounted for. Alternatively, discrepancy could demonstrate the utility of this work in planning health services.

To conclude, more research is needed on after-hours primary care, but it is possible to improve patient access to care by considering regular and after-hours care simultaneously in facility location problems. In particular, this research is expected to benefit rural Canadians at significant distance from primary care providers. By improving the spatial accessibility of health services, use of these services and health outcomes are likely to improve.

Bibliography

[1] Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews. 2015 jan;4:1. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25554246 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4320440.

[2] Guagliardo MF. Spatial accessibility of primary care: Concepts, methods and challenges. International Journal of Health Geographics. 2004;3:1–13. Available from: http://www.ij-healthgeographics.com/content/3/1/3.

[3] Morgan JS, Graber-Naidich A. Small system dynamics model for alleviating the general practitioners rural care gap in Ontario, Canada. Socio-Economic Planning Sciences. 2019;66(April 2018):10–23. Available from: https://doi.org/10.1016/j.seps.2018.04.005.

[4] Güneş ED, Yaman H, Çekyay B, Verter V. Matching patient and physician preferences in designing a primary care facility network. Journal of the Operational Research Society. 2014;65(4):483–496.

[5] O'Malley AS, Samuel D, Bond AM, Carrier E. After-Hours Care and its Coordination with Primary Care in the U.S. Journal of General Internal Medicine. 2012 nov;27(11):1406–1415. Available from: http://link.springer.com/10.1007/s11606-012-2087-4.

[6] Crighton EJ, Bordman R, Wheler D, Franssen E, White D, Bovett M, et al. After-hours care in Canada: analysis of the 2001 National Family Physician Workforce Survey. Canadian family physician Medecin de famille canadien. 2005 nov;51(November 2005):1504–1505. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16926942 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1479478.

[7] Nova Scotia Department of Health and Wellness. Base Case - Workforce Forecast 2016 to 2025; 2016. Available from: https://novascotia.ca/dhw/ShapingPhysicianWorkforce/docs/BASECASE-Forecast-TOTAL-2016-to-2025.pdf.

[8] Doctors Nova Scotia. Fixing Nova Scotia's Primary Health Care Problem: Physicians' Recommendations to Improve Primary Care In Nova Scotia; 2017. Available from: https://doctorsns.com/sites/default/files/2018-04/PrimaryCarePosition.pdf.

[9] Güneş ED, Nickel S. Location Problems in Healthcare. Laporte G, Nickel S, Saldanha da Gama F, editors. Cham: Springer International Publishing; 2015. Available from: http://link.springer.com/10.1007/978-3-319-13111-5.

[10] Khan AA, Bhardwaj SM. Access to Health Care: A Conceptual Framework and its Relevance to Health Care Planning. Evaluation & the Health Professions. 1994 mar;17(1):60–76. Available from: http://journals.sagepub.com/doi/10.1177/016327879401700104.

[11] Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372. Available from: https://www.bmj.com/content/372/bmj.n160.

[12] Reuter-Oppermann M, Vanberkel P, Ziedins I. Day and night location planning for general practitioners [Unpublished manuscript]; 2019.

[13] Reuter-Oppermann M, Nickel S, Steinhäuser J. Operations research meets need related planning: Approaches for locating general practitioners' practices. PLoS ONE. 2019;14(1):1–10.

[14] Ahmadi-Javid A, Seyedi P, Syam SS. A survey of healthcare facility location. Computers & Operations Research. 2017 mar;79:223–263. Available from: http://dx.doi.org/10.1016/j.cor.2016.05.018 https://linkinghub.elsevier.com/retrieve/pii/S0305054816301253.

[15] Graber-Naidich A, Carter MW, Verter V. Primary care network development: The regulator's perspective. Journal of the Operational Research Society. 2015;66(9):1519–1532.

[16] Reuter-Oppermann M, Rockemann D, Steinhäuser J. A GIS-Based Decision Support System for Locating Primary Care Facilities. In: Za S, Drăgoicea M, Cavallari M, editors. Exploring Services Science. vol. 279 of Lecture Notes in Business Information Processing. Cham: Springer International Publishing; 2017. p. 210–222. Available from: http://link.springer.com/10.1007/978-3-642-36356-6.

[17] Daskin MS, Maass KL. The p-median problem. In: Location science. Springer; 2015. p. 21–45.

[18] Beasley JE. OR-Library; 2018. Available from: http://people.brunel.ac.uk/ mastjjb/jeb/orlib/pmedinfo.html.

[19] Cormen TH, Leiserson CE, Rivest RL, Stein C. The Floyd-Warshall Algorithm. In: Introduction to Algorithms. 3rd ed. Cambridge: MIT Press; 2009. p. 693–700.

