August 2018

Stochastic Optimization Models for Perishable Products

Zahra Azadi
Clemson University, zazadi@clemson.edu

Stochastic Optimization Models for Perishable Products

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Industrial Engineering

---

by
Zahra Azadi
August 2018

Accepted by:
Dr. Sandra Eksioglu, Committee Chair
Dr. Burak Eksioglu
Dr. David Neyens
Dr. Kevin Taaffe
Dr. Khoa Truong

Abstract

For many years, researchers have focused on developing optimization models to design and manage supply chains. These models have helped companies in different industries to minimize costs, maximize performance while balancing their social and environmental impacts. There is an increasing interest in developing models which optimize supply chain decisions of perishable products. This is mainly because many of the products we use today are perishable, managing their inventory is challenging due to their short shelf life, and out-dated products become waste. Therefore, these supply chain decisions impact profitability and sustainability of companies and the quality of the environment. Perishable products wastage is inevitable when demand is not known beforehand. A number of models in the literature use simulation and probabilistic models to capture supply chain uncertainties. However, when demand distribution cannot be described using standard distributions, probabilistic models are not effective. In this case, using stochastic optimization methods is preferred over obtaining approximate inventory management policies through simulation.

This dissertation proposes models to help businesses and non-profit organizations make inventory replenishment, pricing and transportation decisions that improve the performance of their system. These models focus on perishable products which either deteriorate over time or have a fixed shelf life. The demand and/or supply for these products and/or, the remaining shelf life are stochastic. Stochastic optimization models, including a two-stage stochastic mixed integer linear program, a two-stage stochastic mixed integer non linear program, and a chance constraint program are proposed to capture uncertainties. The objective is to minimize the total replenishment costs which impact profits and service rate. These models are motivated by applications in the vaccine distribution supply chain, and other supply chains used to distribute perishable products.

This dissertation also focuses on developing solution algorithms to solve the proposed optimization models. The computational complexity of these models motivated the development of

extensions to standard models used to solve stochastic optimization problems. These algorithms use sample average approximation (SAA) to represent uncertainty. The algorithms proposed are extensions of the stochastic Benders decomposition algorithm, the L-shaped method (LS). These extensions use Gomory mixed integer cuts, mixed-integer rounding cuts, and piecewise linear relaxation of bilinear terms. These extensions lead to the development of linear approximations of the models developed. Computational results reveal that the solution approach presented here outperforms the standard LS method.

Finally, this dissertation develops case studies using real-life data from the Demographic Health Surveys in Niger and Bangladesh to build predictive models to meet requirements for various childhood immunization vaccines. The results of this study provide support tools for policymakers to design vaccine distribution networks.

Dedication

This dissertation is dedicated to the memory of my grandfather Ibrahim Mansouri, who loved to see this day. I also dedicate this to my parents, Morteza Azadi and Kobra Mansouri, and my siblings, Mona and Amir.

Acknowledgments

I would like to express my sincere appreciation to my advisor, Dr. Sandra D. Eksioglu, for her continuous support, guidance and patience throughout my studies. Under her supervision, I learned how to define a research problem, find a solution, and publish the results. On a personal level, she inspired me by her hardworking and passionate attitude. I am indebted to her for being such a great friend and a role model.

I would like to extend my gratitude to Drs. Burak Eksioglu, David Neyens, Kevin Taaffe, and Khoa Truong for being a part of my committee and providing invaluable comments for my dissertation. My sincere thanks go to Dr. Harsha Gangammanavar for his time and ideas on my research. I appreciate his working closely and sharing with me his experience in the stochastic optimization field. I am grateful to department head Dr. Cole Smith for providing me the opportunity to teach a class.

I am very thankful to my best friend Aida, for always being there and making me feel happy. I also would like to verbalize my appreciation to all of my sincere friends at the IE Department and in Clemson.

There are no words that can express my thanks to my parents, Morteza Azadi and Kobra Mansouri, for their continuous love, motivation, and encouragement. I would have never made this journey without their support and sacrifice. A heartfelt appreciation is for my sister, Mona, and her husband, Hamed during their four years stay with me at Tehran, they did give an enormous moral support for my undergraduate studies. Thanks to my brother, Amir, for his unconditional love and patience.

Last, but most importantly, I would like to express my deepest gratitude to my husband, Ershad for his endless support, continued patience, and warm love. He has always been around at times I thought that it is impossible to continue.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Inefficient inventory management of perishable products negatively impacts companies economic prosperity. A new study reveals that about $5 billion worth of drugs are thrown away in unopened packs due to expiration date [77]. Economic Research Service (ERS) of United States Department of Agriculture (USDA) estimates that, in 2010, 45 billion pounds of available food at retail stores in the United States was wasted [25]. This amount of wastage sums to $41.9 million [138]. The Secretary of Agriculture and the Deputy Administrator of Environmental Protection Agency announced that the food waste in United States, should decrease by 2030 [4]. Poor inventory management of perishable products contributes to this waste. Waste negatively impacts the environmental and is one of the sources of social problems like hunger. In the United States, food waste accounts for 10% of the energy supplied, 80% of water consumed, and 50% of farm land used [56]. Disposed pharmaceutical wastage often reaches the sewage system. This practice is dangerous since pharmaceutical products contain active chemicals which impact the environment, animal and human health. The Associated Press reported that, in 2008, drinking water used by 41 million people in United States was contaminated by pharmaceutical residues [44]. Perishable product wastage is also critical in developing countries. For example, vaccine wastage negatively impacts immunization coverage by decreasing availability. Wastage occurs when a vaccine vial is physically damaged or exposed to extreme temperatures, or when doses from an open vial are discarded after their safe use time expires. The latter is referred to as open vial wastage (OVW). Clinics can use single-dose vials to reduce OVW; however, such an approach is more expensive than using multi-dose vials.

For decades, optimization models have been developed to design and manage supply chains.

These models have helped companies in different industries to minimize costs, maximize performance while balancing their social and environmental impacts. There is an increasing interest in developing models to support supply chain-related decisions for perishable products. Perishables are products which either become obsolete within a fixed period of time, or loose value with time. The first group of products is referred to in the literature as "perishable products with fixed shelf life", and the second group is referred to as "deteriorating products". Some examples of perishable products with fixed shelf life include dairy, pharmaceuticals, and blood. Electronics, foods, and groceries are examples of deteriorating products. Perishable products are an essential part of our life, and their out-dating is a great threat to the profitability and sustainability of companies, and quality of the environment. The statistics about the large amounts of waste generated indicate that companies are not doing a great job in managing perishable inventories. Based on our review of literature, there is a limited number of models which companies can use to manage the inventory of perishable products. However, these models do not consider integrating inventory replenishment and markdown decisions, as well as, trade-offs between the cost of storage and transportation of perishable products when demand and/or supply is stochastic. In order to fill this gap in the literature, we propose models and solution algorithms which capture the specific characteristics of perishable inventories. The challenges associated with decision making about inventory management of perishable products, both for businesses and non-profit organizations, can be categorized as follows:

- *Transportation and storage of perishable products.* There are trade-offs between the cost of storage and transportation of perishable products and the remaining product shelf life. These decisions impact costs since excess supply leads to spoilage, but shortage results in lost sales. Questions such as, which supplier(s) to select, how much to order, and how frequently to replenish the inventory need to be addressed in order to optimize the system performance.

- *Price markdowns.* The importance of price determination for perishable products is highlighted in many studies such as [32] and [40]. The goal is to price products in such a way that the profits of an organization are maximized. This is not an easy task since price impacts demand. Markdown policies are used to stimulate demand for perishable products by decreasing their selling price. This leads to decrease in waste and increase in cost savings. Unit selling price of a product, and the timing of markdown decisions are important and need to be addressed.

- *Uncertainties in demand.* Perishable products wastage is inevitable when demand is not known beforehand. Therefore, deterministic inventory models which do not take uncertainty into account do not find robust solutions. Although there are studies that consider demand uncertainties, for example [36], most of these studies obtain optimal policies through approximation methods or simulation. When demand uncertainties cannot be represented using standard distributions, probabilistic models are not effective. Developing stochastic models and exact solution approaches is a challenging task which needs to be addressed.

- *Uncertainties in supply.* Replenishment decisions for perishable products are further complicated by uncertainties in the amount and quality of shipments received from suppliers. Dual-sourcing is a policy often used by retailers to mitigate the risks of supply uncertainties [84]. In this setting, usually, one of the suppliers is reliable but expensive. The other supplier is less reliable and less expensive. Supplier reliability is impacted by his limiting and varying capacity. In this regard, considering the trade-offs between supply chain costs and reliability is important and needs to be addressed.

In this dissertation, we use operations research (OR) tools to develop stochastic optimization models and solution algorithms in order to address the above mentioned challenges. These models can enable companies and organizations which manage perishable products to identify policies which allow them to fulfill customer demand at a minimum cost. The proposed models capture the impact of demand uncertainty on decision making. We propose stochastic optimization models, including a two-stage stochastic mixed integer linear program, a two-stage stochastic mixed integer non linear program, and a chance constraint program to capture uncertainties. To solve the two-stage stochastic programs and analyze the results, we develop different solution frameworks based on well-known algorithms such as stochastic Benders decomposition. We use the SAA method to solve the chance constrained program.

In Chapter 2, we study the vaccine vial replenishment problem. Vaccines are perishable products which have a fixed shelf-life. In the last century, many infectious diseases have been completely eradicated or significantly reduced because of childhood vaccinations. Ample evidence suggests that low vaccination coverage in developing countries is caused by vaccine stock out and high vaccine wastage. Wastage occurs when a vaccine vial is physically damaged or exposed to extreme temperatures, or as OVW. Clinics can use single-dose vials to reduce OVW; however, such an approach is more expensive than using multi-dose vials. The focus of this chapter is to develop new policies that support vaccine administration and inventory replenishment. These policies are expected to reduce OVW, reduce the cost of vaccinations and improve vaccination coverage levels in developing countries. We propose a two-stage stochastic programming model that identifies an optimal mix of different-sized vaccine vials, and the corresponding decisions that clinics make about opening vials in the face of uncertain patient arrivals. This work develops a case study with data gathered from Bangladesh. Experimental results indicate that using a mix of vials of different sizes reduces OVW, compared to the current practice of using single-sized multi-dose vials. Experimental results also point to simple and economic vaccine administration policies that health care administrators can use to minimize OVW. The model is solved using an extension of the stochastic Benders decomposition algorithm, the LS method. This algorithm uses GMI and MIR cuts to address the problem's non-convexity. Computational results reveal that the solution approach presented here outperforms the standard LS method.

In Chapter 3, we analyze the joint pricing and inventory management for deteriorating products. This research is motivated by the opportunities we see to reduce waste and increase profitability of perishables in retailing using pricing. We propose a two-stage stochastic optimization model that selects suppliers; identifies a replenishment schedule for a periodic-review inventory system with non-stationary demand and supply; and identifies the timing and size of a price markdown in order to maximize retailer's profits. In this model, the first-stage problem is bilinear since it captures the additive relationship between price and demand. Thus, we develop a solution approach which extends Benders decomposition algorithm via a piecewise linear approximation method to solve the first-stage problem. We develop a case study to validate the model. Numerical experiments point to the benefits of integrating inventory management and pricing decision in the supply chain.

Effective distribution of vaccines from manufacturers to clinics is challenging. One of the main challenges is the stochastic nature of demand for vaccines. In many developing countries the demand for vaccine is growing exponentially because of an increasing birth rate. This demand is met via supplies from international organizations that coordinate shipments from multiple different resources. This complicates vaccine distribution planning. An efficient distribution plan requires an accurate demand forecast. In Chapter 4 we build predictive models to meet requirements for various childhood immunization vaccines (CIV). Data from the Demographic Health Surveys in Niger and Niger Census are used in regression models to predict monthly demand for CIV per region. The population size, the percentage of population under poverty line, the adult literacy rate, and the number of clinics were selected as independent variables. The results suggest that, vaccine type, as well as, the social and economic characteristics of a region impact the demand for vaccination, and should be considered in vaccine distribution planning. The results of this study provide support tools for policymakers to design vaccine distribution networks.

In Chapter 5, we propose a chance constraint programming model which identifies optimal vaccine supply chain designs and management strategies. This is a data-driven model built upon the regression models in Chapter 4. The proposed model considers the limited shelf life of vaccines, facility and transportation storage capacities, as well as variations in patient arrivals at health clinics. The SAA method is used to approximate the chance constraints. We use the model to analyze different supply chain network designs and vaccine administration policies for vaccine vial distribution in Niger. The existing distribution network has multiple layers. Decision makers in each layer identify inventory levels and timing of replenishment. Removing a level in the hierarchy of the network impacts on the system-wide costs and reduces the chance of vial breakage due to fewer times vaccines are touched. Moreover, when a single-dose vaccine is replaced with a multi-dose vaccine, the OVW rate increases. In this regard, we evaluate the impact of converting the current Niger's four-tier vaccine supply chain to a three-tier one by removing the regional stores, and changing Measles's vial size from multi-dose to single-dose on vaccine availability at clinics and immunization coverage.

Finally, in Chapter 6, we summarize our observations and provide concluding remarks.

Chapter 2

Developing Childhood Vaccine Administration and Inventory Replenishment Policies that Minimize Open Vial Wastage

2.1 Introduction

The spread of infectious diseases has significantly receded over the last century, and these widespread immunities can largely be attributed to childhood vaccinations [117]. Childhood vaccinations have helped eradicate diseases, like smallpox, and have severely restricted diseases, such as polio, measles, and tetanus. Immunization programs, led by World Health Organization (WHO) and United Nations Children's Emergency Fund (UNICEF), have been instrumental in delivering global eradication efforts. However, achieving immunization targets in developing countries has been particularly challenging. As recently as 2014, 60% of the 18.7 million children worldwide who did not receive the diphtheria-tetanus-pertussis (DTP) lived in ten developing countries [3]. Vaccine stockout is one of the primary reasons for the low vaccination coverage in these places. In 2014, 26% of WHO-member countries experienced a national level vaccine shortage for at least one month

[127]. Stockout occurs, in part, because of vaccine wastage. WHO estimated that, more than half of the total vaccination supply distributed around the world is wasted [1]. Wastage occurs when a vaccine vial is physically damaged or exposed to extreme temperatures, or when doses from an open vial are discarded after their safe use time expires. The latter is referred to as OVW. The Global Alliance for Vaccines and Immunizations (GAVI) requests that countries take measures to bring down their vaccine wastage rates [1], and lowering OVW rates can be one effective approach.

In many developing countries, immunizations are administered in health care centers or at outreach sessions. For example, in Bangladesh, 94% of immunizations are accomplished at outreach sessions [55]. These sessions are organized through programs, such as Expanded Programme on Immunization (EPI), and delivered by trained nurses. Health care centers are often located in remote areas that lack proper transportation and basic amenities, including refrigeration.Childhood vaccines, including bacille Calmette-Guérin (BCG), measles, DTP, and tentanus (TT), are typically produced in multi-dose vials. Outreach sessions conducted by EPI follow a strict open vial policy, which requires that, opened vials of multi-dose vaccines be discarded at the end of the day or six hours, whichever comes first [145]. Discarded doses contribute to OVW. While single-dose vials have zero OVW, they are more expensive than multi-dose vials because of additional costs for packaging, holding, and transportation. A 2010 study in Bangladesh estimated OVW for BCG at 85%, measles at 71%, DTP at 44.2%, and TT at 36% [55].

Organizations managing immunization programs in developing countries do make vaccine replenishment decisions for a single vial size [43, 55]. Vaccines are then distributed to clinics based on a fixed replenishment schedule [52]. Intuition tells us that, allowing clinics to use a mix of vials of different sizes should result in OVW reductions. Better inventory management should result when clinics decide their own replenishment schedules based on observed patient arrivals, and when clinics have easy-to-implement vaccine administration policies.

In this study, "inventory replenishment policy" refers to the mix of multi-dose vaccine vials, e.g., single-, two-, and ten-dose vials, that a clinic should order to reduce OVW during outreach sessions and maintain low vaccination costs. "Vaccine administration policy" refers to the conditions under which a nurse should open a new vial of a particular size during an outreach session. Such a policy accounts for session duration, the number of patients present, and/or the time of day when a multi-dose vial of a particular size should be opened. In the absence of such policies, deciding whether to open or not to open a multi-dose vial vaccine is challenging due to an uncertain number of patient arrivals. This study develops a two-stage stochastic programming model to identify the optimal mix of multi-dose vaccine vials to order, the timing of orders and the corresponding vial opening decisions in the face of uncertain patient arrivals. The goal is to minimize OVW and achieve lower vaccination costs in developing countries. The models proposed here are tested via an extensive numerical analysis using data from Bangladesh. Numerical results inspired the development of a number of vaccine administration policies which, as demonstrated in this study, can reduce OVW.

2.1.1 Literature Review

In developing countries, EPI vaccines are purchased by governmental and non-profit agencies to satisfy annual vaccine requirements. In these countries, EPI delivers vaccines to a central store from which they are shipped to regional stores, then to district stores, and, finally, to local clinics. The number of downstream tiers, or supply chain stages, which handle vaccines varies by country. Some studies in the literature focus on national-level decisions made about how many vaccines should be purchased to ensure high vaccine coverage levels [34]. Other studies focus on the structure of the corresponding supply chain to minimize immunization costs [14, 13, 24]. A few studies focus on the decisions made about clinic-level vaccine administration that can minimize OVW [97, 96, 43]. Similarly, this work examines clinic-level inventory replenishment decisions and vaccine administration policies with the goal of minimizing OVW and inventory replenishment costs.

The literature on vaccine supply chain management has increased in recent years. Research focused on OVW management started as vaccine vial wastage increased in outreach immunization sessions held under programs such as EPI in developing countries. The work conducted by [45] was the first that discussed the general differences between single- and ten-dose vaccine vials in terms of cost, distribution, coverage, and safety. A survey of vaccine wastage in Bangladesh, provided in [55], detailed the usage and wastage rates of different-sized vaccine types in randomly selected outreach and non-outreach sessions. The aims of both studies were to estimate vaccine wastage, but neither considered the associated supply chain costs. However, these costs were considered in the analysis conducted by [115]. The scope of these studies was restricted to analyzing existing vaccine administration practices, and, hence, they do not prescribe any plan for efficient inventory management. However, they recognize that "there is an urgent need for more rigorous and systematic wastage monitoring" [115].

The following studies have used simulation-based approaches to assess the impacts of different-sized multi-dose vials on OVW. [78] proposes a simulation model for Thailand's Trang Province that evaluates the cost of replacing ten-dose with single-dose vial vaccines. The results of this study show that, given a fixed patient arrival rate, the cost associated with medical waste disposal is greater than the cost savings resulting from OVW reduction. This finding, followed their earlier simulation work that models patient arrivals using a Poisson distribution [80]. The authors conclude that the suitable number of doses per vial may vary by region and patient arrival rate. However, the Poisson arrival assumption may not necessarily capture the arrival process in outreach immunization sessions due to the particular challenges of organizing these sessions in particular communities. A similar simulation study was carried out by [147] to analyze the impact of different patient arrival rates on OVW for five- and ten-dose vaccine vials. In this case, the authors use real-life data to estimate parameters for simulation. These studies examine the current practice of using a single multi-dose vial, and they recommend using a particular vial size based on patients arrival rates. These recommendations are based on simulated cost estimates, but once again, these works do not prescribe any vaccine inventory management schemes.

In order to prescribe a vaccine management plan, optimization models are necessary. Such approaches are limited and include deterministic mixed-integer programs (MIP) and finite state probabilistic models. In [43], the total vaccination cost is minimized via optimal ordering decisions which are identified by solving a deterministic MIP integrated within a Monte Carlo simulation setup. The authors used the simulation model to generate patient arrivals and compute OVW under a fixed vaccine administration policy and for a given multi-dose vaccine vial. The estimated OVW is then used in the optimization model to determine the optimal ordering decisions. [97] proposed a Markov Decision Process (MDP) model to identify a vial-opening policy which minimizes expected OVW cost and maximizes vaccine coverage over a finite horizon. The authors used a backward induction algorithm to solve their probabilistic model. This work is further extended in [96] as the MDP model is integrated with simulation to analyze how session duration impacts the optimal policy obtained in their previous work. In both studies, the authors determine the clinic closing time or the time when patients should be rejected. However, they do not consider loss in opportunity due to unserved patients. Moreover, they assume that an infinite amount of vaccines can be replenished, which is not always the case at outreach sessions. These studies address the inventory management of a single multi-dose vaccine vial, and patient arrival is characterized either probabilistically or through simulation. The optimization models focus either on ordering decisions or vial-opening decisions.

However, these decisions are interrelated, and, therefore, a model which considers both and accounts for uncertainty in patient arrivals should be very effective in reducing OVW costs and increasing immunization rates.

Vaccines are perishable products that expire after their safe-use time. Thus, the study presented here is related to the literature on inventory replenishment models for perishable products with fixed shelf-life. Works by [101], [119], and [17] model these problems by extending the classical economic lot sizing (ELS) problem formulation. ELS and its variations are often formulated using extensions of the minimum cost network flow model [113]. While some ELS models for perishable products with fixed shelf-life consider demand to be deterministic, others have considered stochastic demand; [17] surveys this work. [135] and [103] find optimal ordering policies, when demand follows continuous and differentiable distributions, by incorporating shortage and ordering costs in the cost function. [103] also account for wastage costs due to expiry in their cost function. An extended m-period dynamic programming model is presented both in [51] and [100]. The ELS for perishable products with a fixed shelf-life is solved with well-known ordering policies, such as (s, S) and (Q, r). A Markov renewal process is used to obtain an ordering policy in [86], which uses a closed-form cost function for an (s, S) policy with back-orders. [140] shows the optimality of (s, S) policy when the demand process is compound Poisson. In these studies, the inventory replenishment model is restricted to follow a specific policy that is optimal only under certain demand distributions and cost functions. For example, in [140], the proposed inventory ordering policy is optimal only if the demand is a compound Poisson process.

Using exact methods to derive optimal policies for stochastic ELS models and their extensions is computationally complex which has resulted in the development of many heuristic policies. For example, [104] develop a myopic-based inventory policy using bounds on wastage costs. Following this approximation method, [36] develop two heuristic policies for a joint inventory control and pricing model, but these policies perform well only when demand is non-stationary. Simulation-based methods have also been employed to obtain approximate policies; see [133]. A combination of dynamic programming and simulation is used to reduce the state space of the stochastic ELS model in [59]. However, optimization models that provide exact solutions to the ELS problem and its extensions have been limited to deterministic models; see [47]. Stochastic demand is handled through optimal policies only when highly simplifying assumptions have been made, or through sub-optimal heuristic policies under slightly general settings [86, 104, 140, 133].

2.1.2 Our contributions

When stochastic patient arrivals cannot be described using standard distributions, probabilistic models are not effective. In this case, using optimization methods is preferred over obtaining approximate policies through simulation. Moreover, when vaccines are distributed via a mix of multiple-dose vials, inventory replenishment and vaccine administration decisions, estimating inventory states, and their interdependencies become complicated. Such realistic requirements can be captured using stochastic programming (SP) models. Furthermore, SP provides effective tools to design algorithms that can handle these comprehensive models. In this regard, the contributions of our study are as follows:

- *Vaccine administration and inventory replenishment model:* We propose a two-stage SP model for vaccine vial replenishment which captures (a) the order frequency for respective quantities of different-sized vials, (b) the opening schedule for these vials, and (c) the administration of available doses to patients. Experimental results indicate that, using a mix of different-sized vials compared to using vials of a single size, results in monetary savings, better service, and OVW reductions at outreach sessions.

- *Vaccine administration policies:* Motivated by the solutions obtained from our model, we propose simple and economic vaccine administration policies for outreach sessions. These policies depend on the population size and birth rate in the regions served. This study compares their effectiveness by replicating a simulation process.

- *Scalable algorithm:* We develop a new solution approach for two-stage stochastic integer programs (2-SIP) with continuous recourse. Our algorithm is motivated by the well-known stochastic Benders method; uses GMI and MIR cuts to address the non-convexity of the first-stage problem. Our approach significantly reduces the computational requirement compared to the standard method. Such computational enhancements allow our algorithm to be applied to real problems.

In the following sections, the SP model developed and tested. The model is presented in §2.2, and the solution approach is discussed in §2.3. . Data and computational results are presented in §2.4, and ideas regarding extensions of the current model and the solution approach are presented in §2.5.

2.2 Problem Formulation

2.2.1 Problem description and assumptions

This section details a model that aids the design and management of immunization programs conducted through outreach sessions in developing countries. The structure of vaccine distribution system differs by country. Typically, clinics replenish inventories via shipments from a regional distribution center. Clinics run two types of immunization sessions: drop-in or outreach. In drop-in sessions, individuals visit the clinic to get vaccinated. In outreach sessions, health workers travel to provide vaccinations in rural communities during temporary gatherings [43]. For example, drop-in sessions account for 6%, and outreach sessions account for 94% of immunizations in Bangladesh [55]. This study focuses on short outreach sessions, usually lasting a few weeks, that target the infant population in rural areas. Assume that patient arrivals in outreach sessions are random. Due to the stochastic nature of patient arrivals and the open vial policy, identifying replenishment schedules and administration policies is critical.

The model presented below is based on the following assumptions: (i) Vaccines are distributed in multi-dose vials of different sizes. (ii) Vaccine purchase cost is proportional to its size because of additional packing costs. (iii) Clinics maintain inventories of multi-dose vials of different sizes and inventories of doses from opened vials. (iv) OVW and storage cost the clinic.

Transportation lead time for vaccine replenishment is constant since deliveries are received from the same distribution center. Thus, we can easily transform this into a problem with zero lead time as follows: (a) the order for a replenishment that is to arrive in time t is submitted in time $t-t'$, where t' is the lead time; and (b) the unit procurement cost is increased by the unit inventory cost during transit time. Thus, the remainder of this study assumes negligible lead times.

Figure 5.1 provides a schematic representation of the modeling approach used for this problem. The darker shaded area represents inventory replenishment decisions for multi-dose vials during the planning horizon. These decisions are made weekly. Each box represents the type of vials in the inventory over a particular week. Boxes in successive time periods are connected to represent the inventory flow. The lighter shaded area represents vaccine administration decisions. These decisions are made daily. Each box shows when a vial is opened and how the doses are used during a given day. Boxes in successive time periods are not connected since unused doses are discarded, which contribute to OVW.

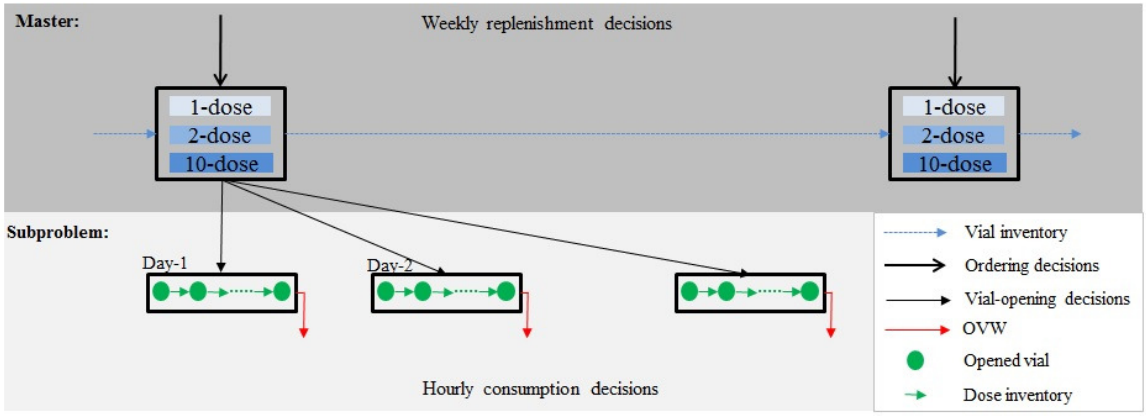


Figure 2.1: Inventory dynamics in vaccine administration and inventory replenishment model.

2.2.2 Vaccine administration and inventory replenishment model

For the model formulation presented here, note the use of lower case for vector and matrix entries and that bold font represents the entire vector or matrix.

Our model considers two major decisions: (a) inventory replenishment or ordering decisions, and (b) vial administration or opening decisions. Ordering decisions help to identify the optimal replenishment schedule, order mix and vial sizes. These decisions are made in a weekly basis. On the other hand, opening decisions are a function of the vaccine administration policy which depends on patient arrivals, and thus, these decisions are made at a much finer timescale: on an hourly basis. Here $\mathcal{T} = 1, \dots, T$ denotes the ordering decision epochs, and each ordering period consists of N opening decision epochs. Therefore, $\mathcal{N} = 0, \dots, NT$ denotes all the opening decision epochs over the planning horizon.

We first describe modeling of the inventory replenishment decisions. Let $\mathcal{V}$ represent the set of multi-dose vial sizes, e.g., one-dose, five-dose and ten-dose vials, available for order. Once the vials are opened they must be used within their safe use time, and τ denotes this limit. At $t \in \mathcal{T}$, ordering decisions z_t are made at fixed cost f_t. If an order is placed ($z_t = 1$), then the replenishment quantity for each multi-dose vial size can be determined. These decisions are represented by $r_{\nu t}$, and the corresponding variable purchase cost by $c_{\nu t}, \forall \nu \in \mathcal{V}$. At each opening decision epoch $n \in \mathcal{N}$, vial-opening decisions are made. Let $u_{\nu n}$ represent the number of vials of size ν opened. Replenishment quantity decisions $r_{\nu t}$ and vial-opening decisions $u_{\nu n}$ together determine the state of vial inventory, represented here as $s_{\nu n}, \forall \nu \in \mathcal{V}$. Let $d_{\nu t}$ represent the corresponding unit inventory holding cost. The evolution of inventory for each vial size $\nu \in \mathcal{V}$ is captured by the following flow-balance equations:

$$s_{\nu Nt} = s_{\nu(Nt-1)} + r_{\nu t} - u_{\nu Nt} \qquad \forall t \in \mathcal{T}, \tag{2.1a}$$

$$s_{\nu n} = s_{\nu n-1} - u_{\nu n} \qquad \forall n \in \mathcal{N} \setminus \{N, 2N, \ldots, TN\}, \tag{2.1b}$$

The initial inventory $s_{\nu 0}$ is assumed to be known. Since the inventory of vials is replenished weekly, equation (2.1a) is used to represent the corresponding inventory balance constraints. However, since dose administration decisions are made at a finer time scale, and those decisions impact the inventory of vials, equation (2.1b) is used to capture inventory balance during vaccine administration time periods. Since open vial policy does not apply to unopened vials, loss of inventories is not considered in these constraints.

$$\sum_{\nu \in \mathcal{V}} r_{\nu t} \leq M_t z_t \qquad \forall t \in \mathcal{T}. \tag{2.2}$$

Let M_t be an upper bound (UB) on the quantity of vials ordered in a time period. Constraints (2.2) ensure that the inventory of vials is replenished only if an order is placed in period t. Note that, decisions about vial replenishment and inventory ($\mathbf{z}, \mathbf{r}, \mathbf{u}$, and $\mathbf{s}$) are made before the number of patient arrivals is realized.

We now describe modeling of the vial administration decisions. The vials opened in period n can vaccinate children in period $m < n + \tau$, or before the end of a session, whichever occurs first. With y_{nm} to denote the number of doses obtained from vials opened in period n, and used in period m, then $y_{n(n+\tau)}$ represent the number of doses that expire after their safe use time expires, and hence contributes to OVW. The relationship between the number of vials opened and the number of doses available is captured by the following equation:

$$\sum_{m=n}^{n+\tau-1} y_{nm} + y_{n(n+\tau)} = \sum_{\nu \in \mathcal{V}} q_\nu u_{\nu n} \qquad \forall n \in \mathcal{N}, \tag{2.3}$$

where q_ν is the number of doses in a multi-dose vial of size ν. The right-hand side of equation (2.3) represents the total number of doses obtained by opening vials, while the left-hand side is the sum of utilized and expired doses in period n.

The total number of patients arriving in time period n is denoted by $\tilde{\omega}_n$. If the total number of patients arriving during period t exceeds the number of doses available, then these patients will

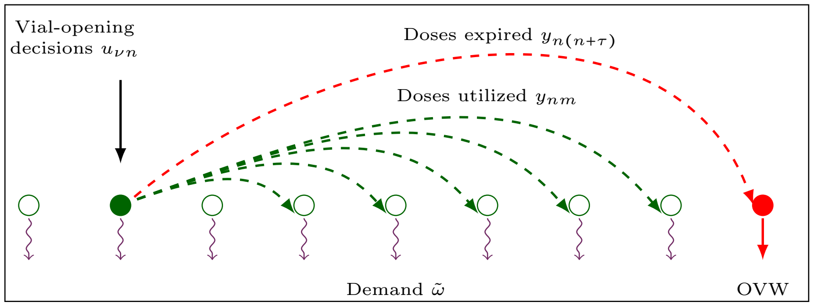


Figure 2.2: Dose utilization for $n = 2$ with $\tau = 7$.

* Filled circles indicate opened vials and unfilled circles indicate unopened vials.

not be served. This model assumes that patients do not wait for vaccinations. This assumption is relaxed in §2.4.4 when the wait-to-open policy is described. A penalty is associated with each unserved patient, which is denoted by ℓ_n. With these, the dose inventory must satisfy the following flow-balance equations:

$$\sum_{m=n-\tau+1}^{n} y_{mn} + \ell_n = \omega_n \qquad \forall n \in \mathcal{N}. \tag{2.4}$$

Here, ω_n is a realization of the random variable $\tilde{\omega}_n$. Equation (2.4) indicates that, of the total number of patients arriving during a session, the number of patients vaccinated depends on the total number of doses available, denoted by the first term in the left-hand side of the equation. If the number of open doses is smaller than the number of patients present during period t, then the unserved patients leave the session. Figure 2.2 provides a schematic representation of how doses obtained from opened vials in period $n = 2$ are utilized over the next $\tau = 6$ periods. OVW is represented by the arc in red.

We finally describe modeling of the objective function. The objective function includes the fixed (f_t) and variable (c_ν) purchasing costs, the inventory holding costs (d_ν), and the expected value of wastage (g) and penalty (p) costs. This objective function is given by:

$$\sum_{t\in\mathcal{T}} (f_t z_t + \sum_{\nu\in\mathcal{V}} c_\nu r_{\nu t}) + \sum_{\nu\in\mathcal{V}} \sum_{n\in\mathcal{N}} d_\nu s_{\nu n} + \mathbb{E}\{\sum_{n\in\mathcal{N}} (g y_{n(n+\tau)} + p\ell_n)\}. \tag{2.5}$$

The expectation function is taken over the distribution of $\tilde{\omega}$ and its argument captures the penalty associated with unserved patients and OVW disposal. Note that the total disposal cost depends on patient arrivals and is bounded below by zero.

The vaccine vial replenishment problem can be written as a 2-SIP to capture the impact of the stochastic nature of patient arrivals on vaccine inventory replenishment and vaccine administration decisions. The complete 2-SIP is presented in Appendix (A). To simplify the exposition of the algorithm, the following succinct representation of the model formulation is used in the remainder of the chapter. A single decision vector $\mathbf{x} \in \mathcal{X}$ collectively denotes the ordering decision z_t, the replenishment quantity $r_{\nu t}$, vial-opening decisions $u_{\nu n}$, and the state variables $s_{\nu n}$, i.e., $\mathbf{x} = (\mathbf{z}, \mathbf{r}, \mathbf{u}, \mathbf{s})^\intercal$. Note that the decision vector $\mathbf{x}$ consists of binary (z_t), as well as pure-integer decisions[1]. Since decision $\mathbf{x}$ is made prior to realizations of patient arrivals, which is non-anticipative in nature [20], $\mathbf{x}$ represents the first-stage decision. On the other hand, vial opening decisions y_{nm} depend on patient arrivals, and therefore, these decisions are made in an adaptive manner. We collate all the adaptive decisions in a single decision vector $\boldsymbol{\gamma} = (\mathbf{y}, \ell)^\intercal$. This allows us to write the replenishment model as:

$$\begin{aligned} \min \ & F(\mathbf{x}) + \mathbb{E}(H(\mathbf{x}, \tilde{\omega})) \\ s.t. \ & \mathbf{x} \in \mathcal{X} \end{aligned} \tag{2.6}$$

where the recourse function $H(\mathbf{x}, \omega)$ takes the first-stage decisions $\mathbf{x}$ and realization ω of random variable $\tilde{\omega}$ as input. The resulting model is:

$$\begin{aligned} H(\mathbf{x}, \omega) = \min \ & G(\boldsymbol{\gamma}) \\ s.t. \ & \mathbf{W}\boldsymbol{\gamma} = \rho(\boldsymbol{\omega}) - \mathbf{T}\mathbf{x}. \end{aligned} \tag{2.7}$$

In the equations above, functions $F(\cdot)$ and $G(\cdot)$ are linear functions; $\mathbf{W}$ is the recourse matrix; and $\mathbf{T}$ is the technology matrix. In the SP literature, formulation (2.6) is referred to as the first-stage problem, and formulation (2.7) is the second-stage problem.

Equations (2.3) and (2.4) represent the constraints set of formulation (2.7). The first-stage decision variable $\mathbf{x}$ affects the right-hand side of equation (2.3). $\mathbf{Tx}$ in formulation (2.7) represents this relationship. The uncertainty affects the right-hand side of equation (2.4). $\rho(\boldsymbol{\omega})$ in formulation (2.7) denotes this relationship. Consequently, the recourse matrix $\mathbf{W}$ represents the left-hand sides

[1] We use the notation $\mathbf{v}$ to denote a single column vector obtained from elements $(v_1, v_2, \ldots, v_i)^\intercal$ for one dimensional vectors, column-wise concatenation for two-dimensional matrices $(v_{11}, \ldots, v_{i1}, v_{12}, \ldots, v_{i2}, \ldots, v_{1j}, \ldots, v_{ij})^\intercal$, and so on.

of equations (2.3), and equation (2.4) is independent of randomness. Such SPs are said to have a fixed recourse [20]. Finally, the first- and second-stage problems have discrete decision variables. However, the structure of the second-stage constraint set satisfies the following proposition:c

Proposition 1 *The linear programming in second-stage problem* (2.7) *has at least one integer optimal solution.*

Proof: Second-stage problem (2.7) is an uncapacitated minimum cost network flow model. Therefore, the recourse matrix satisfies the total unimodularity property [11]. In other words, since the recourse matrix is integer, every basic feasible solution to the linear relaxation of second-stage problem is integer for any integer vector $(\sum_{\nu \in \mathcal{V}} q_\nu u_{\nu n}, \omega, \forall n \in \mathcal{N})$ which appears on the right-hand side of (2.7). Since the number of patients and number of doses available are integer, this condition is satisfied. ■

As a consequence of Proposition 1, the integer constraints are relaxed in the second-stage problem. Thus, this problem is solved as a linear program for a given input of $(\mathbf{x}, \boldsymbol{\omega})$. Linearizing second stage problem (2.7) is necessary for the implementation of the LS algorithm which relies on the duality of the second-stage problem. From this point on, our notation 2-SIP refers to an integer first-stage problem (2.6) and the linear relaxation of the second-stage problem (2.7). The solution algorithm is presented in the next section.