[20] Dissemination Area Boundary File, 2011 Census. Statistics Canada Catalogue no 92-160-X. 2021. Available from: https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2011-eng.cfm.

[21] McNamara L, Vanberkel P, Petrie D, Carter A. An Application and Framework for Evaluating Emergency Department Networks using Location Analysis and Geographic Information Systems; 2018.

[22] Calik H, Labbé M, Yaman H. p-Center Problems. In: Location Science. Cham: Springer International Publishing; 2015. p. 79–92. Available from: http://link.springer.com/10.1007/978-3-319-13111-5_4.

[23] ReVelle C, Toregas C, Falkson L. Applications of the Location Set-covering Problem. Geographical Analysis. 2010 sep;8(1):65–76. Available from: http://doi.wiley.com/10.1111/j.1538-4632.1976.tb00529.x.

[24] Church RL, ReVelle C. The maximal location covering problem. Papers in Regional Science. 1974;32(1):101–118.

[25] Arnolds I, Nickel S. Layout Planning Problems in Health Care. In: Eiselt HA, Marianov V, editors. Applications of Location Analysis. vol. 232. Switzerland: Springer International Publishing; 2015. p. 109–152. Available from: http://link.springer.com/10.1007/978-3-319-20282-2_5.

[26] Asada Y, Kephart G. Equity in health services use and intensity of use in Canada. BMC Health Services Research. 2007;7. Available from: http://www.biomedcentral.com/1472-6963/7/41.

[27] Bruno G, Giannikos I. Location and GIS. In: Laporte G, Nickel S, Saldanha da Gama F, editors. Location Science. Switzerland: Springer International Publishing; 2015. p. 509–536. Available from: http://link.springer.com/10.1007/978-3-319-13111-5_19.

[28] Burkey ML, Bhadury J, Eiselt HA. Voronoi Diagrams and Their Uses. In: Eiselt HA, Marianov V, editors. International Series in Operations Research and Management Science. vol. 155. Boston, MA: Springer; 2011. p. 445–470. Available from: http://link.springer.com/10.1007/978-1-4419-7572-0_19.

[29] Canadian Institute for Health Information. How Canada Compares: Results From The Commonwealth Fund 2015 International Health Policy Survey of Primary Care Physicians. Ottawa, ON: CIHI; 2016. Available from: https://www.cihi.ca/en/health-system-performance/performance-reporting/international/commonwealth-fund-survey-2015.

[30] Church RL. Geographical information systems and location science. Computers & Operations Research. 2002 may;29(6):541–562. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0305054899001045.

[31] Doctors Nova Scotia, Nova Scotia College of Family Physicians. The Backbone of Primary Health Care: The Role and Value of Family Physicians in Nova Scotia; 2019.

[32] Farahani RZ, SteadieSeifi M, Asgari N. Multiple criteria facility location problems: A survey. Applied Mathematical Modelling. 2010 jul;34(7):1689–1709. Available from: http://dx.doi.org/10.1016/j.apm.2009.10.005 https://linkinghub.elsevier.com/retrieve/pii/S0307904X09003242.

[33] Graber-Naidich A, Carter MW, Verter V. Restructuring the resident training system for improving the equity of access to primary care. European Journal of Operational Research. 2017;258(3):1143–1155.

[34] Kephart G, Asada Y. Need-based resource allocation: Different need indicators, different results? BMC Health Services Research. 2009;9. Available from: http://www.biomedcentral.com/1472-6963/9/122.

[35] Marchildon GP, Hutchison B. Primary care in Ontario, Canada: New proposals after 15 years of reform. Health Policy. 2016 jul;120(7):732–738. Available from: http://dx.doi.org/10.1016/j.healthpol.2016.04.010 https://linkinghub.elsevier.com/retrieve/pii/S0168851016300872.

[36] Marianov V. Location Models for Emergency Service Applications. In INFORMS Tutorials in Operations Research. 2017;(October):237–262.

[37] Sridharan R. The capacitated plant location problem. European Journal of Operational Research. 1995;87(2):203–213. Available from: https://ac.els-cdn.com/037722179500042O/1-s2.0-037722179500042O-main.pdf?_tid=113db3f8-4575-4c9a-b904-7928caa98b51&acdnat=1536143130_7edc3dbc783cb9595c2bae5ed08d37a5.

[38] Terashima M, Jones P, Levy A, Christian E, Ahrens B, Dowling L, et al. Developing Nova Scotian community clusters for health service planning and research: An inter-sectoral, collaborative approach to the development of community cluster boundaries. Halifax, Nova Scotia: Maritime SPOR SUPPORT Unit; 2016. Available from: www.healthatlas.ca.