2.3 Solution Approach: An extended L-shaped method

2-SIPs have been used to model many applications in the fields of financial planning, capacity expansion, manufacturing, resource allocation, etc. Over the past several decades, different algorithms have been proposed to tackle these problems. To achieve computational tractability, many algorithms represent uncertainty through a finite number of realizations (scenarios). If $\boldsymbol{\Omega} = \{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_S\}$ represents this finite set of scenarios with respective probabilities $p_i,\ i = 1, \ldots, S$, then the expectation function in first-stage problem (2.6) can be written as:

$$\mathbb{E}\{H(\mathbf{x}, \tilde{\boldsymbol{\omega}})\} = \sum_{i=1}^{S} p_i H(\mathbf{x}, \boldsymbol{\omega}_i). \tag{2.8}$$

Discretizing ω allows modeling of the corresponding deterministic equivalent formulation (DEF) of the proposed 2-SIP. In general, even if S is relatively small, the size of DEF can increase quickly. Since the decisions variables of the first-stage problem are integers, DEF is a network flow problem with fixed charge costs, known as an NP-Hard problem [65]. Therefore, solving DEF is computationally challenging.

Decomposition-based methods, notably, the stochastic Benders decomposition [134], Dantzig-Wolfe decomposition [41], and progressive hedging [123] have effectively addressed this computational challenge. These methods iteratively build approximations to the expected recourse function. Dantzig-Wolfe decomposition is an inner linearization method that iteratively solves the dual of the first-stage problem. However, this method is not applicable to the class of problems with discrete first-stage variables, which is the case here. Progressive hedging is a primal-dual method where, in each iteration, a penalty is associated with a deviation from a feasible solution. However, using this method requires selecting an appropriate proximal parameter that is both instance dependent and hard to discern [137]. Therefore, the solution approach proposed here is based on stochastic Benders decomposition which is also known as the LS method.

In order to build the DEF of a 2-SIP, whose random variable has a large number of realizations, or its underlying distribution is continuous (as is the case in our application), it is recommended to use a sampling-based approach [20]. In this case, one can generate S realizations of random vector $\tilde{\boldsymbol{\omega}}$ using Monte Carlo simulation. Then, the set $\boldsymbol{\Omega}$ will be comprised of these simulated vectors, and each will have the same probability $p_i = 1/S$. Given this set of realizations, the SAA problem of the DEF can be written as follows:

$$\min_{\mathbf{x}\in\mathcal{X}} \; \hat{F}_S(\mathbf{x}) \coloneqq F(\mathbf{x}) + \frac{1}{S}\sum_{i=1}^{S} H(\mathbf{x}, \boldsymbol{\omega}_i). \tag{2.9}$$

The second term in problem (2.9) is an unbiased estimator of the expectation function in formulation (2.8). The SAA problem is a 2-SIP with discrete distribution and therefore can be solved using the LS method with a branch-and-cut procedure to recover the integrality of the first-stage problem. Since the SAA problem is constructed using random realization obtained by sampling using Monte-Carlo or other techniques, verifying the quality of the solutions obtained from the SAA is imperative. Thus, this study uses estimates of lower bounds (LBs) and upper bounds (UBs) computed across multiple replications, as suggested by [92].

The LS method is an iterative method used to solve SP problems with continuous recourse. It is well known that, in these problems, the expected recourse function is piecewise linear and convex [20]. Therefore, the dual solutions of the second-stage problem are used to generate a LB and provide an outer approximation of this expected recourse function, in each iteration of the LS method. When second-stage problem variables are discrete valued, then additional steps are necessary to achieve convexity; see [125] for details. Due to the total unimodularity of the recourse matrix $\mathbf{W}$, employing such procedures is not necessary. In order to solve problems with binary constraints in the first stage, Laporte and Louveaux proposed the integer LS method in 1993 [75]. Later on, improved optimality cuts were introduced by [64] to strengthen the method. This method was further improved by using local branching techniques developed by [122]. These methods work for problems with only first-stage binary variables and hence do not apply to the problem presented here.

In order to improve the computational time of solving large-scale mixed-integer two-stage SPs with continuous recourse, [22] add valid integer cuts within the LS method. The cuts are added to the LP relaxation of the first-stage problem. Their computational results show an improvement in the LS algorithm's performance. The solution algorithm developed here is an extension of the LS method to solve the SAA formulation in which the first-stage problem is a mixed-integer program. This work uses GMI and MIR cuts to strengthen Benders' cuts. In this regard, our approach resembles the work of [22] in which cuts are added to the DEF of their problem.

2.3.1 Extended LS Method:

When the classical LS method of [134] is applied to 2-SIPs with continuous recourse, optimization is achieved by solving the first-stage problem as an MIP in every iteration. These MIP solvers often tend to be computationally expensive, prohibiting extensive numerical experimentations necessary to conducing reliable statistical analyses of the results obtained by solving an SAA-based formulation of the problem. Herein lies the motivation for the algorithm presented here.

Our algorithm operates in one of two modes: (a) *optimization* mode and (b) *integer-feasibility* mode. In the optimization mode, the goal is to develop acceptable LB approximations of the first-stage objective function. In this regard, the procedure adopted is akin to the classical LS method applied to a two-stage stochastic linear program (2-SLP). The goal of the integer-feasibility mode is to determine an integer feasible solution.

The algorithm is initialized by setting the first-stage feasible region $\overline{\mathcal{X}}^0 = \mathcal{X}$, the set of original constraints. Over the course of the algorithm, lower-bounding affine functions are computed in the optimization mode and integer cuts are added in the integer-feasibility mode. In iteration k, the first-stage feasible region $\overline{\mathcal{X}}^{k-1}$ is characterized by the set of original constraints $\mathcal{X}$, a set of lower bounding cuts $\mathcal{L}_{opt}^{k-1}$, and a certain set of integer cuts $\mathcal{L}_{int}^{k-1}$. The lower bounding cuts are expressed using an auxiliary variable η, as in the case of the classical LS method. While $\mathbf{x}$ is an integer decision variable, η is a free continuous variable. Therefore, the feasible region for iterations $\overline{\mathcal{X}}^{k-1}$, $k > 1$ is described in an extended space which includes both $\mathbf{x}$ and η, a mixed-integer set. The second-stage problem in formulation (2.7) satisfies the relatively complete recourse assumption; that is, the second-stage problem is feasible for all $\boldsymbol{\omega}_i \in \boldsymbol{\Omega}$ and $\mathbf{x} \in \overline{\mathcal{X}}^k$, which mitigates the need to generate feasibility cuts. Feasibility is achieved because not every patient is assumed to be vaccinated, and waste is allowed and accounted for. However, the relatively complete recourse assumption can be relaxed, and a mechanism to generate feasibility cuts can be incorporated in the optimization mode. These feasibility cuts will be generated in a manner similar to the classical LS method, and therefore, they are not discussed here. First, the steps involved in the optimization mode procedure are described, followed by the approach used in the integer-feasibility mode.

2.3.1.1 Optimization Mode.

In this mode, iteration k starts by solving the following problem that identifies a candidate solution:

$$\min \ \{F(\mathbf{x}) + \boldsymbol{\eta} \mid (\mathbf{x}, \boldsymbol{\eta}) \in \overline{\mathcal{X}}_{lp}^{k-1}\} \tag{M_{lp}^k}$$

where $\overline{\mathcal{X}}_{lp}^{k-1}$ is a linear relaxation of the first-stage feasible region $\overline{\mathcal{X}}^{k-1}$. The optimal solution of this problem is given by $(\mathbf{x}^k, \boldsymbol{\eta}^k)$, with objective function value v_S^k. Using the solution $\mathbf{x}^k$ and a realization $\boldsymbol{\omega}_i \in \Omega$, the second-stage problem $H(\mathbf{x}^k, \boldsymbol{\omega}_i)$ is solved to obtain the optimal dual solution, $\boldsymbol{\pi}^k$. This procedure is enumerated for every $\boldsymbol{\omega}_i \in \boldsymbol{\Omega}$. Using these dual solutions, a lower bounding affine function cut is computed by:

$$l_{opt}^k(\mathbf{x}, \boldsymbol{\eta}) := \frac{1}{S}\sum_{i=1}^{S}(\boldsymbol{\pi}^k)^\top[\rho(\boldsymbol{\omega}_i) - \mathbf{T}\mathbf{x}] - \boldsymbol{\eta} \leq 0. \tag{2.10}$$

This affine function is used to obtain the updated set $\mathcal{L}^k_{opt} = \mathcal{L}^{k-1}_{opt} \cup \{l^k_{opt}(\mathbf{x}, \boldsymbol{\eta})\}$ which provides a LB approximation to the objective function.

The optimal objective function value v^k_S provides a LB to (M^k_{lp}). A UB to (M^k_{lp}) is $\hat{F}_S(\mathbf{x}^k) = l^k_{opt}(\mathbf{x}^k, \eta^k)$. Let Δ^k denote the relative gap between the current UB and LB defined as $\Delta^k(\mathbf{x}^k) = (\hat{F}^k_S(\mathbf{x}^k) - \hat{v}^k)/\hat{F}^k(\mathbf{x}^k)$. If $\Delta^k(\mathbf{x}^k)$ is less than a predefined error ϵ, then switch to the integer-feasibility mode. Otherwise, if $\Delta^k(\mathbf{x}^k) \geq \epsilon$, then set $\mathcal{L}^k_{int} = \mathcal{L}^{k-1}_{int}$, i.e., no new integer cuts are added in iteration k. This results in an updated first-stage feasible region $\overline{\mathcal{X}}^k$, and one iteration in optimization mode is complete.

Remark : While the algorithm operates in the optimization mode, no new integer cuts are added. Therefore, when the relative gap $\Delta^k(\mathbf{x}^k) < \epsilon$, the algorithm has solved a 2-SLP with a first-stage problem characterized by a feasible region $\{\mathbf{x}^k \mid \mathbf{x}^k \in \mathcal{X}^k \cap \{l(\mathbf{x}^k, \boldsymbol{\eta}^k) \leq 0\}, l \in \mathcal{L}^\kappa_{int}\}$, where κ is the last iteration when the algorithm operated in integer-feasibility mode. This feasible region is defined by the original set of constraints and the integer cuts added up to iteration κ. Therefore, the algorithm presented here converges to the optimal solution of this problem in a finite number of iterations if ω has a finite support [20].

2.3.1.2 Integer-feasibility mode.

Let κ denote the iteration when switching to integer-feasibility mode. At this point, the current solution is assigned as an incumbent solution, i.e $\hat{\mathbf{x}}^\kappa = \mathbf{x}^\kappa$, with the corresponding objective function value $\hat{v}^\kappa = v^\kappa$. If the incumbent solution satisfies the integer requirement, then we have an integer feasible solution within an acceptable relative gap. Therefore, the algorithm is terminated. Otherwise, a mixed-integer program is solved:

$$\min \ \{F(\mathbf{x}) + \boldsymbol{\eta} \mid (\mathbf{x}, \boldsymbol{\eta}) \in \overline{\mathcal{X}}^{k-1}_{lp} \cap l^k_{opt}(\mathbf{x}, \boldsymbol{\eta}) \leq 0\}. \qquad (M^\kappa_{mip})$$

Note that the constraint set in the above problem does includes the latest optimality cut added in the optimization mode. Let $\hat{\mathbf{x}}^\kappa_S$ be the integer optimal solution with the objective function value $\hat{v}^\kappa_S$. Since the solution is updated, the gap estimate $\Delta^\kappa(\hat{\mathbf{x}}^\kappa)$ is recomputed with $\hat{F}_S(\hat{\mathbf{x}}^\kappa)$ as the UB. This update might result in Δ^κ becoming greater than error ϵ. In this case, we conclude that the lower bounding approximation for $\hat{\mathbf{x}}^\kappa$ is not acceptable, and hence, the optimization mode is necessary to improve the approximation.

Before returning to the optimization mode, GMI and MIR cuts are constructed at $(\mathbf{x}^\kappa, \boldsymbol{\eta}^\kappa)$ and added to the feasible region [63]. As a result, the feasible region no longer contains the non-integer solution $\mathbf{x}^\kappa$. Note that integer cuts are generated based on the original constraint set, as well as the affine lower bounding functions added so far. Particularly, GMI cuts are generated for fractional basic solutions to $(\overline{M}_{lp}^{k})$. Therefore, multiple GMI cuts are generated up to iteration κ using fractional basic solutions. Let $\mathcal{L}_{gmi}^\kappa$ denote the set of all GMI cuts generated in iteration κ. The MIR cuts are generated only for constraints with mixed-integer variables. In the SAA model, the variables used in the constraint set of the first-stage problem ((2.1), (2.2)) are integers. Therefore, MIR cuts are only used for the lower bounding affine functions in $\mathcal{L}_{opt}^\kappa$, which contain the continuous variable $\boldsymbol{\eta}$. Moreover, the observation that the recourse function $H(\cdot)$ is lower bounded by zero almost surely plays a critical role in deriving the MIR inequalities. The set of MIR inequalities added in iteration κ is denoted by $\mathcal{L}_{mir}^\kappa$. Together, the updated set of integer cuts becomes:

$$\mathcal{L}_{int}^\kappa = \mathcal{L}_{int}^{\kappa-1} \cup \mathcal{L}_{gmi}^\kappa \cup \mathcal{L}_{mir}^\kappa. \tag{2.11}$$

This formulation completes the steps involved in the integer-feasibility mode of the proposed algorithm, which results in an updated first-stage feasible region $\overline{\mathcal{X}}^\kappa$. The algorithm returns to a new iteration in optimization mode.

2.3.2 Statistical Evaluation

Notice that $(\hat{\mathbf{x}}_S, \hat{v}_S)$ is the optimal solution-value pair to the SAA problem in formulation (2.9) when sample size is S. However, the sample size which is required to ensure that $\hat{\mathbf{x}}$ is an optimal solution to the true problem (2.6), is problem specific. Moreover, since the samples used to setup the SAA are chosen randomly through the Monte Carlo sampling method, the optimal objective function value obtained is also a random quantity. Therefore, concluding that the solution obtained from a single SAA is optimal to first-stage problem (2.6), may be erroneous. In order to overcome such premature conclusions on optimality, the experiments should be replicated with different sets of scenarios. The idea of using multiple replications for estimating the LBs and UBs, which are then used to obtain estimates on an optimality gap, was proposed by [93]. This idea has been incorporated successfully in the SP literature to provide statistical performance guarantees for 2-SLPs [85]. Here, a similar experimental setting is used for the SAA approach to develop a 2-SIP

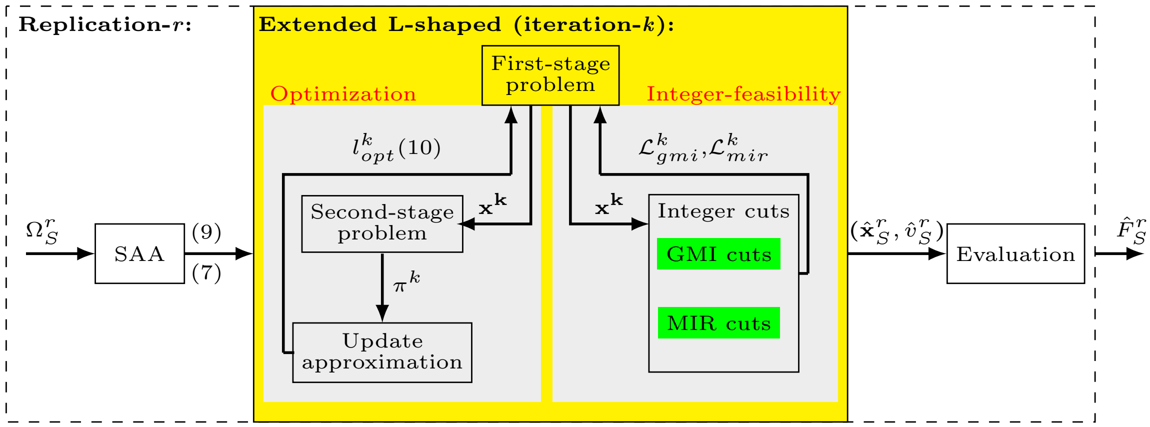


Figure 2.4: Schematic representation of extended L-shaped algorithm.

with continuous recourse.

For a given selection of S, let Ω^r_S denote the set of scenarios generated, and let $(\hat{\mathbf{x}}^r_S, \hat{v}^r_S)$ be the optimal solution pair obtained using the extended LS algorithm in replication $r = 1, \ldots, R$. In this regard, an estimator of the optimal objective function value of the true problem (2.6), is given by $\bar{v}^R_S = \frac{1}{R}\sum_{r=1}^R \hat{v}^r_S$. Note that $\hat{v}^r_S$ is an LB, and therefore $\bar{v}^R_S$ is a biased estimator of the objective function value.

In each replication, the quality of solution $\hat{\mathbf{x}}^r_S$ is evaluated by simulating the second-stage problem using scenarios different from the ones in Ω^r_S. This is an out-of-sample evaluation process. The evaluation continues until a sample of size S' is found, such that the corresponding $\hat{F}^r_{S'}$ is contained in the $(1-\alpha)$ confidence interval created using $\hat{F}^r_s$ for $s = 1, \ldots, S'$. Consider the estimator $\Delta_r = \hat{F}^r_S(\hat{\mathbf{x}}^r_S) - \hat{v}^r_S$ of the optimality gap. The mean and variance of this estimator are given by:

$$\bar{\Delta}^R = \frac{1}{R}\sum_{r=1}^R \Delta^r, \quad \sigma^2_\Delta = \frac{1}{R(R-1)}\sum_{r=1}^R (\Delta^r - \bar{\Delta}^R)^2. \tag{2.12}$$

Notice that $\bar{\Delta}^R$ overestimates the optimality gap and therefore, can be viewed as a pessimistic gap. In our computational study, we report this gap and the distribution of $\hat{F}^r_{S'}$ over replications.

Before the conclude this section, we present the distinguishing features of our algorithmic framework, compared to prior works.

1. If the LS method is implemented to solve the SAA, then the MIP first-stage problem presented in this study would be solved in every iteration of the algorithm. Instead, the algorithm proposed here solves a linear first-stage problem during the optimization mode, which

Algorithm 1 Pseudocode for one replication of the extended L-shaped algorithm.