[39] Tomblin Murphy G, Kephart G, Lethbridge L, O'Brien-Pallas L, Birch S. Planning for what? Challenging the assumptions of health human resources planning. Health Policy. 2009;92(2-3):225–233.

Appendix A

Full Selection of Papers for Review

Table A.1 details the categorization of the texts selected for review, and whether each study was included for analysis. The categories in this table are study type, Day (relevant to regular operations), AH (relevant to after-hours operations), LS (location science), OR (operations research), HC (healthcare), and Selected (whether the reviewed text was included for further analysis or not).

Table A.1: Papers Selected For Review

Study ID (Author)	Year	Type	Day	AH	LS	OR	HC	Selected
Ahmadi-Javid et al. [14]	2017	Survey of studies	•		•	•	•	•
Arnolds, I., & Nickel, S. [25]	2015	Book Chapter	•			•	•	
Asada, Y., & Kephart, G. [26]	2007	Journal Article					•	
Bruno, G., & Giannikos, I. [27]	2015	Book Chapter	•		•	•		•
Burkey, M. L., Bhadury, J., & Eiselt, H. A. [28]	2011	Book Chapter			•			
Canadian Institute for Health Information [29]	2016	Government report					•	
Church, R. L. [30]	2002	Journal Article			•	•		
Crighton, E. J. et al. [6]	2005	Journal Article					•	•
Doctors Nova Scotia [8]	2017	Report					•	
Doctors Nova Scotia [31]	2019	Report					•	
Farahani, R. Z., SteadieSeifi, M., & Asgari, N. [32]	2010	Survey of Studies			•	•		•
Graber-Naidich, A. et al.[15]	2015	Journal Article	•		•	•	•	•
Graber-Naidich, A. et al. [33]	2017	Journal Article	•		•	•	•	
Guagliardo, M. F. [2]	2004	Journal Article			•		•	•
Güneş, E. D. et al. [4]	2014	Journal Article	•		•	•	•	•
Güneş, E. D., & Nickel, S. [9]	2015	Book Chapter	•		•	•	•	•
Kephart, G., & Asada, Y. [34]	2009	Journal Article					•	
Khan, A. A., & Bhardwaj, S. M. [10]	1994	Journal Article					•	•
Marchildon, G. P., & Hutchison, B. [35]	2016	Journal Article	•	•			•	
Marianov, V. [36]	2017	Journal Article	•		•	•	•	
Mcnamara, L., et al. [21]	2018	Working Paper	•		•	•	•	
Morgan, J. S., & Graber-Naidich, A. [3]	2019	Journal Article	•				•	•
Nova Scotia Department of Health and Wellness [7]	2016	Report					•	
O'Malley, A. S. et al. [5]	2012	Journal Article		•			•	•
Reuter-Oppermann, M. et al. [13]	2019	Journal article	•		•	•	•	•
Reuter-Oppermann, M. et al. [16]	2017	Book Chapter	•		•	•	•	•
Reuter-Oppermann, M. et al. [12]	n.d.	Working Paper	•	•	•	•	•	•
Sridharan, R. [37]	1995	Journal Article	•		•	•		•
Terashima, M. et al. [38]	2016	Report			•		•	
Tomblin Murphy, G. et al. [39]	2009	Journal Article					•	

Appendix B

p-Median Problem Class Code

This section contains the full pmed class code that was used.

```
class pmed:
    def __init__(self, data):
        self.facilities = data['p']
        self.csts = data['c']

    def defprob(self, d = None):
        self.tstart = time.time()
        p = self.facilities
        costs = self.csts

        if d == None:
            d = [1 for i in range(len(costs))]

        #Create list of nodes
        sites = [x for x in range(len(costs))]

        Routes = [(i,j) for i in sites for j in sites]

        route_vars = LpVariable.dicts("Route",(sites,sites),0,None)
        site_vars = LpVariable.dicts("Site",(sites),0,1,LpInteger)

        self.prob = LpProblem("pmedprob",LpMinimize)

        #Objective Function
        self.prob += lpSum([costs[i][j]*route_vars[i][j] for (i,j)
    in Routes]), "Sum of Costs"

        #Constraints
        for j in sites:
```

```
            self.prob += lpSum([route_vars[i][j] for (i) in sites]) == 1, "Sum of Demand"

        self.prob += lpSum([site_vars[i] for (i) in sites]) == p

        for j in sites:
            for i in sites:
                self.prob += route_vars[j][i] - site_vars[j] <= 0

        self.tdef = time.time() - self.tstart

    def solveprob(self, solver = None, solvepath = None):

        self.tstart = time.time()

        if solver == None:
            self.prob.solve()

        else:
            self.prob.solve(solver)

        self.tsolve = time.time() - self.tstart

    def soln(self):
        return value(self.prob.objective)

    def fullsoln(self):
        print("Status:", LpStatus[prob.status])
        for v in prob.variables():
            if v.varValue != 0:
                print(v.name, "=", v.varValue)
        print("Total Cost = ", value(prob.objective))

    def times(self):
        return self.tdef, self.tsolve
```

Listing B.1: pmed Code

Appendix C

Images Used in Subfigures

This section contains figures that were placed in subfigures. They are included in larger size here.