Initilize ϵ, $k \leftarrow 0$, $\mathcal{L}_{int}^{0} = \varnothing$, $\mathcal{L}_{opt}^{0} = \varnothing$
Optimality mode
$k \leftarrow k + 1$
Solve (M_{lp}^{k}) to obtain $(\mathbf{x}^k, \boldsymbol{\eta}^k)$ with the objective function value of v^k
for all $i \in \{1, \dots, S\}$ **do**
    Solve the second-stage problem $H(\mathbf{x}^k, \omega_i)$ to obtain $\boldsymbol{\pi}^k$.
end for
Compute $\ell_{opt}^{k}(\mathbf{x}^k, \boldsymbol{\eta}^k)$ using (3.11) and add to $\mathcal{L}_{opt}^{k}$
$\hat{F}_S(\mathbf{x}^k) = l_{opt}^{k}(\mathbf{x}^k, \boldsymbol{\eta}^k)$
if $\Delta^k(\mathbf{x}^k) = (\hat{F}_S^k(\mathbf{x}^k) - \hat{v}^k)/\hat{F}^k(\mathbf{x}^k) < \epsilon$ **then**
    Go to integer-feasibility mode
else
    $\mathcal{L}_{int}^{k} = \mathcal{L}_{int}^{k-1}$
end if
Integer-feasibility mode
Set $k = \kappa$, $\hat{\mathbf{x}}^\kappa = \mathbf{x}^\kappa$, $\hat{v}^\kappa = v^\kappa$
if $\hat{\mathbf{x}}^\kappa \in \mathbb{Z}^+$ **then**
    STOP
else
    Solve (M_{mip}^{κ}) to obtain $\hat{\mathbf{x}}_S^\kappa$ with the objective function value $\hat{v}_S^\kappa$
end if
if $\Delta^\kappa(\mathbf{x}^\kappa) = (\hat{F}_S^\kappa(\mathbf{x}^\kappa) - \hat{v}^\kappa)/\hat{F}^\kappa(\mathbf{x}^\kappa) < \epsilon$ **then**
    STOP
else
    Add GMI and MIR cuts, $\mathcal{L}_{int}^{k} = \mathcal{L}_{int}^{k-1} \cup \mathcal{L}_{gmi}^{k} \cup \mathcal{L}_{mir}^{k}$
    Go to optimization mode
end if

significantly reduces its running time. The algorithm seeks an integer solution only during the integer-feasibility mode. The computational results in Table (2.4) illustrate that the proposed algorithm leads to lower running times.

2. The effectiveness of using integer cuts within the LS method has previously been explored in [22]. The authors add GMI cuts in every iteration of the LS method. In addition to GMI cuts, the algorithm proposed here exploits the impact of adding MIR cuts as well. These cuts are only added during the integer-feasibility mode, which reduces the computational time of the algorithm. The results in Table (2.4) indicate that using both GMI and MIR cuts results in finding solutions of higher quality, compared to using GMI cuts only.

3. For the application considered in this study, limited data is available to estimate the underlying distribution of the random variable involved. When this is the case, an evaluation process of the SP algorithms is required to validate and verify the quality of the solutions obtained [93]. While evaluation procedures have been developed for a 2-SLP [85], to the best of our knowledge, the work presented here reflects the first adoption of these procedures for 2-SIPs.

Figure 2.4 shows the schematic representation of the algorithmic framework developed in this study.

2.4 Computational Experiments: A case study for Bangladesh

Instances of the vaccine administration and inventory replenishment problem were created with real-life data from different regions in Bangladesh. These instances were created to highlight the impact of data, deterministic as well as stochastic, on model decisions and to draw insights about the problem as it actually exists in the real world. The optimal solution of the SAA is used as a benchmark for the simple, easy-to-implement inventory replenishment and vaccine administration policies proposed here.

This section is organized as follows: First, the input data is presented; second, the experimental setup is explained; third, the results of solving the SAA are analyzed. The inventory replenishment and vaccine administration policies derived from solving the SAA are referred to as the “base policy” to distinguish them from some easy-to-implement policies proposed at the end of this section.

Table 2.1: Cost and weight of different vaccine vial sizes of Pentavalent.

Vaccine size	Purchase cost per dose ($)			Vaccine weight per dose (gr)		
	Lower limit	Mean	Upper limit	Lower limit	Mean	Upper limit
1 dose	3.24	3.60	3.96	1.542	1.713	1.885
2 dose	3.25	3.50	3.75	1.723	1.914	2.106
10 dose	2.00	1.80	2.20	3.169	3.522	3.874

2.4.1 Data input model

Two different data sources were used in this case study to generate problem instances for an outreach session conducted by EPI in Bangladesh. Two regions, based on varying demographic data, were selected: urban Chittagong and rural Rajshahi. Deterministic data, including purchase and inventory holding costs, was complied using the 2012 WHO vaccine volume calculator [2]. Stochastic data was compiled from the Demographic and Health Survey (DHS) of Bangladesh, a population-based survey executed by National Institute for Population Research and Training in 2011 [105] that includes individual household level DHS data. The details of deterministic and stochastic parameters are described next.

2.4.1.1 Deterministic parameters:

The EPI outreach sessions in rural areas are scheduled on a weekly basis [52]. Thus, this study terms the planning horizon to include 2 weeks with 5 working days per week and 2 to 8 working hours per day. Based on the proposed model, inventory replenishment decisions are made at the beginning of each week, and the vaccine administration decisions are made hourly. Experiments for the Pentavalent vaccine, which is distributed in one-, two-, and ten-dose vials, were conducted. However, this setup can apply to other pediatric vaccines as well. The Pentavalent vaccines has 6 hours safe use time before expiration (i.e. $\tau = 6$). The 2012 WHO vaccine volume calculator [2] provides data about vial purchasing costs and corresponding weights, which are summarized in Table (2.1). Inventory holding cost is estimated based on the weight of the Pentavalent vaccine. Sensitivity analyses, with respect to purchase and inventory holding cost, were conducted. The lower, mean and upper limit of the values are presented in Table (2.1). In all the other experiments conducted, the mean values of purchase and inventory holding costs are used. Order setup cost is assumed to be constant over the planning horizon and is set to $10. This value is estimated based on interviews with health care workers in a developing country. The penalty cost associated with

Table 2.2: Data used to predict the daily demand distribution of Pentavalent through different regions in Bangladesh.

Region	Number of clinics	Infant population	Daily fitted distribution	Hourly estimated distribution
Barisal	41	8,147,000	NB(0.65,0.43)	NB(5,0.43)
Chittagong	141	288,079,000	NB(0.53, 0.81)	NB(286, 0.81)
Dhaka	171	46,729,000	NB(2.08, 0.16)	NB(2, 0.16)
Khulna	70	15,563,000	NB(0.95, 0.33)	NB(4, 0.33)
Rajshahi	182	18,329,000	NB(0.84, 0.31)	NB(2, 0.31)
Rangpur	43	15,665,000	NB(5, 0.93)	NB(327, 0.93)
Sylhet	62	9,807,000	NB(0.78,0.38)	NB(5,0.38)

unserved patients is estimated to be $100 and is based on a sensitivity analysis conducted as part of this study. A higher penalty value resulted in scalability issues, and a smaller value resulted in a large number of unserved patients. The selected penalty avoided scalability problems and provided realistic numbers of unserved patients. The UB of order size M_t is set equal to 2000 for all t. This value is large enough to capture vaccination shortage dues to OVW only, not capacity. Finally, the penalty cost for OVW equals the vaccine's purchase cost.

2.4.1.2 Stochastic parameters:

The only stochastic element in our model is patient arrival. In order to determine how many arrivals an outreach program might expect, this study uses the DHS data set, which includes both demographic information and interviews with families from each region of Bangladesh. The interviews provide information about vaccination history of children younger than 5. We used each child's vaccination date to count the number of daily observations, or patient arrivals, over a 1-year time interval. Since the outreach sessions are held only on weekdays, only weekday data is considered. The data is used to estimate the distribution of the total number of patients arriving to an outreach session in one day. We conducted goodness-of-fit tests to find the best distribution for this data, and summaries of results are listed in Table (2.2). Here NB refers to Negative Binomial distribution.

Since no information is available about the timing of vaccinations during the day, NB distribution is applied to estimate the total number of patients arriving within the hour. Furthermore, this work assumes that patients' arrivals within an hour follow a uniform distribution. Such an assumption is not restrictive since the model does not penalize for patient waiting time. The last

Table 2.3: Paired t-test between different number of scenarios.

	***p*–value (95% CI)**							$\left\Vert \hat{\mathbf{x}}^{\mathbf{S}} - \hat{\mathbf{x}}^{\mathbf{S}'} \right\Vert / \left\Vert \hat{\mathbf{x}}^{\mathbf{S}} \right\Vert$						
S; S′	Rajshahi	Barisal	Khuluna	Sylhet	Dhaka	Rangpur	Chittagong	Rajshahi	Barisal	Khuluna	Sylhet	Dhaka	Rangpur	Chittagong
100; 200	0.99	0.00	0.00	0.00	0.14	0.00	0.44	0.30	0.07	0.04	0.06	0.02	0.02	0.05
200; 500	0.99	0.84	0.29	0.80	0.03	0.08	0.01	0.33	0.05	0.01	0.05	0.03	0.02	0.05
500; 1,000	0.99	0.67	0.89	0.35	0.48	0.00	0.96	0.22	0.05	0.03	0.05	0.02	0.02	0.02
1,000; 2,000	0.99	0.74	0.66	0.99	0.58	0.39	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00

column of Table (2.2) shows the estimated arrival rates.

Scenario selection: Selecting an appropriate number of scenarios for our SAA formulation is critical, since, generating a large number of scenarios increases the running time of the algorithm, and generating a small number of scenarios may not appropriately represent uncertainty. Two different sets of analyses are used to assess the difference between the optimal solutions, $\hat{\mathbf{x}}^S$ and $\hat{\mathbf{x}}^{S'}$, obtained in S and S' scenarios.

The first set of analysis captures the impact of the optimal solution on the recourse cost by fixing the optimal solution and simulating the recourse function over the two sets of scenarios. A paired t-test on the function values was performed. The corresponding p–values for each region are presented in Table (2.3). A p–value greater than 0.05 implies that the two solution $\hat{\mathbf{x}}^S$ and $\hat{\mathbf{x}}^{S'}$ have a statistically indistinguishable impact on the recourse cost. The second set of analysis is based on the Euclidean norm of the two solutions. Table (2.3) shows the results. A high value indicates that the solutions are significantly different. As Table (2.3) shows, $\hat{\mathbf{x}}^{1,000}$ and $\hat{\mathbf{x}}^{2,000}$ have acceptable p–values and relatively small Euclidean distances. For all regions considered in this study, increasing the number of scenario from $1,000$ to $2,000$ is not necessary and therefore, the number of scenarios used is $S = 1,000$.

2.4.2 Experimental setup

The solution algorithm presented here was implemented in C programming language on a 64-bit Intel Core i5 processor @ 1.9 GHz with 8 GB RAM. All first-stage and second-stage problems were solved using CPLEX callable subroutines.

Our experiments began by finding an optimal or near optimal first-stage solution (error gap < 0.01) using the extended LS method. Next, a posterior analysis was conducted to evaluate the quality of the first-stage solutions by fixing the first-stage decisions and simulating the second-stage problem (2.7). The scenarios used for optimization and evaluation were generated using the same distribution. Evaluation terminated when the standard deviation of the value of the recourse function

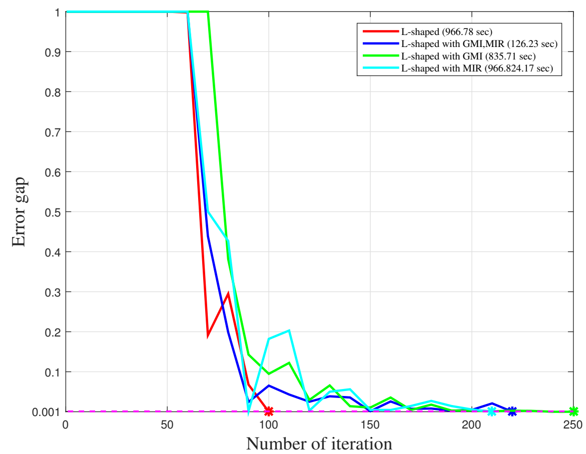


Figure 2.5: Base policy: analysis of solution algorithm.

Table 2.4: Comparison of solution quality.

Algorithm	LB ($)	UB ($)	CPU time (s)	No. of iter.	No. of LP	No. of MILP
LS	$21,571.7 \pm 12.1$	$22,047.9 \pm 33.8$	859.7 ± 245.9	233.8 ± 11.2	0	232.7 ± 11.2
LSM	$21,574.9 \pm 11.8$	$22,048.9 \pm 30.7$	350.4 ± 86.1	836.3 ± 83.5	756.4 ± 79.3	78.9 ± 11.8
LSG	$21,571.9 \pm 12.1$	$22,047.1 \pm 33.6$	412.2 ± 160.1	837.9 ± 93.8	756.4 ± 88.5	80.5 ± 12.7
LSMG	$21,571.7 \pm 12.1$	$22,047.9 \pm 33.9$	137.2 ± 47.3	431.1 ± 26.3	335.6 ± 16.6	94.5 ± 12.4

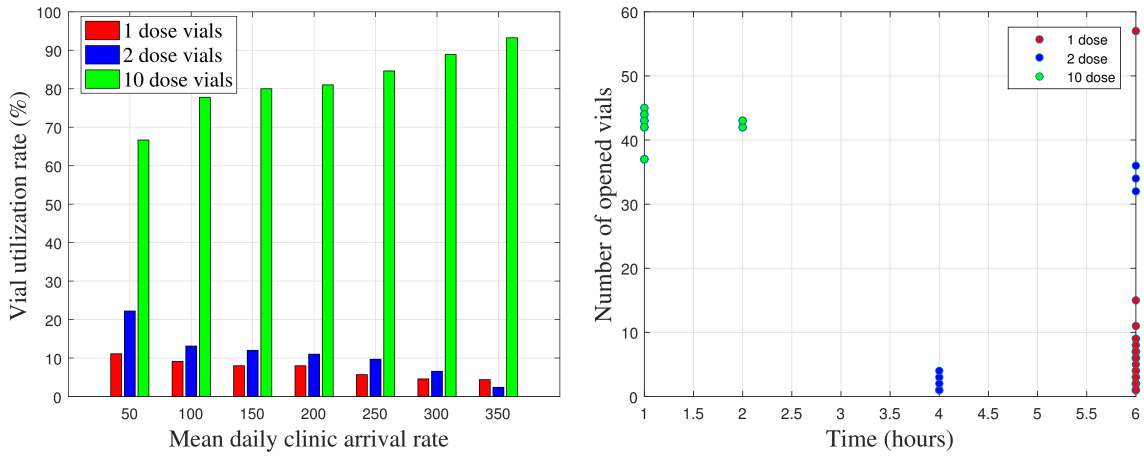
came within a tolerable limit which is equal to 0.01. This process was replicated using independent samples. In our experiments we conducted $R = 30$ replications. The results are presented in terms of the empirical distribution over replications, and heuristic policies are also treated similarly.

2.4.3 Base policy

First, the performance of the solution algorithm and the model's adequacy must be verified. The inventory replenishment and vaccine administration policies identified by solving the extended LS method outlined in §2.3.1 are called "base policy". Results are provided for the base policy and are compared with the results obtained by the classical LS method in Table (2.4).

2.4.3.1 Performance of solution algorithms:

This set of experiments is implemented across the Chittagong dataset with vaccines of a single vial size, such as a ten-dose Pentavalent. The goal is to evaluate the impact of MIR and GMI cuts in the performance of the extended LS method. The quality of solutions obtained under the following conditions are compared with the classical LS method: (a) use only MIR cuts (LSM), (b)



(a) Vial portfolio.

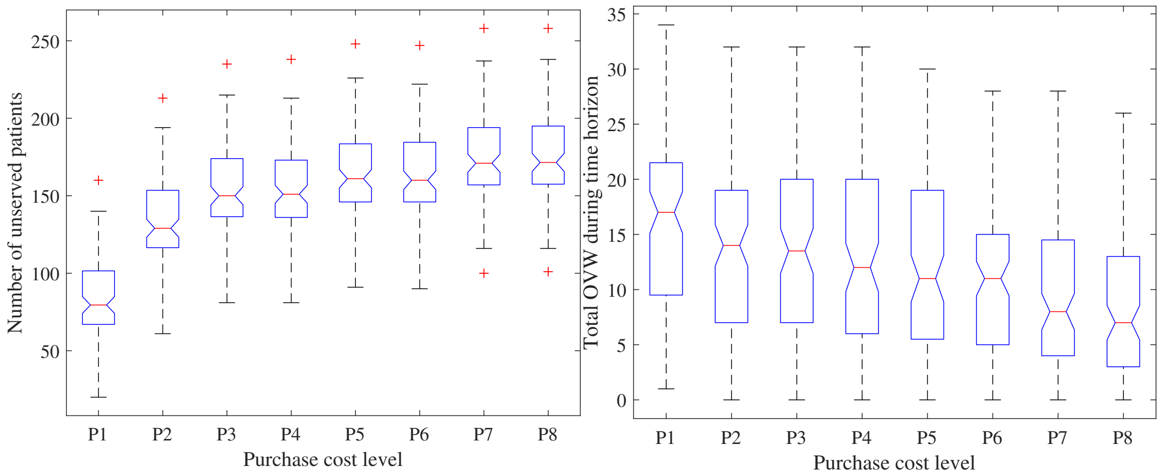
(b) Daily consumption schedule.

Figure 2.6: Base policy: analyzing the impact of mean daily patient arrival rate on vial portfolio and daily consumption schedule on number of opened vials.

use only GMI cuts (LSG), and (c) use both MIR and GMI cuts (LSMG). Table (2.4) summarizes the solution quality of each algorithm in terms of the CPU time, total number of iterations, LB, UB, total number of LPs solved, and total number of MILPs solved over 30 replications. Note that the values of LB and UB for all algorithms are similar, and the LS method uses the least number of iterations. However, in every iteration of the LS method, an MILP is solved, and therefore, the CPU time is significantly high. In algorithms that require integer cuts, MILPs are solved only when the error gap is below a threshold $\epsilon = 0.1\%$ and $\mathbf{x} \notin \mathbb{Z}^{+}$. Therefore, the first-stage problem is solved as an LP in most iterations, particularly in the first iterations of the algorithm, which results in not only a higher number of iterations, but also a shorter computational time.

Experimental results indicate that using MIR cuts results in a greater decrease in computational time than GMI cuts. The greatest computational time savings, by a factor 6.2, was observed when both types of integer cuts were incorporated. Figure (2.5) illustrates the error gaps over algorithm iterations for one of the replications.

This analysis establishes that the LSMG algorithm provides solutions similar to those obtained from LS method in significantly less CPU time. Such computational enhancements are necessary to satisfy the need to replicate SPs. This also allows the algorithm presented here to be applicable to real-world problem instances, such as the one considered in our study.



(a) Unserved patients

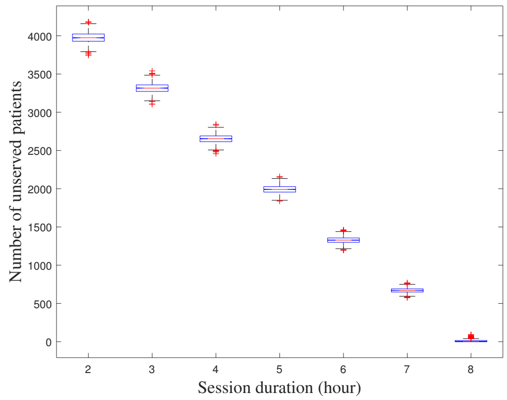
(b) Wastage

Figure 2.7: Effect of varying purchase cost on dose utilization.

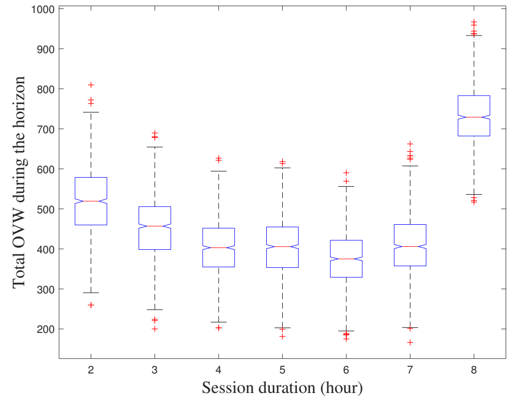
2.4.3.2 Analysis of parameters:

Solutions to the vaccine administration and inventory replenishment model are sensitive to problem parameters, such as patient arrival rate, purchase cost, and duration of a session. Analysis of these parameters helps health care policy makers to evaluate the effectiveness of outreach sessions and design appropriate policies. In the first set of experiments, the impact of different patient arrivals on vial portfolio selection is evaluated. The second experiment is developed to study the effects of varying purchase costs on OVW rates and the number of unserved patients. Finally, we analyze the impacts of session duration on OVW and the number of unserved patients.

1. ***Patient arrival***: This experiment was conducted using regions with different patient arrival rates. For example, the patient arrival rate in Rajshahi is lowest since this is the least populated region with available data, and patient arrival rate in Chittagong is highest, and this is the most populated. Figure 2.6a depicts the relationship among utilization rates of multi-dose vaccines of different sizes and patient arrival rates. Going from Rajshahi to Chittagong, the arrival rate increases, as does the number of ten-dose vials ordered. This increase is attributed to the low cost per dose of multi-dose vials. At high arrival rates, the one-dose vials are ordered to supplement the multi-dose vials, particularly, to meet the end-of-day arrivals. Based on the results presented in Figure 2.6a, procuring larger size vials for moderately and highly populated regions is economically efficient. Likewise, the usage of one- and two-dose vials is higher in



(a) Unserved patients in Chittagong



(b) Wastage in Chittagong

Figure 2.8: Impact of session duration on OVW and number of unserved patients in Chittagong

sessions with small patient arrival rates, such as outreach centers in remote locations. Figure 2.6b shows the number of multi-dose vials opened at different times within the duration of a session. The results indicate that multi-dose vials should be opened in the early hours of a session and one-dose vials at the final hours.

2. ***Purchase cost***: To conduct these experiments 8 problem instances, P1-P8, were developed. Each instance has the same patient arrival rate as in Rajashahi, but each instance has different purchase cost. Different values of purchase costs were used within the lower and upper limits presented in Table 2.1. Instance P1 has the lowest purchase costs, and costs increase from P1 to P8. Figures 2.7a and 2.7b present the number of unserved patients and OVW observed during the evaluation step of one of the replications. As purchase cost increases, vial opening decisions become conservative in order to reduce OVW, which, in turn, increases the number of unserved patients. This analysis provides a tool to aid policy makers identifying the necessary subsidies for vaccine purchase costs and achieving a certain level of immunization in the region. For example, when purchase costs equals the lower limit in instance P1 ($c_1 = 1.6, c_2 = 1.5, c_{10} = 1$), the total number of unserved patients during the 2-weeks planning horizon and in all the outreach sessions in the region is less than 100 for 99% of scenarios tested, as indicated by the top whisker. This number increased to 175 for P3-P4 went up to 190 for P8.

3. ***Session duration***: The choice of session duration affects the numbers of patients who attend these sessions. Since vaccines expire after their safe use time, the choice of session duration also

impacts OVW. To evaluate these impacts, we created problem instances for different regions by changing session duration from 2 to 8 hours. When session duration is less than 8 hours, assume that patients, who cannot attend the session during its operational hours, are lost or unserved. As a result, a shorter session duration increases the number of unserved patients; see Figure 2.8a. The results indicate that OVW reaches its minimum value when the session duration equals the vaccine's safe use time (τ = 6 hours). When the session duration is short, unexpired doses also contribute to OVW. While increasing the session duration decreases the number of unserved patients, the need to open new vials to serve patients arriving in later hours also impacts OVW. Note that increasing session duration also increases operational costs, which were not considered in the model presented here.

Based on this work's analysis of the base policy, the usage of multi-dose vials with complementary one-dose vials for use in last operating hours (when $t > \tau$), is recommended for highly populated and well connected regions, like Chittagong. Demographic data about Chittagong indicates that the majority of the population is employed in jobs which do not provide flexible working hours, and therefore, using longer outreach sessions is beneficial. Different from Chittagong, Rajshahi region is mostly rural, and the majority of its population works in the agricultural sector [68]. This provides mothers with greater flexibility to attend the outreach sessions. Thus, having a 6-hour session duration in this location is appropriate to achieve high coverage and maintain low OVW.

The model presented here and the subsequent analysis targets developing countries, and hence, purchase cost is a critical factor in inventory replenishment and vaccine administration decisions. Therefore, health care policy makers should negotiate the necessary subsidies to achieve the targeted vaccination coverage levels in a cost efficient manner. Our setup provides not only a statistically optimal policy but also a systematic tool to conduct analyses that support decisions related to vaccination in developing countries.

2.4.4 Heuristic policies

Experimental results presented in §2.4.3 indicate that solving the proposed inventory replenishment and vaccine administration model to develop a base policy requires extensive use of engineering tools which are commercially expensive and are rarely accessible to health care administrators in developing countries. In order to address these concerns, this study develops and evaluates four simple, easy-to-use policies based on the base policy. These policies are: (a) first multi-dose, last single-dose (FMLS) policy, (b) shorter duration session (SDS) policy, (c) single vial size policy, and (d) wait to open (WO) policy. These policies impact both vaccine order replenishment and administration. To implement these policies, the first-stage (2.6) and the second-stage (2.7) problems developed above are modified and described below:

a *FMLS*: The principle of this policy follows the observation made in Figure 2.6b regarding the use of one-dose and multi-dose vials: Health care practitioners should open multi-dose vials in the early hours of the session and one-dose vials as the session comes to an end. To implement this approach, the session length is divided into three equal intervals, such that, for $t \in \mathcal{T}$ (a) ten-dose vials are opened during the first two hours of a session ($\mathcal{N}_1 = \{Nt-5, Nt-4\}$), (b) two-dose vials are opened during the next two hours ($\mathcal{N}_2 = \{Nt-3, Nt-2\}$), and (c) one-dose vials are opened in the last two hours ($\mathcal{N}_3 = \{Nt-1, Nt\}$). Mathematically, this approach corresponds to setting $u_{\nu n} = 0$ for $n \in \mathcal{N} \setminus \mathcal{N}_i$ when $q_\nu = i$ in (2.6). For example, for time periods $n \in \mathcal{N}_3$, when only 10-dose vials are opened, $u_{\nu n} = 0$ if $q_\nu = 1$ or 2.

b *SDS*: Results in Figure 2.8 indicate that sessions of 6 hours duration have the minimum OVW, but longer sessions have the lowest number of unserved patients. The SDS policy achieves low OVW by trimming duration of the immunization session from 8 to 6 hours. In the model presented here, we assume that only a certain percentage of patients, who would be arriving during the last 2 hours of a session, are attending the immunization session earlier in the day. The remaining patients are not served.

c *Single vial size*: Current vaccination practice relies on the use of single multi-dose vials [43, 55]. This policy can highlight the advantage of using a mix of multi-dose vials. We implement the single multi-dose vial policy using one-, two, or ten-dose vials by adjusting set $\mathcal{V}$ in the first-stage and second-stage problems.

d *WO*: It is a common practice to delay vaccine administration until a sufficient number of patients have arrived at a session [23]. Such a practice reduces OVW, but if the waiting time is too long, patients may leave without vaccination, which adversely effects immunization coverage. In order to study the effects of this policy, an extension to our model is presented. This extension captures the impacts of patient waiting times on vaccine vial utilization decisions and considers the trade-offs among OVW, unserved patients, and waiting times.

Including patient waiting times affects only the second-stage problem in formulation (2.7). To capture this difference, a new variable must be introduced. Let w_{jn} be the number of patients who arrived in period j and were served in period n. Thus, $j \le n < j + L$, where L is the maximum waiting time in hours. Patients not served within this time window leave the session. A non-linear function $v(j - n)$ for the total patient waiting time is introduced, and a unit cost v is used to estimate the total cost of waiting and not serving patients. Figure 2.9 illustrates how vaccines are administered to patients arriving in time period n. Mathematically, the flow-balance equation (2.4) will be changed to include the following constraints:

$$\sum_{m=n-\tau+1}^{n} y_{mn} = \sum_{j=n-L+1}^{n} w_{jn} \qquad \forall n \in \mathcal{N}, \tag{2.13a}$$

$$\sum_{j=n}^{n+L-1} w_{nj} + \ell_n = \tilde{\omega} \qquad \forall n \in \mathcal{N}. \tag{2.13b}$$

The left-hand side of constraints (2.13a) represents the inventory of doses on hand and the right-hand side represents the total number of patients that arrived within the last L periods. This inventory contains doses from vials opened within the last τ periods. For example, when $L = 2$, the doses available in period n are used to vaccinate patients who arrived in periods $(n - 2)$, $(n - 1)$, and n, as shown in Figure (2.9). Patients who arrive in the current time period, $\tilde{\omega}$, can be served within the next L hours. Patients waiting longer than L hours leave the session. These patients are captured by the terms on the left-hand side of equation (2.13b) and represented by the incoming arcs in Figure 2.9. Now the recourse function $H(\mathbf{x}, \omega)$ can be rewritten as follows:

$$\begin{aligned} H(\mathbf{x}, \omega) = \min \quad & \sum_{n \in \mathcal{N}} \Big(g y_{n(n+\tau)} + p\ell_n + \sum_{j=n+1}^{n+L} v(j-n) w_{nj}\Big) \\ s.t. \quad & (2.3), (2.13) \\ & y_{nm}, \ell_n, w_{nj} \in \mathbb{Z}^+ \qquad \forall n \in \mathcal{N}, m \in \{n, \ldots, n+\tau+1\}, j \in \{n, \ldots, n+L\}. \end{aligned} \tag{2.14}$$

In these experiments, the maximum waiting time was set to $L = 2$ hours, and the penalty for waiting to $v = 0.24$\$/hour, the average hourly wage in Bangladesh.

To implement this policy, we solve the 2-SIP with (2.6) as first-stage problem and (2.14), instead of (2.7), as the second-stage problem. To experiment with this policy, the experimental

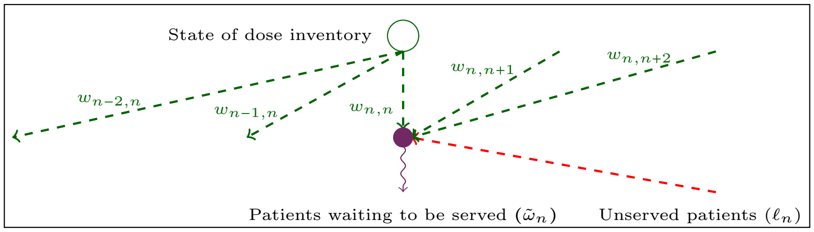


Figure 2.9: Dose utilization and patients waiting for $n = 2$ with $L = 2$.

setup described in §2.4.2 was adopted.

2.4.4.1 Policy analysis and comparison:

The performance of policies presented is compared based on the total costs, OVW, and number of unserved patients. Figure 2.10 illustrates the cumulative distribution function (CMF) of the average total cost, number of unserved patients, and OVW over 30 replications.

1. *FMLS*: Results in Figure 2.10a indicate that the probability of achieving a certain cost target under FMLS policy is only slightly lower than the base policy. For example, the probability that the total cost is less than \$22,200 is 0.81 for base policy and 0.8 for FMLS policy. Moreover, the results in Figures 2.10b and 2.10c indicate that the number of unserved patients and OVW are comparable to the base policy, which can be attributed to the similarity in vial-opening decisions under base policy and the vial-opening windows set under FMLS policy; see Figure 2.6b. However, FMLS policy mandates using one-dose vials even when multiple patients arrive in the last time window, although using multi-dose vials is more economical. In summary, FMLS policy is easy to implement and fairly economical.

2. *SDS*: SDS policy increases the effective arrival rate in the 6 hours time window by allowing a fraction of patients, originally scheduled to attend in the last seventh and eighth hours of the initial 8-hour session, to attend a session during the first 6 hours. This increased arrival rate increases the number of vaccines used within the first 6 hours, but it has only a marginal impact on OVW and the fraction of patients left unserved. The increase in the number of vaccines used and the number of unserved patients results in a higher cost, which is demonstrated in Figure 2.10a.

   In this experimental setup with SDS policy, assume that an incentive of \$1 was provided to patients willing to be rescheduled earlier in the day. Assume that 65% of patients are willing to be rescheduled. This percentage could be higher if the incentive pay is bigger, which would result in an increase of vaccination coverage. The incentive pay of \$1 is expected to cover the cost of a snack and drink . This approach is motivated by an experience in Nicaragua, where, during 1985, the attendance in immunization sessions increased from 77% to 94% because food was served to patients waiting for vaccinations [87]. Using such an incentive makes economical sense when the total amount paid is smaller than the penalty associated with unserved patients. We re-ran our experiments to estimate the penalty from unserved patients as a function of the percentage of patients willing to move their appointment to earlier in the day. Figure 2.11a shows the relationship between the fraction of patients rescheduled and the corresponding unit incentive costs, which are calculated based on the penalty costs. For example, an incentive pay of \$4 is economical if at least 20% of the patients are incentivized to reschedule.

3. *Single vial size*: If a region must use a single multi-dose vial, then the choice of vial size should depend on patient arrival rate. In regions with higher patient arrival rates, it is cost efficient to use ten-dose vials. In regions with lower patient arrival rates, one- and two-dose vials are more economical compared to using ten-dose vials, since the latter results in higher unused doses, which contribute to OVW. Figure 2.10 summarizes these observations for a region with high patient-arrival rate, like Chittagong. The results in Figure 2.10b show that, with a probability 0.95, the total number of unserved patients, for the duration of the planning horizon, is as high as 45 when two-dose vials are utilized. This number is smaller when ten-dose vials are uzed instead. The use of vials of small size reduces OVW (it is in fact 0 for one-dose policy). Since the penalty for letting a patient go unserved is higher than the penalty for OVW, the total cost for the ten-dose policy is lower than the two-dose policy. In fact, the probability that the total cost is smaller than $22,000$ is 0.7 for the ten-dose policy. This probability is 0.81 for the optimal policy. This indicates that a policy that relies on the use of ten-dose vials only is near optimal for this region.

4. *WO*: This policy, by requesting that patients wait until a multi-dose vial is opened, contributes to reducing OVW and associated costs. WO policy reduces the number of unserved patients significantly; see Figure 2.10b, although a few patients do leave if the waiting times are too long. This policy makes the utilization of larger-sized vials economical, as seen in Figure 2.10c. For these reasons, the total cost of WO policy is smallest; see Figure 2.10a. Moreover, Figure

2.11b shows that WO policy, in an effort to reduce the total waiting time, recommends health care practitioners use ten-dose vials for patients arriving in the early hours of the session, and smaller-sized vials near the end.

In order to compare the proposed policies with our base policy, we conducted a two-sample paired t-test on results obtained over 30 replications. These results are summarized in Table 2.5. The second column of the table presents the estimated difference in total costs between the base and heuristic policies. In this table, the negative numbers indicate that the base policy has a lower total cost. The table also presents the 95% CI of the difference in objective function values and the corresponding $p-$values. A $p-$value greater than 0.05 indicates that the null hypothesis, which states that the objective function values are statistically indistinguishable from one another, cannot be rejected at a 95% confidence level. These results indicate that WO is the only policy that has a total cost lower that the base policy. Moreover, the $p-$values indicate that the total cost of FMLS policy is statistically indistinguishable from the total cost of the base policy. This is further corroborated by the fact that 95% CI for estimated difference includes zero. However, the null hypothesis is rejected for other policies, meaning that they are significantly different from the base policy.

Our analysis suggests that the WO policy should be implemented during outreach sessions because it achieves the highest coverage with the least cost. However, implementation of this policy must take into consideration economic and demographic characteristics of the population in a region which impact patients' willingness to wait. The success of the FMLS policy applied depends on selecting the appropriate time periods to determine which different-sized vials should be opened. The ease of implementing FMLS policy can be enhanced by combining the FMLS and WO policies. Vaccine coverage rates are greatest under the SDS policy, which thrives with appropriate incentive programs. The additional costs incurred from these incentives are justified when the tangible benefits of disease eradication considered and realized. Finally, in regions where vaccines are distributed in a single multi-vial size, choosing the right size is important. Factors to weigh include population size, birth rate, and the number of clinics that organize outreach sessions in the region.

Table 2.5: Two-sample paired t-test for differences between the base and heuristic policies in Chittagong.

Two-sample t-test	Estimated difference (%)	Minimum difference (%)	Maximum difference (%)	95% CI for difference	p–value
Base, FMLS	−0.2	−0.4	−0.1	$[-26.4, 0]$	0.19
Base, SDS	−18.2	−19.1	−17.6	$[-3,946, -3,892]$	0.000
Base, WO	0.2	0.1	0.3	$[26.9, 58.9]$	0.00
Base, 1 dose	−73.5	−74.1	−73.2	$[-15,861, -15,819]$	0.00
Base, 2 dose	−68.9	−69.5	−68.7	$[-14,876, -14,834]$	0.00
Base, 10 dose	−0.1	−0.5	0.0	$[-46.9, -6.0]$	0.01

2.5 Conclusion and Future Research

This study presents a 2-SIP model for inventory replenishment and the administration of childhood vaccines in targeted immunization outreach sessions. To the best of our knowledge, this is the first stochastic optimization model which captures the relationships that exist among these decisions . The model presented here minimizes replenishment and OVW costs. OVW represent the doses not used by the end of an immunization session. Unlike related works in the literature and the current practice that relies on the use of a single multi-dose vial, this study models the performance of an inventory replenishment policy that allows a mix of multi-dose vials for vaccination. An extensive numerical analysis is conducted to evaluate the performance of the proposed policy and to compare to other simple-to-implement vaccine administration policies.

In order to solve the proposed 2-SIP, the LS method is extended by incorporating GMI and MRI cuts in the first-stage problem. Via an extensive numerical study we show that the proposed algorithm is scalable; it outperforms the LS method by providing high quality solutions in a much shorter CPU time.

After developing a case study using real-world data from Bangladesh, a sensitivity analysis is conducted to evaluate the system's behavior. Our observations can be summarized as follows:

1. Population size impacts decisions about the mix of multi-dose vials to use in a region. The use of multi-dose with complementary single-dose vials is recommended in highly populated regions.

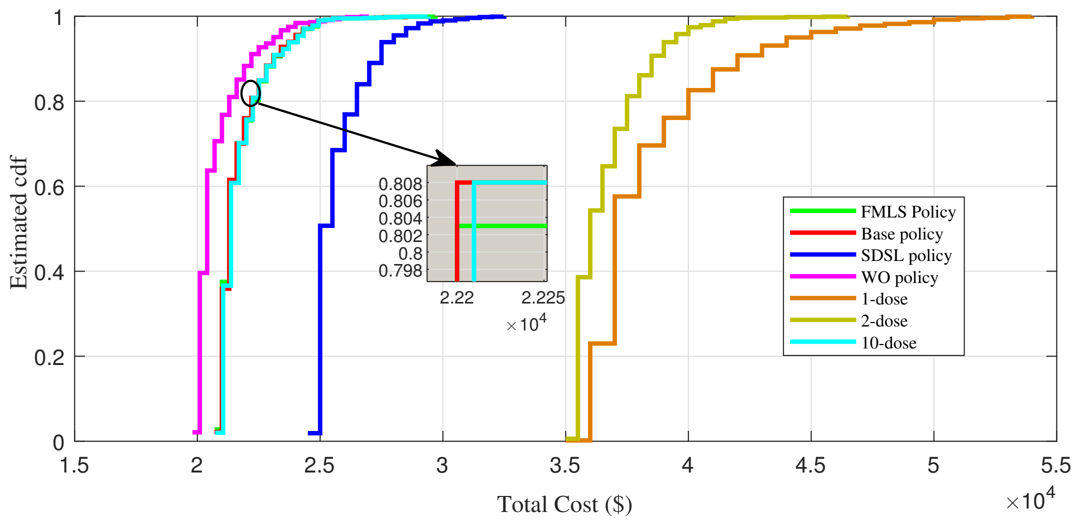
2. Vaccine purchasing costs impact the achievable immunization levels within a given budget limit. Thus, the models presented here aid policy makers in negotiating the necessary subsidies to achieve the targeted vaccination coverage levels.