C.1 Figure 4.6 (Halifax Approaches Comparison)

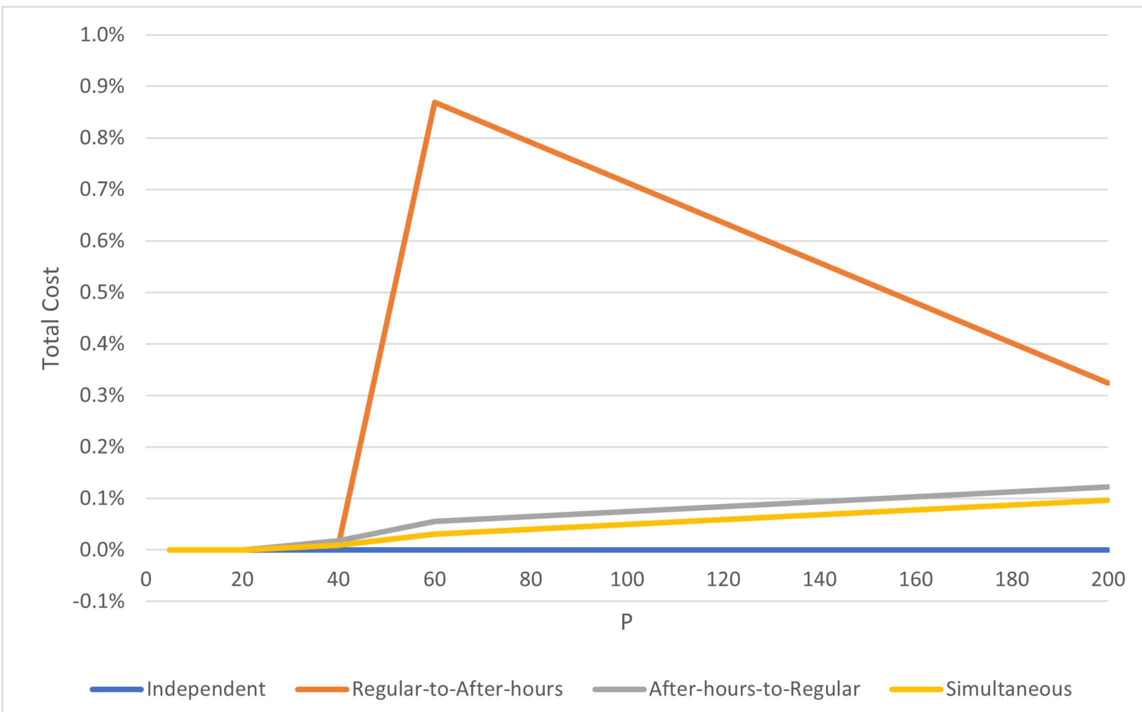


Figure C.1: Percent cost increase of each approach compared to independent ($p'/p = 0.2$)