3. Session length impacts replenishment costs, the number of patients served, and OVW. Session

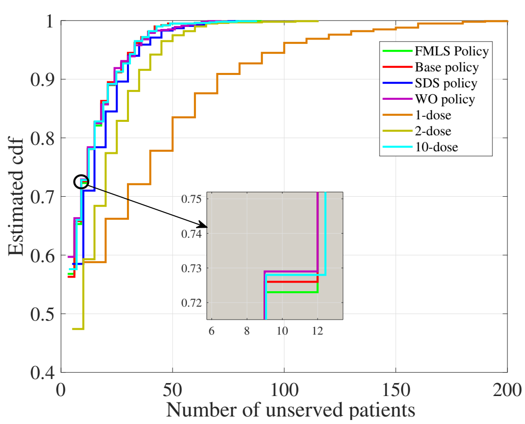
lengths that are the same length as the vaccine's safe use time do minimize the total cost. Short sessions in sparsely populated regions also minimize costs and OVW.

These observations motivated the design of vaccine administration policies that are simple and economical. Numerical results demonstrate that the WO policy has the lowest total costs and therefore, is highly recommended. For highly populated and well-connected regions, FMLS policy works well since it provides high vaccine coverage level at a lower cost. Moreover, in regions that can only use single multi-dose vials, the decision about the size of a vial to use should be based on population size, birth rate, and the number of clinics in the region (see Table (2.5)).

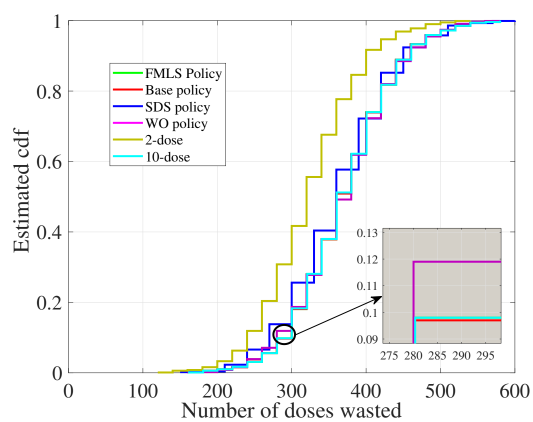
We plan to extend this research in the following ways. First, since no clear guidelines determine the number of scenarios used, investigating applications of sequential sampling algorithms, such as the two-stage stochastic decomposition (SD) method [62], is necessary. The SD method was originally designed for 2-SLPs and does not require *a priori* selection of scenarios. Since the model presented here is an MILP, we plan to extend this method to accommodate discrete decision variables in the first stage. Second, this proposed model identifies inventory replenishment decisions of a single outreach session, so we plan to extend this model to consider multiple sessions organized by the same clinic, as well as multiple clinics within a region. We expect that these clinics will coordinate their own decisions about inventory and operating hours to minimize costs and OVW. Third, we plan to develop an extension of the proposed model to aid replenishment decisions in clinics that handle different types of vaccines with different safe use times, such as liquid with 28 days and lyophilized with 6 hours.



(a) Total cost.

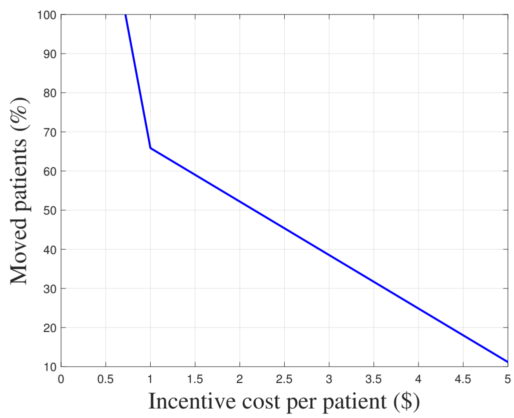


(b) Unserved patients.

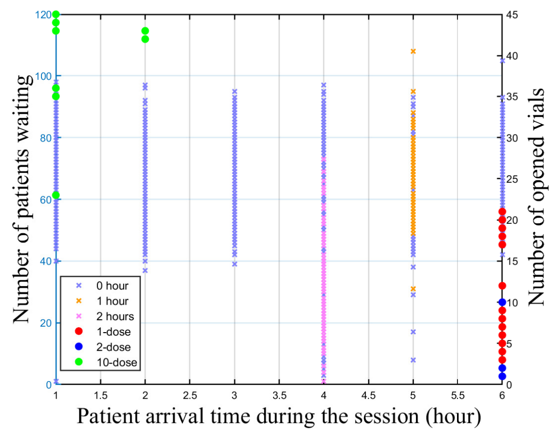


(c) OVW.

Figure 2.10: Comparison of heuristic policies.



(a) Incentive cost versus the percent of people who attend.



(b) The hourly consumption schedule and waited people.

Figure 2.11: SDS and WO analysis results.

Chapter 3

A Two-Stage Stochastic Model for Joint Pricing Inventory Replenishment with Deteriorating Products

3.1 Introduction

This work focuses on the analysis of joint pricing and inventory management decisions for age-dependent perishable products in a periodic-review inventory system. An age-dependent perishable product loses its quality/value and quantity the longer it stays in the shelves, and it is disposed after a certain time. Examples of such products are fruits, and vegetables in grocery stores; or baked goods in bakeries. Since demand for perishable products is price sensitive, businesses offer price markdowns to stimulate demand for products which are approaching the end of their shelf life. The same product type, at different stage of shelf life and price, coexist in the market. Thus, different from the case when the inventory does not perish, models for joint pricing and inventory management of perishable products take into account the age of inventories. The proposed model supports replenishment and pricing decisions with the aim of maximizing profits while minimizing waste and disposal cost for perished products.

The motivation for this research are the opportunities we see to reduce waste of perishable products by integrating inventory management policies with pricing and markdown decisions; and to manage the profitability of perishables in retailing using pricing. For example, ERS of USDA estimated that, in 2010, 45 billion pounds of available food at retail stores in the United States was wasted [25]. Moreover, about 40% of the annual agricultural production was wasted while 17% of the population was undernourished in 2014 in India [21]. In addition to social impacts, wastage has negative environmental impacts. For example, in the United States food waste accounts for 10% of the energy supplied, 80% of consumed water, and 50% of land used [56].

A number of studies in the literature propose control policies to optimize the performance of inventory management systems. Early works in this area assumed that the product has a single static price which is exogenous to the inventory management problem [48]. In these works, inventory management and pricing decisions were made in isolation for two main reasons. First, the data available was insufficient to characterize the impact that price and markdowns have on demand. Second, the benefits from improved inventory management were perceived as additive to the benefits from pricing. In recent works, we have seen an increased interest on integrating pricing and inventory management decisions in retail and other industries [49]. This is due to the increased availability of data and the development of decision support tools for analyzing the impacts of pricing and markdowns on inventory management decisions. Technologies available today (such as, point-of-sale data and loyalty programs) provide companies with ample data about customers' purchasing history and preferences which can be used to estimate the impact of price on customer demand. Indeed, inventory replenishment strategies control the supply side of a business whereas pricing policies control the demand side. Integrating these decisions mitigates the risk of mismatch in supply and demand, and increases profitability [128].

It has been noted in the literature [130, 124] that customers are willing to pay less for perishable products which are approaching their expiration date since they may perceive these products as of lower quality. Many industries nowadays are dealing with shorter product life-cycles. Thus, even when a product's quality is not impacted (such as smart phones), the willingness of customers to pay the full price decreases when a new version of the product appears in the market. Businesses use price markdowns to reduce losses from wastage of perishable products. Supermarkets and bakeries, such as Walmart, Publix, Foodstuffs, etc,. have discount racks. These retailers markdown vegetables, fruits, dairy, and bakery goods by as much as 40% to 50%.

Replenishment decisions for perishable products are challenging due to uncertainties of customer demand which are caused by limited product shelf life and price markdowns. These decisions are further complicated by uncertainties in the amount and quality of shipments received from suppliers. Dual-sourcing is a policy often used by retailers to mitigate the risks of supply and demand uncertainties [84].

This study proposes a stochastic optimization model that integrates inventory replenishment and pricing decisions for age-dependent perishable products in a periodic-review inventory system. We express the stochastic demands for new and old products as linear functions of their prices. The model determines suppliers and corresponding replenishment quantities which balance product waste (as a result of too much inventory) and product shortage (as a result of too little inventory). We consider a single markdown which is typically the case in grocery stores. This perishable product is produced by two suppliers. One of the suppliers is reliable but expensive. The other supplier, is not reliable since it has limited and varying capacity. However, this supplier is less expensive. The proposed model captures the trade-offs that exits between timing of markdown and waste, size of markdown and waste, size of markdown and profits, and supply chain costs and reliability.

The proposed model captures many interrelated and conflicting relationships. We develop a case study and conduct an extensive sensitivity analysis to demonstrate the nature of these relationships. This analysis reveals trends which are neither intuitive, nor easy to estimate. We expect that our results will provide insightful perspectives to managers making inventory replenishment decisions for perishable products.

3.2 Literature Review

The main contributions of this study are related to three streams of literature: stochastic inventory management of perishable products, joint pricing and inventory management and dual sourcing.

The literature on inventory management of perishable products with stochastic demand is considerably broad. [101] provided a review of early work in this area. More recently [102, 71, 17, 70] indicate an increasing interest in this topic due to some new and interesting applications. This literature distinguishes perishable products with fixed shelf life from deteriorating products.

Deteriorating products are further classified into perishable products with random life expectancy, and age-dependent perishable products. Age-dependent perishable products' quality/value and/or quantity degrade over time at a constant or dynamic rate based on age.

Early works in the field was focused on managing the inventory of perishable products with fixed shelf-life and stationary demand [74]. Stationary demand models assume that demand in successive periods are represented by independent and identically distributed random variables. Periodic review models, such as, $(R,\ S)$ or $(R,\ s,\ S)$, and continuous review models are typically used to model problems with stationary demand [31, 36, 83]. Using these models when demand is non-stationary would result in waste –when demand is low–, or stock outs –when demand is high [131]. Other approaches used in the literature to model the problem include Markov process [30] or dynamic programming [38, 81, 72]. [116] propose a model for inventory management of perishable products with fixed shelf-life and non-stationary demand.

Early works in inventory management modeling assumed that customer demand followed exogenous distributions [149, 118]. Typically, the objective of these models was to improve operational efficiency by minimizing expected costs. Another common assumption was that sales price did not change with time, was exogenous and imposed [26]. A few studies, however, pointed to the importance of modeling jointly pricing and inventory management decisions [48, 90, 37].

Modeling inventory replenishment for age-dependent perishable products with stochastic demand is a challenge. [66] proposed an ELS model for the deterministic version of this problem and developed a dynamic programming solution approach. Extensions of this model focus on deterministic demand. This is mainly because it is challenging to capture the age of the inventory in a stochastic setting. The work of [141] is one of the first studies that discusses a joint pricing and inventory replenishment problem for perishable products with stochastic demand. [141] extends the news-vendor model by incorporating pricing as a decision variable. [95] extends this model by considering a price-dependent demand (x) which is affected additively by a random term (ϵ) independent of price, i.e., $x = \mu(p) + \epsilon$, where $\mu(p)$ is mean demand as a function of market price p. Their analysis concluded that, for additive demands (i.e., $x = a + bp$) the optimal price in the stochastic setting is always lower than the optimal price in the deterministic setting. [76] extend this model considering price-demand relationships of various levels of complexities. These models determine a single price and order quantity in a single period setting. Thus, the shelf life of these products is fixed to one period. Work by [139] and [27] consider joint pricing and inventory management for deteriorating products. Both works assume a constant decay rate and derive a fixed-pricing policy.

The news-vendor model has been extended to consider multiple period planning horizons. For example, [31] present a joint pricing and replenishment model for a perishable product over a multi-period planning horizon. The authors first consider the single-period problem and identify optimality conditions for the case when demand function is linear. Next, the authors propose a dynamic programming model for the multi-period problem. The number of states in a dynamic program typically grows exponentially with the problem size. For this reason, only problems of small size are solved optimally. To solve larger problems, [31] propose heuristic approaches that provide approximate solutions. [36] provide a structural analysis of a multi-period, periodic review model. The authors also develop analytical bounds for the optimal order-up-to level, and propose a heuristic policy that applies to problems with stationary and non-stationary demand. Other notable works are by [67] who formulate a multi-period dynamic pricing and inventory model for perishable products with one-period shelf life and strategic customers; [38] who identify optimality conditions for a pricing an inventory allocation problem with two-period shelf life. The authors propose heuristics to identify quality decisions for products with longer shelf life. Recently, [72] use a dynamic programming model to identify optimal order quantity and price for a perishable product with fixed shelf life. Works by [91, 98] solve a joint pricing and inventory problem for deteriorating products with constant perishability rate. Note that, these models listed here either assume a fixed shelf life or constant perishability rate.

The following papers focus on perishable products with random life expectancy. Work by Mandal and Pal (1998), which was further investigated by Wu and Ouyang (2000) and then extended by Wu (2001), focuses on inventory management of a perishable product with ramp type demand function and Weibull distributed deterioration rate. [146] extend these models by considering a general demand rate function and Weibull distributed deterioration rate. To the best of our knowledge, there are no papers in the literature that address joint pricing and inventory replenishment for products with random life expectancy.

A number of works in the literature focus on dual sourcing as a strategy to reduce risk in the supply chain [136, 15]. [57] integrate supplier selection and inventory replenishment decisions via a mixed integer nonlinear programming model. These suppliers differ by transportation costs and lead times. Stochasticity of lead time impacts suppliers' reliability. This study does not consider pricing decisions and the product is not perishable. [12] develop a two-stage stochastic programming models to aid supplier selection decisions in the processed food industry. The first stage problem decides about the quantities to procure from each supplier. The second stage problem identifies procurement and transportation plans. This model assumes uncertain supply and demand. However, this study does not consider pricing decisions. A few studies model joint pricing and inventory management problems with supply uncertainty and non-perishable products [82, 28, 50, 35, 128, 29]. To the best of our knowledge the problem of joint pricing and inventory management for deteriorating products under demand and supply uncertainty and dual sourcing has not been investigated in the literature.

3.3 Our Contributions

Modeling: We propose a two-stage stochastic optimization model that integrates inventory replenishment and pricing decisions for an age-dependent perishable product in a periodic-review inventory system. We express the stochastic demands for new and old products via linear functions of their prices. We consider markdowns for older products to stimulate sales and minimize waste. The model uses dual sourcing to mitigate supply chain risks due to random supply. As indicated by our extensive literature review, this problem -although relevant for grocery stores and bakeries- has not been studied in the literature for these reasons. First, this problem is challenging to solve. Second, modeling the relationship between demand and price requires extensive amount of data. Third, due to lack of technical support, businesses favored simple models which provided simple rules of thumb that are easy to implement. However, current advancements in technology (such as RFID tags, point-of-sale data, loyalty programs) allow businesses to track the age of products and provide the means to collect ample amount of data necessary to estimate the impact of pricing on sales; etc. Thus, businesses have the means to employ relevant models and decision support tools to help with pricing decisions which better align demand and supply. **Solution approach:** We propose a solution approach for a two-stage stochastic, bilinear model with linear recourse. Our algorithm is an extension of the Benders-decomposition method. We approximate the non-linear first stage problem using extensions of McCormick relaxation method. Our approach significantly reduces the computational time as compared to using commercial solvers.

3.4 Problem Statement and Formulation

3.4.1 Notations and Assumptions

We consider a single retailer who makes inventory replenishment decisions for one perishable product in a periodic-review inventory system. We assume that a discrete and finite planning horizon of length T is a typical one and repeats itself over time. Thus, all problem data are assumed cyclic with cycle length equal to T ($\psi_{T+t} = \psi_t$ for $t = 1, \ldots, T$, where ψ_t is a problem parameter in period t). As a result, the inventory replenishment decisions and inventory levels are cyclic.

This retailer receives shipments from $\mathcal{I}$ suppliers. Each supplier i $(= 1, \ldots, \mathcal{I})$ charges a fixed cost (s_i) per shipment and a variable cost (r_i) per unit of product delivered. The fixed cost is due to order replenishment and loading/unloading; and the unit cost is due to purchasing and transportation. We assume zero transportation lead time since refrigerated transportation vehicles are typically used to deliver perishable products, thus, we assume no deterioration during lead time.

We assume that the longer a unit of this product is carried in inventory, the faster it deteriorates. This is indeed the case for most practical applications [66, 112]. Let $\rho_{t\tau}$ denote deterioration rate in period τ of a unit which was received in period t. Thus, for $t \leq \tau \leq j$, we assume that $\rho_{t\tau} \leq \rho_{tj}$. Let $\kappa_{t\tau}$ represent the portion of a demand in period τ satisfied via a shipment received in period t. Thus, $\kappa_{t\tau} = \prod_{j=t}^{\tau-1}(1 - \rho_{tj})$ and $1 = \kappa_{tt} \geq \kappa_{t\tau} \geq \kappa_{tj}$ for $t \leq \tau \leq j$.

We consider a single price markdown of this perishable product. Thus, we assume that a product which has been in shelves for less than $l(< T)$ periods is "new" and sells at full price p^n. A product which has been in shelves longer is "old" and sells at a discounted price p^o. Products which have been in shelves for f periods of time ($\ell \leq f < T$), or longer, are disposed. It is expected that the discounted price of old products will impact the demand for new products, and vice-versa. We assume demands for new and old products in each time period are stochastic, independent and non-stationary. Within the same period however, demand for new and old products are interdependent and represented via equations (3.1) and (3.2).

$$\mathcal{D}^n(p^o, p^n, t) = \alpha^n - \beta^n p^n + \beta^o p^o + \omega_t^n, \qquad \forall t = 1, \ldots, T; \tag{3.1}$$

$$\mathcal{D}^o(p^o, p^n, t) = \alpha^o - \theta^o p^o + \theta^n p^n + \omega_t^o, \qquad \forall t = 1, \ldots, T; \tag{3.2}$$

Where,

$$\mu^n(p^o, p^n) = \alpha^n - \beta^n p^n + \beta^o p^o; \tag{3.3}$$

$$\mu^o(p^o, p^n) = \alpha^o - \theta^o p^o + \theta^n p^n; \tag{3.4}$$

estimate the mean demand for new and old products, and ω_t^n, ω_t^o are the corresponding random elements. We assume that these random elements are independent, with generic distributions and continuous density functions that may change with time. The coefficient $\beta^n(\beta^o)$ represents the sensitivity of the demand for new products to the price of new (old) products; and $\theta^o(\theta^n)$ represents the sensitivity of the demand for old products to the price of old (new) products. We assume that $\beta^n \geq \beta^o$ and $\theta^o \geq \theta^n$ to ensure that an increase in price will result in a decrease of the overall demand. The intercepts α^n, α^o represent the maximum demand for new and old products. In summary, equations (3.3) and (3.4) enforce the mean demand for new (old) products to decrease with the price of new (old) product and increase with the prices of old (new) [48]. Other studies in the literature use equations (3.1) and (3.2) to represent stochastic demands for substitute products [48, 40].