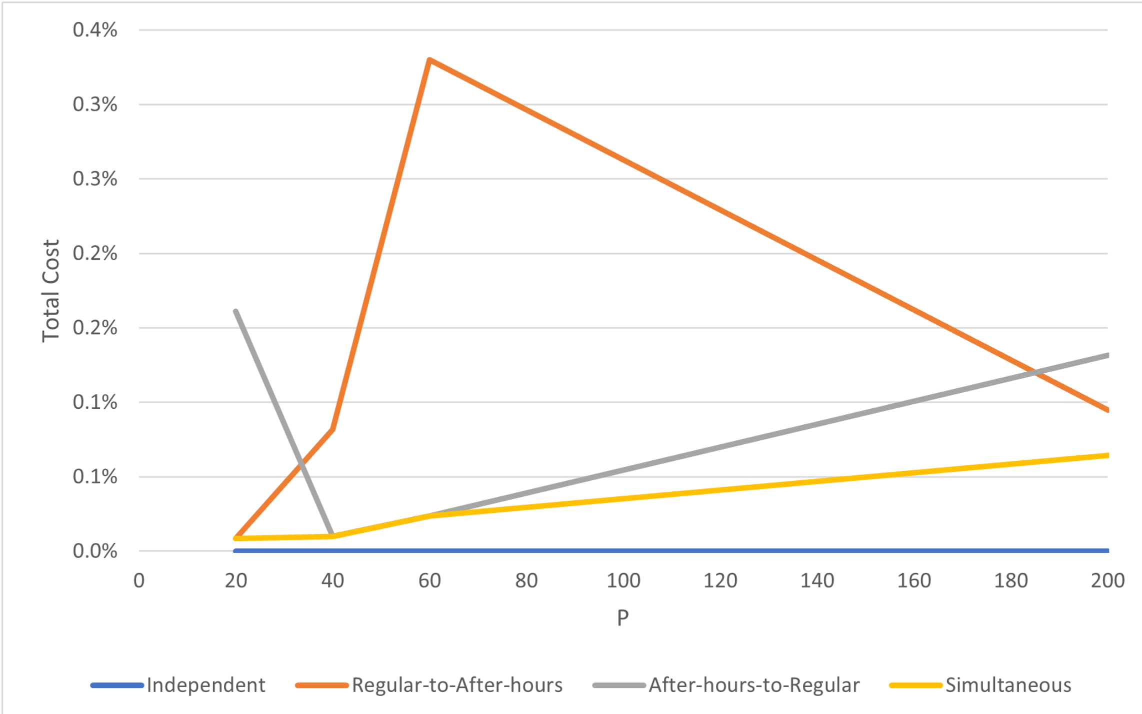


Figure C.2: Percent cost increase of each approach compared to independent ($p'/p = 0.4$)

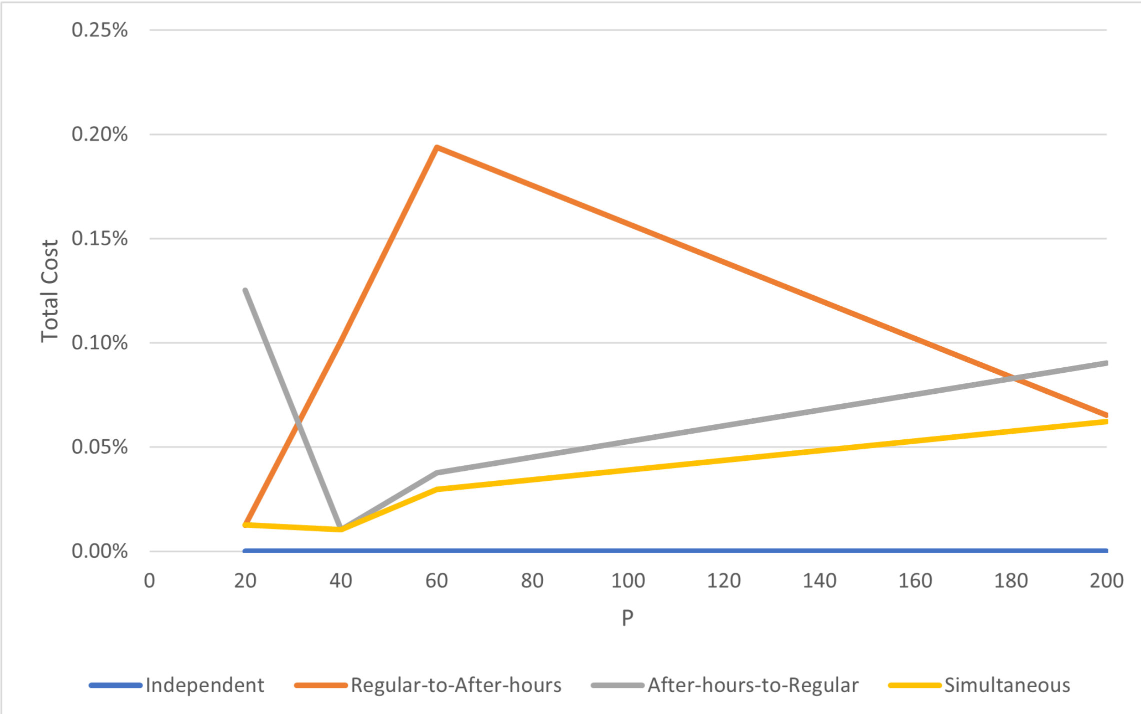


Figure C.3: Percent cost increase of each approach compared to independent ($p'/p = 0.6$)