3.4.2 Problem Formulation

We propose a joint pricing and inventory replenishment model for a perishable product with stochastic, non-stationary demand and age-dependent deterioration rate. We model this problem as a two-stage stochastic program in order to capture the stochastic nature of demand and its impact on the timing of decisions in the supply chain. We consider that supplier selection, inventory replenishment schedule, and pricing decisions are made a *priori* (in the first stage) and before demand is realized. Therefore, these decisions are non-anticipative in nature. The assumption that a replenishment schedule is identified prior to demand realization is not restrictive since the model is uncapacitated. Pricing decisions only impact the mean demand for old and new products. However, inventory replenishment decisions for new and old products follow demand realization. Therefore, these decisions are made in an adaptive manner.

First stage decision variables: Supplier selection and replenishment schedule decisions are captured via y_{it}. This is a binary decision variable which takes the value 1 if supplier i is selected to replenish the inventory in period t ($\mathcal{I} = 2$), and takes the value 0 otherwise. This variable is used to calculate the total fixed order cost during the planning horizon $\sum_{i=1}^{\mathcal{I}} \sum_{t=1}^{T} s_i y_{it}$.

Other first stage decision variables are the price of new and old products p^n, p^o. The relationship between prices and mean demands are captured via equations (3.3) and (3.4). The total cost due to first stage decisions is

$$\zeta(p^n, p^o, y_{it}) = \sum_{t=1}^{T} \sum_{i=1}^{\mathcal{I}} s_i y_{it} - (p^n \mu^n + p^o \mu^o) * T.$$

Since there is a one-to-one relationship between price and mean demand, we substitute mean demand with the corresponding functions defined in (3.3) and (3.4). Thus, the total cost becomes

$$\zeta(p^n, p^o, y_{it}) = \sum_{t=1}^{T} \sum_{i=1}^{\mathcal{I}} s_i y_{it} - (\alpha^n p^n + \alpha^o p^o - \beta^n (p^n)^2 - \theta^o (p^o)^2 + (\beta^o + \theta^n) p^n p^o) * T.$$

We use a single decision vector $x \in \mathcal{X}$ to collectively denote the first stage decision variables (i.e., $x = (p^n, p^o, y_{it} \ \forall i = 1, \ldots, \mathcal{I}, \ t = 1, \ldots, T)$). Thus, the corresponding total cost is denoted by $\zeta(x)$.

Second stage decision variables: Inventory replenishment decisions are denoted by $q_{it\tau}$ and represent how much of the demand in period τ is satisfied via shipment from supplier i in period t. These decision variables facilitate tracking the age of the inventory (the age of a product in period τ equals $\tau - t$). For example, q_{itt} represents how much of the demand in period t is satisfied via a shipment received in the same period; $q_{i,t-1,t}$ represents how much of the demand in period t is satisfied via 1 period old inventory, etc.

Let $c_{it\tau}$ denote the replenishment cost per each unit of $q_{it\tau}$, and h_τ denote the unit inventory holding cost in period τ. The unit replenishment cost consists of the unit delivery cost r_i and the total unit inventory holding cost for the period $(t, \ldots, \tau - 1)$ during which it was held in inventory: $c_{it\tau} = r_i + \sum_{s=t}^{\tau-1} h_s \kappa_{ts}$. Thus, the total replenishment cost is $\sum_{i,t,\tau} c_{it\tau} q_{it\tau}$. $\mathcal{S}_{it}$ is a random variable which represents the amount available at supplier $i \in \mathcal{I}$ in time period $t = 1, ..., T$.This quantity is a random parameter with specified distribution.

Other decision variables included in the model are u_t and z_t which represent amounts of shortage and wastage due to the stochastic nature of demand. The corresponding penalty costs are f_t and g_t. The service level is defined as the probability of not having a stock-out in a replenishment cycle. One can adjust (increase) the value of f_t to reduce shortages, thus, improve service level.

Similarly, the value of g_t impacts wastage.

Two-stage stochastic model formulation: The objective is to minimize the total cost of the first stage decisions and the expected cost of the second stage decisions. That is,

$$\min \qquad F(x) = \zeta(x) + \mathbb{E}\{\phi(x,\omega)\}, \tag{3.5a}$$

$$s.t. \qquad x \in \mathcal{X} = \{(p^n, p^o, y_{it}) | p^n, p^o \geq 0,\ y_{it} \in \{0,1\}\ \forall i \in \mathcal{I},\ t = 1,\dots,T\} \tag{3.5b}$$

The objective function $F(x)$ includes bilinear terms, thus, the problem is not convex.

The recourse function $\phi(x,\omega)$ for a realization $\tilde{\omega} = (\tilde{\omega}^n, \tilde{\omega}^o, \tilde{\mathcal{S}})$ is given by:

$$\phi(x,\tilde{\omega}) = \min \ \sum_{t=1}^{T} \left[f_t u_t + g_t z_t - p^n \tilde{\omega}_t^n - p^o \tilde{\omega}_t^o + \sum_{i=1}^{\mathcal{I}} \sum_{\tau=t}^{T} c_{it\tau} q_{it\tau} \right] \tag{3.6a}$$

$s.t.$

$$\sum_{i=1}^{\mathcal{I}} \sum_{t=[T+(\tau-\ell)+1]}^{\tau} \kappa_{t\tau} q_{it\tau} + u_\tau = \tilde{\mathcal{D}}_\tau^n \qquad 1 \leq \tau \leq T, \tag{3.6b}$$

$$\sum_{i=1}^{\mathcal{I}} \sum_{t=[T+(\tau-\ell)+1]}^{[T+(\tau-\ell)]} \kappa_{t\tau} q_{it\tau} + u_\tau - z_\tau = \tilde{\mathcal{D}}_\tau^o \qquad 1 \leq \tau \leq T, \tag{3.6c}$$

$$\kappa_{t\tau} q_{it\tau} - \tilde{\mathcal{S}}_{it} y_{it} \leq 0 \qquad 1 \leq i \leq \mathcal{I}; 1 \leq t \leq T, 1 \leq \tau \leq T, \tag{3.6d}$$

$$q_{it\tau}, u_\tau, z_\tau \geq 0, \qquad 1 \leq i \leq \mathcal{I}; 1 \leq t \leq T, 1 \leq \tau \leq T. \tag{3.6e}$$

Where $\tilde{\mathcal{D}}_\tau^n = \alpha^n - \beta^n p^n + \beta^o p^o + \tilde{\omega}_\tau^n$ and $\tilde{\mathcal{D}}_\tau^o = \alpha^o - \theta^o p^o + \theta^n p^n + \tilde{\omega}_\tau^o$.

For convenience, we propose the following notation $[t] = (t+1) \bmod T - 1$ and $[0] = T$. Function (3.6a) minimizes the total cost of inventory replenishment, shortage and wastage. This function also maximizes the revenue gain due to the stochastic nature of demand. Constraints (3.6b) are the inventory balance constraints for new products, and (3.6c) for old products. Note that, $q_{it\tau}$ represents the amount received from supplier i in period t to satisfy demand in period τ. However, $(1 - \kappa_{t\tau})q_{it\tau}$ units of inventory are wasted due to storage from periods t to τ and the remaining amount $\kappa_{t\tau} q_{it\tau}$ represent the demand met in period τ. Shortage and wastage occur due to the random nature of demand and are captured via u_τ, z_τ. Constraints (3.6d) set an upper bound on the amount of inventory. In the case when a shipment has been delivered in period t ($y_{it} = 1$), the upper bound is the total amount available at supplier i in period τ; otherwise ($y_{it} = 0$), the upper bound is zero. Constraints (3.6e) are the non-negativity constraints.

In this two-stage stochastic programming model, (3.5) is the first-stage problem and (3.6) is the second stage problem. The decision variables of the first-stage problem, $x \in \mathcal{X}$, are a mix of binary and continuous variables; however, the decision variables of the second stage problem are continuous.

Formulation (3.6) has uncertain parameters in the right-hand-sides of equations (3.6b) to (3.6d) and in the objective function. We reformulate the second stage problem in order to have uncertain parameters only in the right-hand-sides of the constraint sets. This reformulation is a two-stage stochastic program with fixed recourse.

$$\phi(x, \tilde{\omega}) = \min \ \eta \tag{3.7a}$$

s.t.

$$(3.6b), ..., (3.6e)$$

$$\eta - \textstyle\sum_{t=1}^{T} \left[f_t u_t + g_t z_t + \sum_{i=1}^{\mathcal{I}} \sum_{\tau=t}^{T} c_{it\tau} q_{it\tau} \right] \geq \sum_{t=1}^{T} \left[-p^n \tilde{\omega}_t^n - p^o \tilde{\omega}_t^o \right] \tag{3.7b}$$

3.5 Solution Approach

Two-stage stochastic programs with recourse provide a powerful tool to model problems arising in the areas of energy planning, manufacturing, supply chain, etc [20]. The special structure of these models enables the use of decomposition-based algorithms. Many decomposition-based algorithms, so-called scenario-based decomposition methods, use scenarios to capture uncertainties.

This model captures the stochastic behavior of supply via $\mathcal{S}_{it}$ and the stochastic behavior of demand for new and old products via ω_t^n, ω_t^o. These variables represent deviations of demand for new and old products from the corresponding mean demands. We discretize these random variables by creating S scenarios. Let $\Omega = \{\tilde{\omega}_1, \tilde{\omega}_2, \ldots, \tilde{\omega}_S\}$ be the set of scenarios and $\mathcal{P}_s$ be the probability associated with each scenario $s = 1, \ldots, S$. Thus, we can represent the expectation function in (3.5) as:

$$\mathbb{E}\{\phi(x, \omega\} = \sum_{s=1}^{S} \mathcal{P}_s \phi(x, \tilde{\omega}_s). \tag{3.8}$$

This discretization facilitates the formulation of the corresponding deterministic equivalent of (3.5), which can be solved using deterministic, global optimization algorithms, such as, branch-and-reduce [129], nonconvex outer approximation methods [73], and branch and bound [9]. However, these methods are not practical for large scale problems that have a large number of scenarios. This is why we use Benders decomposition algorithm explained in Section (3.5) which is an iterative procedure that generates an outer approximation of the expected recourse function [19].

The recourse function (3.8) can be replaced by a Monte Carlo estimate assuming that all scenarios have the same probability of $\mathcal{P}_s = 1/S$. This yields the SAA problem which follows:

$$\min \quad {}_{x \in \mathcal{X}} F_S(x) = \zeta(x) + \tfrac{1}{S} \textstyle\sum_{s=1}^{S} \phi(x, \tilde{\omega}_s), \tag{3.9a}$$

$$s.t. \quad x \in \mathcal{X} = \{(p^n, p^o, y_{it}) | p^n, p^o \geq 0,\ y_{it} \in \{0,1\}\ \forall i \in \mathcal{I},\ t = 1, ..., T\} \tag{3.9b}$$

Recall that, function $\zeta(x)$ contains the bilinear term $p^n p^o$ and the quadratic terms $(p^n)^2, (p^o)^2$. Due to the bilinear terms $p^n p^o$, this function is neither convex nor concave. The following are a few examples of convex relaxation methods that are used in the literature for bilinear terms: McCormick relaxation [94], outer linearization [129], and α-based Branch and Bound [10]. In Section (3.5.2) we present a relaxation of the problem via McCormick envelopes. This leads to a linear relaxation of nonlinear functions in the first-stage problem. In Section (3.5.3) we use an extension of the McCormick relaxation algorithm which leads to tighter relaxations of nonlinear functions in the first-stage problem. Section (3.5.4) presents an extended Benders decomposition algorithm which employs these relaxations.

3.5.1 Benders Decomposition Algorithm

Let $\mathcal{X}^k$ represent the feasible region of problem (3.9) during the k^{th} iteration of the Benders decomposition algorithm. In the first iteration of this algorithm the feasible region $\mathcal{X}^0$ includes only constraints (3.9b). In the k^{th} iteration, the feasible region $\mathcal{X}^k$ comprises of these constraints as well as additional affine functions. The maximum of the affine functions approximates the expected recourse function. These affine functions are generated using the optimal dual solutions to the second-stage problem in (3.7). Let x^k be the optimal/near optimal solution to the following relaxed first stage problem:

$$\min_{(x,\eta) \in \mathcal{X}^k} F_k(x) = \{\zeta(x) + \eta\} \tag{3.10}$$

where, η is an auxiliary variable employed by the affine functions.

For a given x^k and a realization $\tilde{\omega}_s \in \Omega$, we obtain dual solutions to the second-stage problem $\phi(x^k, \tilde{\omega}_s)$. Let λ, ${\gamma_t^n}^k, {\gamma_t^o}^k$, ${\iota_{it\tau}}^k$, be the dual optimal solutions to (3.7b), (3.6b), (3.6c), and (3.6d) in iteration k. Then, the lower bounding affine function (optimality cut) is obtained as follows:

$$l_{opt}^k(x,\eta) := \frac{1}{S}\sum_{s=1}^{S}\left[\sum_{i\in\mathcal{I}}\sum_{\tau=1}^{T}\sum_{t=1}^{T}\tilde{\mathcal{S}}_{it}{\iota_{it\tau s}}^k y_{it}^k - \right.$$
$$\left. \sum_{t=1}^{T}\left(\tilde{\mathcal{D}}_{ts}^n {\gamma_{\tau s}^n}^k + \tilde{\mathcal{D}}_{ts}^o {\gamma_{\tau s}^o}^k\right) - \lambda_s \sum_{t=1}^{T}(p^n\tilde{\omega}_{ts}^n + p^o\tilde{\omega}_{ts}^o)\right] + \eta, \tag{3.11}$$

where $\tilde{\mathcal{D}}_{\tau s}^n = \alpha^n - \beta^n p^n + \beta^o p^o + \tilde{\omega}_{\tau s}^n$ and $\tilde{\mathcal{D}}_{\tau s}^o = \alpha^o - \theta^o p^o + \theta^n p^n + \tilde{\omega}_{\tau s}^o$. By adding (3.11) to $\mathcal{X}^k$ an iteration of Benders decomposition method is completed and $\mathcal{X}^{k+1}(x) = \mathcal{X}^k(x) \cup (l_{opt}^k(x,\eta) \leq 0)$. With help from slack variables (u_t, z_t) we ensure that the second-stage problem is always feasible. Therefore, the problem is relatively complete, and there is no need to generate feasibility cuts in this implementation of Benders algorithm.

Let x^k be the solution found during the k^{th} iteration of Benders algorithm. Then, the objective function value of (3.10), $F_k(x^k)$, provides a lower bound (L^k) for problem (3.5). The objective function value for (3.9), $F_S(x^k)$, provides an upper bound (U^k) for (3.5). Benders decomposition algorithm terminates when the relative error gap between U^k and L^k is less than a predefined error gap ϵ_1. Let $\hat{x} = x^k$ be the optimal solution and $F_S(\hat{x})$, be the corresponding optimal objective function value.

3.5.2 McCormick Relaxation Algorithm

The first-stage problem (3.9) contains the bilinear terms $p^n p^o$ in the objective function. Therefore, the problem is nonlinear and non-convex. We use McCormick envelopes to linearize these terms [94].

Let $\underline{p}^n \leq p^n \leq \overline{p}^n, \underline{p}^o \leq p^o \leq \overline{p}^o$; and $\overline{p}^n, \overline{p}^o$ be upper bounds on the variables. One can calculate $\overline{p}^n$ by maximizing function $\tilde{F}_S(x) = p^n - \frac{1}{S}\sum_{s=1}^{S}\phi(x,\tilde{\omega}_s)$ over the feasible region of formulation (3.9). A similar approach is used to calculate the upper bounds for p^o. Let $W = p^n p^o$ and substitute the bilinear terms in the objective with the newly defined variables. The corresponding McCormick relaxations of these terms are constraints (3) presented in the appendix (C).

The following is a relaxation of formulation (3.9).

$$\min \quad \hat{F}(x) = \sum_{i\in\mathcal{I}}\sum_{t=1}^{T} s_i y_{it} - (\alpha^n p^n + \alpha^o p^o - \beta^n (p^n)^2 - \theta^o (p^o)^2 + (\beta^o + \theta^n) W) * T + \eta, \quad (3.12)$$

$s.t.$

$$(3.9\text{b}), (3.11), (3)$$

Proposition 3.5.1 *A feasible solution $\hat{x}^*$ of* (3.12) *is also feasible to the master problem* (3.5)*.*

Proof: In formulation (3.12) the bilinear term $p^n p^o$ appears only in the objective function (3.12). Thus, McCormick envelopes provide an exact relaxation of the feasible region of (3.12). ■

Corollary 3.5.1 *Let $\hat{x}^*$ be a feasible solution to* (3.12)*. An upper bound to the master problem* (3.5) *is $\hat{U} = F_S(\hat{x}^*)$ which is found by evaluating the objective function of* (3.9) *at $\hat{x}^*$.*

Corollary 3.5.2 *Let $\hat{x}^*$ be a feasible solution to* (3.12)*. A lower bound to the master problem* (3.5) *is $\hat{L} = F_k(\hat{x}^*)$ which is found by evaluating the objective function of* (3.10) *at $\hat{x}^*$.*

3.5.3 Piecewise Linear Approximation Algorithm

McCormick relaxations of the bilinear term $p^n p^o$ is loose. An algorithmic approach proposed by [60], which is based on a piecewise linear relaxation of the bilinear terms using bivariate partitioning, provides tighter bounds. To create the piecewise linear approximation, we partition p^n, p^o into N^n, N^o equal and exclusive segments using $N^n + 1, N^o + 1$ grid points. Let d^n, d^o denote the length of each segment. Let $\mathcal{N} := \{N^n, N^o\}$ denote the collection of the number of grid points and $W = p^n p^o$. The set of constraints necessary to relax the bilinear term are (4), and presented in the Appendix.

The following is a relaxation of formulation (3.5).

$$\min \quad \hat{F}(x) = \sum_{i\in\mathcal{I}}\sum_{t=1}^{T} s_i y_{it} - (\alpha^n p^n + \alpha^o p^o - \beta^n (p^n)^2 - \theta^o (p^o)^2 + (\beta^o + \theta^n) W) * T + \eta, \quad (3.13\text{a})$$

$s.t.$

$$(3.5\text{b}), (3.11), (4)$$

Proposition 3.5.2 *A feasible solution* $\hat{x}^*$ *of* (3.13) *is also a feasible to the master problem* (3.5).

Proof: In formulation (3.13) the bilinear term $p^n p^o$ appears only in the objective function (3.13a). Thus, McCormick envelopes provide an exact relaxation of the feasible region of (3.5). ■

Corollary 3.5.3 *Let* $\hat{x}^*$ *be a feasible solution to* (3.13). *An upper bound to the master problem* (3.5) *is* $\hat{U} = F_S(\hat{x}^*)$ *which is found by evaluating the objective function of* (3.9) *at* $\hat{x}^*$.

Corollary 3.5.4 *Let* $\hat{x}^*$ *be a feasible solution to* (3.13). *A lower bound to the master problem* (3.5) *is* $\hat{L} = F_k(\hat{x}^*)$ *which is found by evaluating the objective function of* (3.10) *at* $\hat{x}^*$.

Proposition 3.5.3 *Formulation* (3.13) *provides a tighter relaxation of* (3.5) *than formulation* (3.12) *when either* $N^n \geq 1$, *or* $N^o \geq 1$, *or both* $N^n \geq 1, N^o \geq 1$.