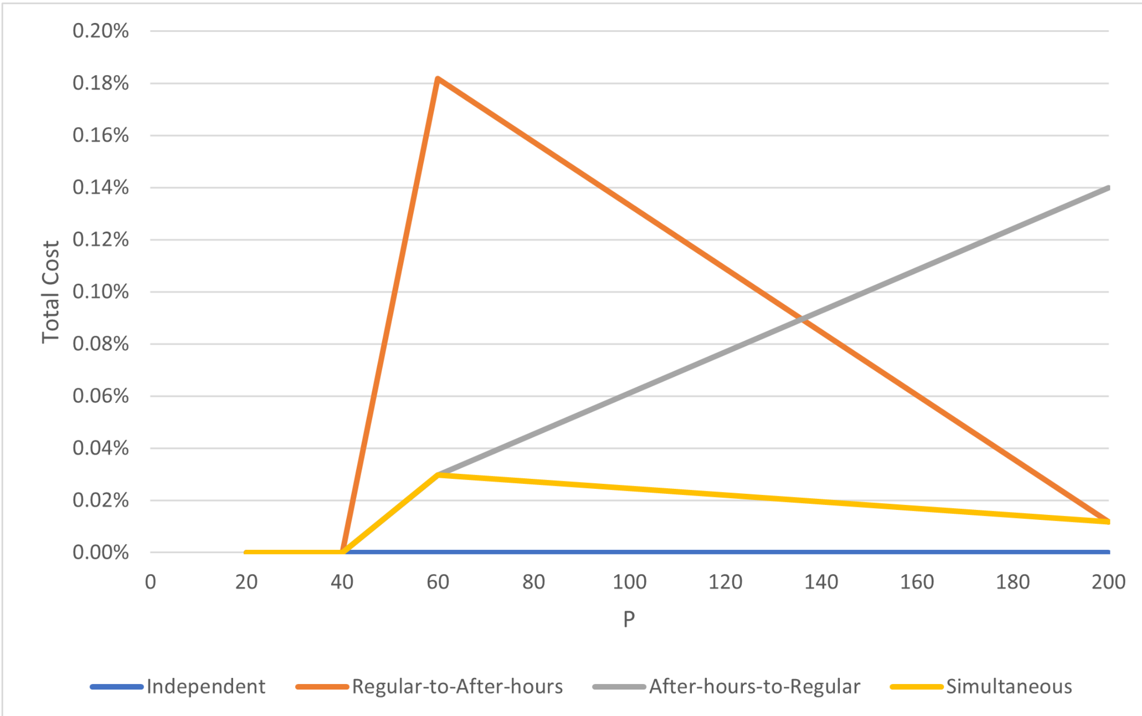


Figure C.4: Percent cost increase of each approach compared to independent ($p'/p = 0.8$)

C.2 Figure 4.9 (Cape Breton Approaches Comparison)

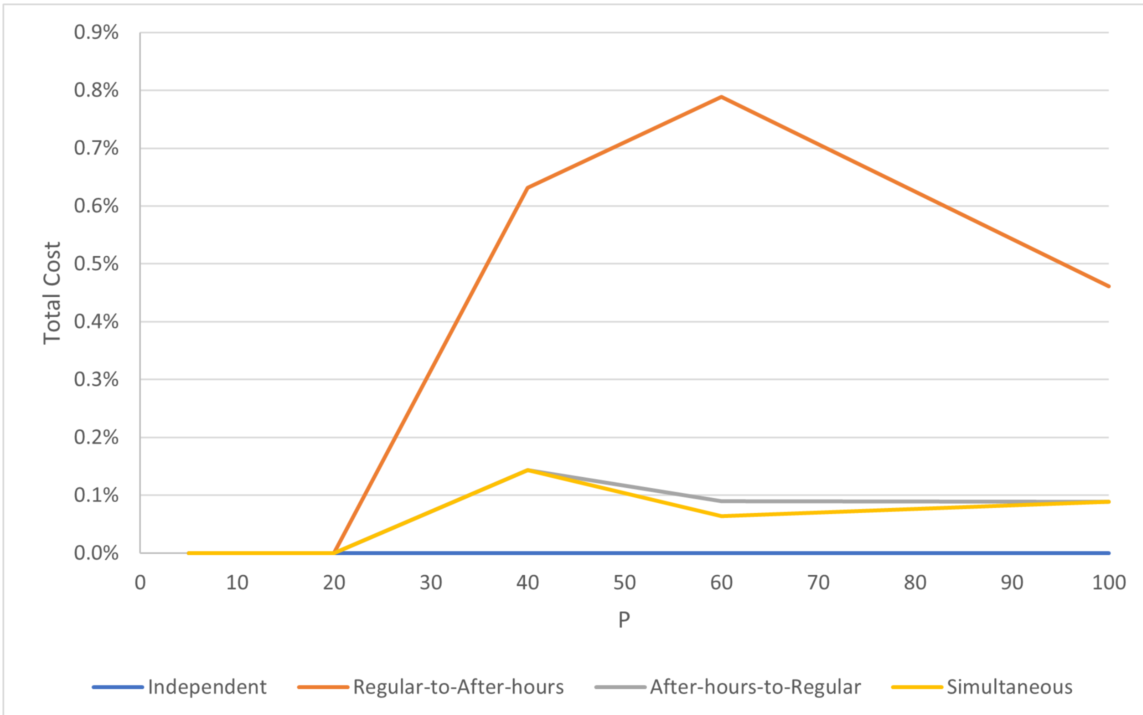


Figure C.5: Percent cost increase of each approach compared to independent ($p'/p = 0.2$)

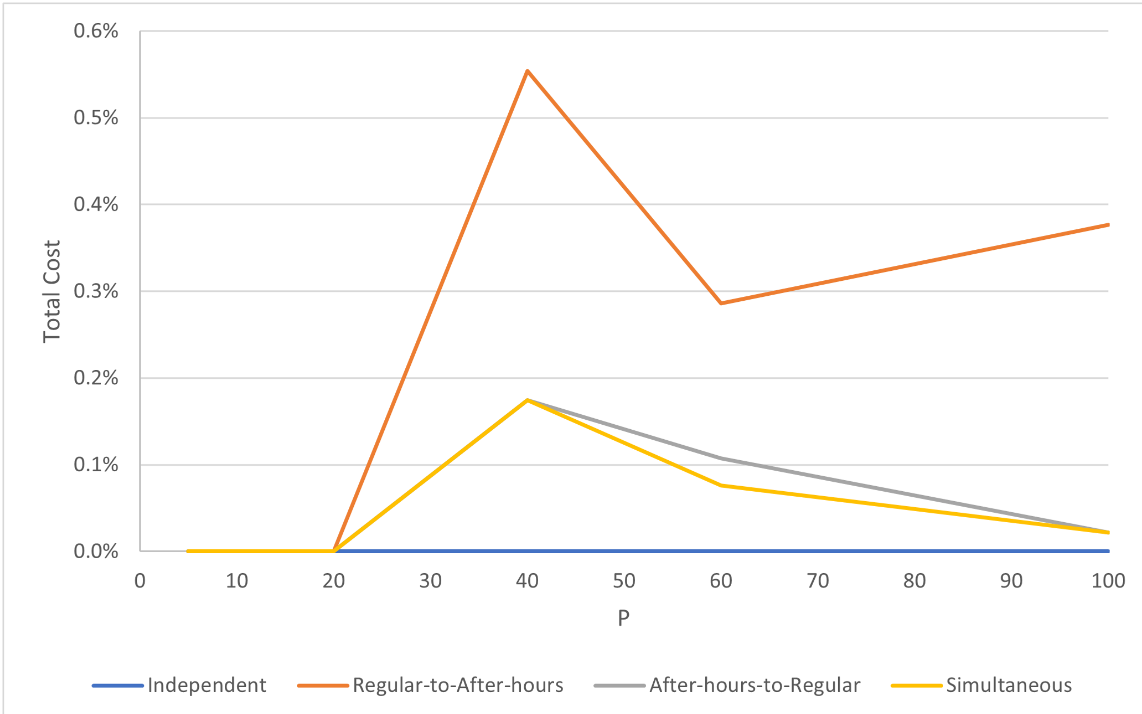


Figure C.6: Percent cost increase of each approach compared to independent ($p'/p = 0.4$)

Figure C.7: Percent cost increase of each approach compared to independent ($p'/p = 0.6$)

Figure C.8: Percent cost increase of each approach compared to independent ($p'/p = 0.8$)

C.3 Figure 4.12 (Cape Breton Modified Approaches Comparison)

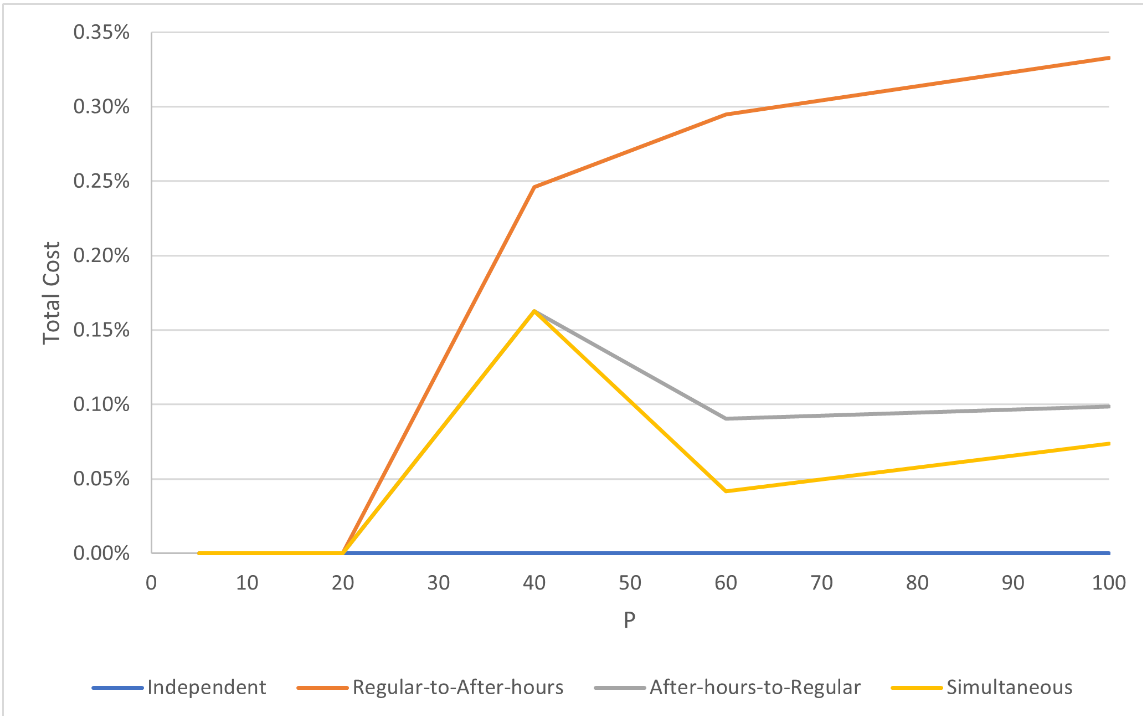


Figure C.9: Percent cost increase of each approach compared to independent ($p'/p = 0.2$)

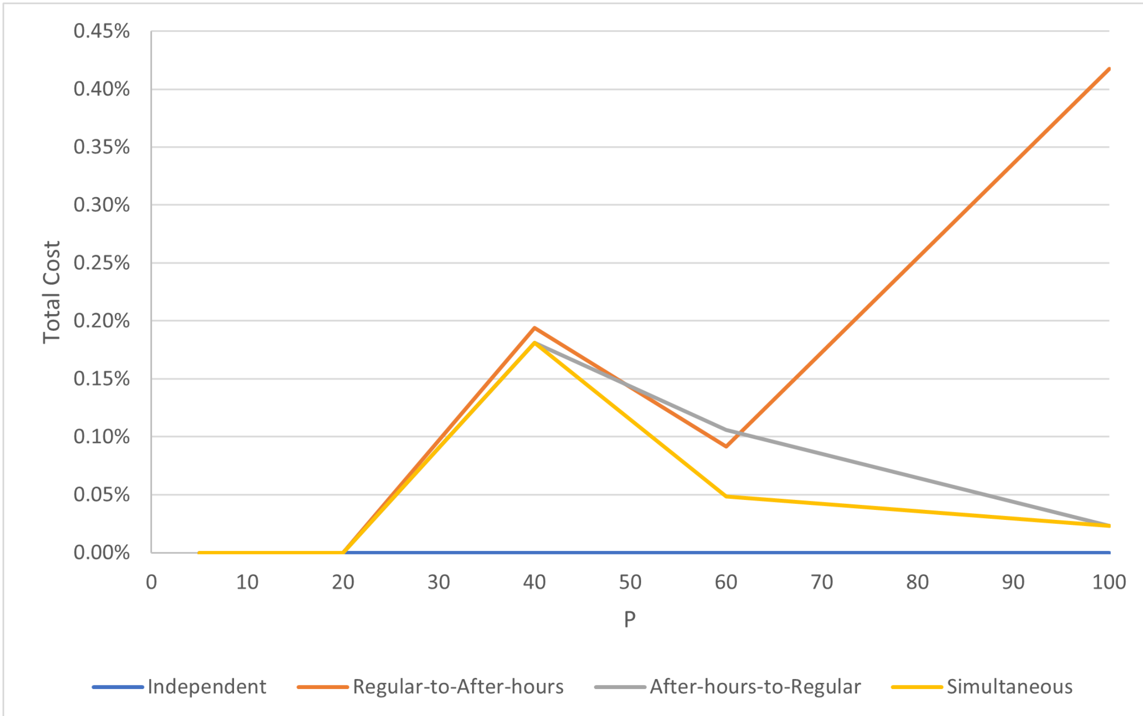


Figure C.10: Percent cost increase of each approach compared to independent (p'/p = 0.4)

Figure C.11: Percent cost increase of each approach compared to independent (p'/p = 0.6)

Figure C.12: Percent cost increase of each approach compared to independent (p'/p = 0.8)