Proof: In the case when $N^n = 0$ and $N^o = 0$, the piecewise linear approximation of $p^n p^o$ provides just the corresponding McCormick relaxation of this term. In this case, formulations (3.12) and (3.13) are exactly the same. Increasing the number of grid points N^n and/or N^o has no impact on the feasible region of (3.12) and (3.13). Based on *Lemma I* in [60], the uniform placement of grid points for bivariate partitioning minimizes the sum of squares of the maximum separation of the bilinear terms $W = p^n p^o$ from its linear relaxation in each segment. Therefore, increasing the number of grid points provides a tighter approximation of the objective function of (3.5). This results in finding solutions which provide tighter bounds for (3.5). ■

Algorithm (2) summarizes an algorithm we use to solve (3.5) by iteratively solving a sequence of relaxations (3.13). In each iteration the number of grid points is increased by 1. Let $\hat{x}^i \in \hat{\mathcal{X}}$ represent a solution to (3.13) during the i-th iteration of the algorithm. We use $\hat{x}^i$ to calculate lower and upper bounds to (3.5). The algorithm terminates when the relative gap between upper and lower bounds is less than a predefined error gap δ. Formulation (3.13) is a mixed integer, quadratic program which we solve via CPLEX.

3.5.4 Proposed Extended Bender's Decomposition Algorithm

Algorithm (3) summarizes the proposed extended Benders decomposition algorithm. During the k-th iteration of this algorithm, a relaxation of (3.5) is solved to obtain the first-stage variables x^k. Given x^k, formulation (3.7) of the subproblem is then solved for each scenario $\tilde{\omega}_s \in \Omega, s = 1, \ldots, S$. The first- and second-stage solutions are used to calculate lower and upper bounds to (3.5). This

Algorithm 2 Piece-wise linear approximation algorithm

Initialize:
  $i \leftarrow 0; \quad \hat{L} \leftarrow -\inf, \quad \hat{U} \leftarrow +\inf, N^n, N^o, \delta > 0$
while $\left|\hat{U} - \hat{L}\right| / \hat{U} > \delta$ **do**
  Solve (3.13) to obtain $\hat{x}^i$
  Set $\hat{L} = F_k(\hat{x}^i)$, $\hat{U} = F_S(\hat{x}^i)$
  $(i, N^n, N^o) + +$
end while

iterative procedure ends when the relative error ($|U^k - L^k|/U^k$) is smaller than some predefined bound (ϵ_2). If the relative error is greater than ϵ_2, an optimality cut (3.11) is added to the master problem, and the procedure repeats.

The solutions obtained from solving relaxations (3.12) and (3.13) of the master problem (3.5) in the initial iterations of the Benders algorithm are typically of low quality. This is mainly because initially, due to the limited number of optimality cuts (3.11), the feasible region of the master problem is not well defined. The quality of the solutions found improves as the algorithm progresses. Therefore, in order to reduce the running time of the algorithm, we initially solve (3.12) since it is computationally less challenging. As the algorithm progresses, we search for better solutions by solving (3.13).

Algorithm 3 Extended Benders Decomposition Algorithm

Initialize:
  $k \leftarrow 0; \quad L^0 \leftarrow -\infty, \quad U^0 \leftarrow +\infty, \epsilon_2 > \epsilon_1 > 0$
while $|U^k - L^k|/U^k > \epsilon_1$ **do**
  if $|U^k - L^k|/U^k \leq \epsilon_2$ **then**
    Call **Algorithm 1** to obtain x^k
  else
    Solve (3.12) to obtain x^k
  end if
  for $\tilde{\omega}_s \in \Omega, s = 1, \ldots, S$ **do**
    Solve (3.7)
    Obtain ${\gamma^n_{\tau s}}^k$, ${\gamma^o_{\tau s}}^k$ ${\iota_{it\tau s}}^k$, and λ_s^k
  end for
  Update (3.11)
  Set $L^k = F_k(x^k)$, $U^k = F_S(x^k)$
  $k + +$
end while

3.5.5 Evaluation of the Algorithm

Since selecting a good sample size S is critical and problem specific, we replicate the whole process (**Algorithm 2**) using different number of scenarios. Doing this, we increase the probability of obtaining a solution which is truly optimal to the original problem. Assume that Ω_S^r represents the set of scenarios and x_S^r, F_S^r are the optimal solution and corresponding objective function value obtained using the extended Bender's decomposition algorithm in replication $r = 1, \ldots, R$. We define $\overline{F}_S^R = \frac{1}{R}\sum_{r=1}^{R} F_S^r$ to be a point estimate of the optimal objective function value to the true problem (3.5).

In addition to this, the first-stage optimal solution of each replication (x_S^r) is evaluated by simulating the second-stage problem using a set of scenarios different than Ω_S^r. This process is termed the *out-of-sample* evaluation. This evaluation is terminated when $(1-\alpha)\%$ confidence interval is built on $F_{S'}^r$ with $\Omega_{S'} \neq \Omega_S$. Let $\Delta_r = F_S^r - F_{S'}^r$ be an estimator of the optimality gap with mean and variance of:

$$\overline{\Delta}^R = \frac{1}{R}\sum_{r=1}^{R} \Delta_r, \quad \sigma_\Delta^2 = \frac{1}{R(R-1)}\sum_{r=1}^{R}(\Delta_r - \overline{\Delta}^R)^2. \tag{3.14}$$

Since this estimator overestimates the optimality gap we can argue that it is a pessimistic gap.

3.6 Numerical Study

This section the results of our numerical study. In Section (3.6.1) we describe the experimental setup. In Section (3.6.2) we present a case study. In Section (3.6.3) we compare the performance of **Algorithm 2** to the solutions found by an exact approach when solving some special cases of problem (3.5), and in Section (3.6.5) we summarize the results of the sensitivity analysis. While the experiments in Section (3.6.6) consider one supplier the experiments in Section (3.6.7) examine dual sourcing.

3.6.1 Experimental setup

Algorithm 2 is implemented on a 64 bit Intel Core i4 processor @ 800 MHz with 8 GB RAM. The linear and mixed integer quadratic programs are solved using CPLEX callable subroutines. In our experimentations, we first obtain an optimal or near optimal first stage-solution with a relative

Table 3.1: Input data

	Suppliers	
Costs	1	2
s_i	$U[150, 250]$	$U[250, 350]$
r_i	$U[25, 30]$	$U[20, 25]$

error $\epsilon_1 \leq 0.01$ using **Algorithm 2**. Next, we verify the quality of these first-stage solutions by fixing them and solving subproblem (3.7) on a different set of scenarios. We use the same distribution to generate scenarios in both, optimization and evaluation processes. The evaluation process is replicated a few times ($R = 10$). In each replication, we use independent samples of random scenarios. Evaluation terminates when a 95% confidence interval is built on the recourse function values. The evaluation results are presented in our computations. The results presented in Figures 3.1 to 3.10 are averages over these replications.

3.6.2 Case Study

We construct our experimental study based on a dataset described in Table 3.1 and used in [114] who study a similar problem. This data considers 2 suppliers and a planning horizon that is (T =)7 periods long. The deterioration rate $\rho_{t,t+1} = 0.02$ and the unit inventory holding cost $h =$ \$1/(unit*time period). Products are marked down after being in the shelf for (ℓ =)3 days.

The random elements ω^n and ω^o follow truncated normal distribution ($[0, 150]$ and $[0, 175]$) with positive standard deviations $\sigma^n = 10$, $\sigma^o = 20$. Thus, demand $\mathcal{D}_t^n$ and $\mathcal{D}_t^o$ follow truncated normal distribution with constant standard deviations σ^n, σ^o. These inputs are taken from [40] who study a similar problem.

The penalties for shortage and wastage are initially set high ($f_t = g_t = 1,000 \ \forall t \in \mathcal{T}$) to minimize lost demand and wastage. We conduct sensitivity analysis to evaluate their impact on the performance of this system.

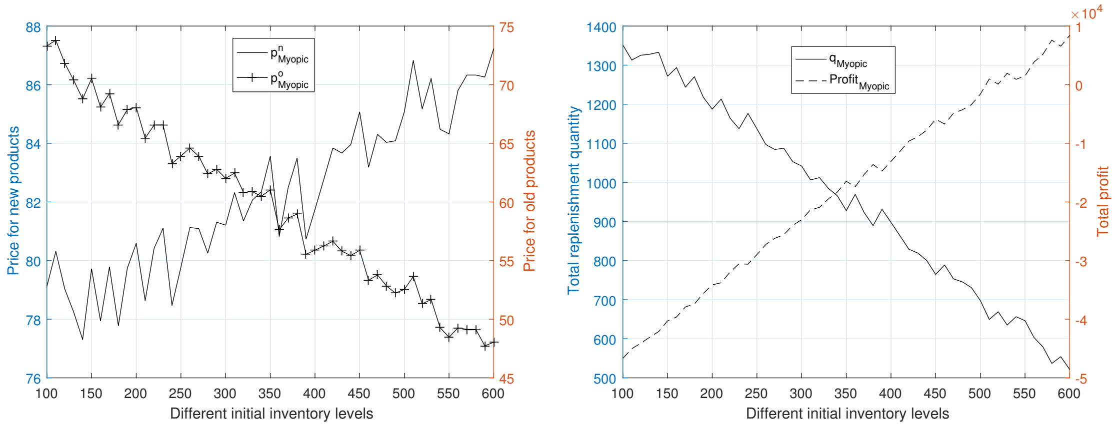
3.6.3 Model Validation

Existing models for joint inventory replenishment and pricing of perishable products make simplifying assumptions about deterioration rate and demand distribution. These models either assume that product shelf life is two periods long, or demand is stationary. In order to demonstrate the advantage of the proposed model we compare it to a recent model by [40].

[40] derived analytical results about the optimal inventory replenishment policy to a special case of our problem when the product shelf life is 2 periods long and the planning horizon is 1 period long, that is, $l = 2, T = 1$. In this model, the relationship between price and demand is represented using the same linear functions (3.1) and (3.2). Given the size of the initial inventory, the model decides on the replenishment quantity which minimizes total costs. [40] shows that the optimal policy derived for this problem is myopic, thus not optimal, when the product's shelf life and planning horizon are longer.

We conducted an experiment using the same data set and we let $l = 2, T = 1$ to mimic the problem setting of [40]. Figure 3.1a depicts the relationship between the optimal price of new and old products for different levels of initial inventory. It is observed that the price of old products decreases as the initial inventory level increases. Figure 3.1b depicts the relationship between the optimal price of new and old products for different levels of initial inventory. It is observed that the price of old products decreases and the price of new products increases as the initial inventory level increases. Figure 1b shows that the size of orders for new products decreases and the total profit increases as the initial inventory level increases. These observations are similar to the results presented in [40]; which is a validation of the model we propose.

We compare the results of an optimal solution from using Algorithm 2 to the results of implementing the myopic policy for the case when the product shelf life is 3 periods and planning horizon is 2 periods long. We test both models for the same levels of initial inventory. Figure 3.2a depicts the relationship between the optimal price of new and old products and the initial inventory level. This relationship is similar in both approaches. Figure 3.2b represents the relationship between replenishment quantity and initial inventory level; and profits and initial inventory level. Since the planning horizon is 2 periods long, both models try to minimize ending inventories. However, the replenishment quantities under the myopic policy are higher since this policy tries to ensure that the requirement about the initial inventory level for the second period is met. Figure 3.2c compares the profits from both models. These profits are consistently smaller for the myopic policy. These results demonstrate the value of the model proposed in this study, and it is expected that the benefits will be greater when solving problems that consider longer planning horizon and non-stationary demand.



(a) Pricing decisions (b) Ordering decisions and profits

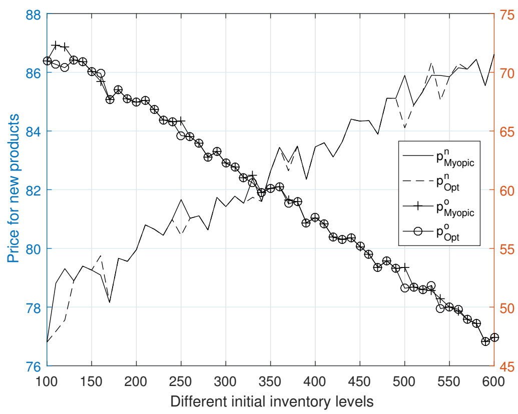
Figure 3.1: Evaluating solution quality of Algorithm 2 when $\ell = 2, T = 1$.

Table 3.2: Comparison of solution quality.

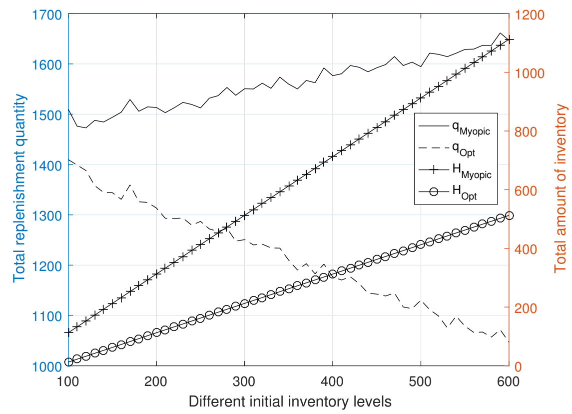
		Total Costs		CPU Time	
T	S	Algorithm 1	COUENNE	Algorithm 1	COUENNE
1	100	-46,663.91	-13,343.96	74	80
2	100	-230,544.38	-88,179.50	86	89
7	100	-310,747.81	-106,943.10	138	162

3.6.4 Analyzing the Performance of Algorithm 1

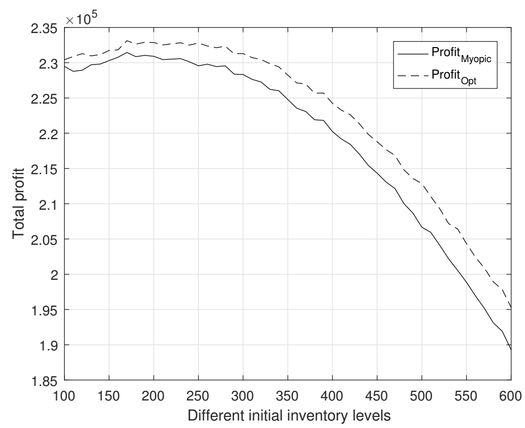
The goal of these experiments is to evaluate the performance of the piecewise linear approximation algorithm (Algorithm 1) which we use in line 4 of Algorithm 2 to solve model (3.13). We also solve model (3.12) using COUENNE [18], an open source software and compare their performance. Note that COUENNE does not guarantee the optimality. We conduct the experiments in rather small problem instances since COUENNE was not able to solve larger problem instances. Table (3.2) summarizes the objective function values (total costs) and the corresponding running times. The results indicate that implementing the piecewise linear approximation algorithm results in higher quality solutions compared to COUENNE.



(a) Pricing decisions



(b) Ordering decisions and inventory



(c) Profits

Figure 3.2: Comparing the quality of solutions from Algorithm 2 to the model by [40] for $\ell = 3, T = 2$.

3.6.5 Sensitivity Analysis

3.6.5.1 Scenario Selection:

To identify an appropriate number of scenarios for our experimentations we use two methods. These methods analyze the difference between the optimal solutions obtained when using different number of scenarios. The first approach assesses the impact of the optimal solution on the recourse objective function value. To accomplish this, during the evaluation phase we use the optimal solution and simulate the subproblem to find the recourse function value over the two sets of scenarios. We then perform a paired t-test on the corresponding objective function values. Table 3.3 presents $p-$values for each pair of scenarios tested. If the $p-$value is greater than 0.05 we conclude that the impacts of solutions $\hat{x}^S$ and $\hat{x}^{S'}$ ($S \neq S'$) on the value of the recourse objective function are

Table 3.3: Paired t-test and relative Euclidean distance for different number of scenarios.

S;S'	p-value (95% CI)	$\left\|\hat{\mathbf{x}}^{\mathbf{S}} - \hat{\mathbf{x}}^{\mathbf{S}'}\right\| / \left\|\hat{\mathbf{x}}^{\mathbf{S}}\right\|$
10;50	0.1	0.0026
50;100	0.01	0.0025
100;200	0.004	0.0002

not statistically different. The second approach finds the relative Euclidean distance between the two solutions. A small value implies that the solutions are statistically indistinguishable. Table 3.3 presents the results.

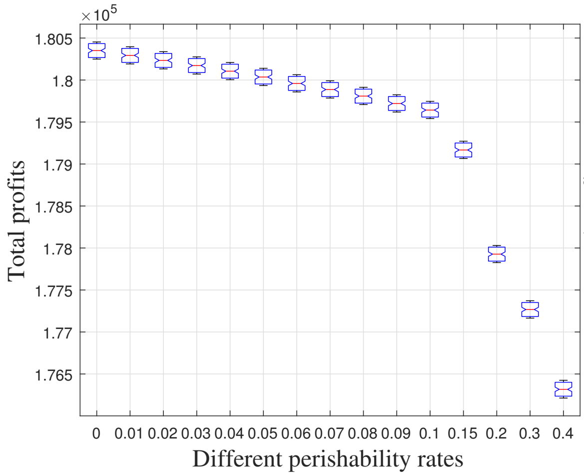
We observe that increasing the number of scenarios from 100 to 200 results in a $p-$value smaller than 0.05 and in the smallest relative Euclidean distances. Thus, it is not necessary to increase the number of scenarios used from 100 to 200. This is the reason why in our experiments we use 100 scenarios.

3.6.6 Does a Price Markdown Impact Profits in the Supply Chain?

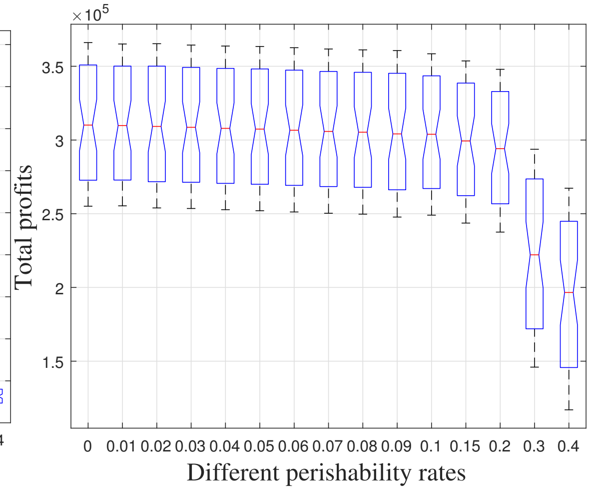
To answer this question we conducted the following two experiments. In the first experiment we consider no price markdown in the supply chain. This implies that all the products are priced as new. Next, we increase the perishability rate and observe the performance of the supply chain. Retailer's profit for each perishability rate are presented in Figure 3.3a. In the second experiments we consider a price markdown in period 3. The remaining problem parameters did not change. We conducted a similar sensitivity of perishability rate and observe the performance of the supply chain. The results are summarized in Figure 3.3b. Figure 3.3c presents the increase in profits in the supply chain due to the price markdown. These results indicate that a price markdown positively impacts profits in the supply chain.

3.6.6.1 Analyzing the Impact of Deterioration Rate:

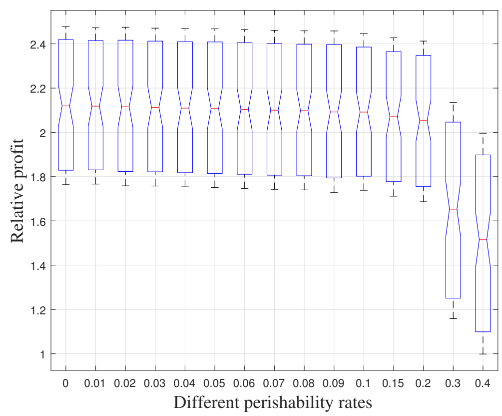
Figure 3.4a indicates that total profits in the supply chain decrease with the perishability rate. This is due to the deterioration of the inventory which impacts waste and the frequency and size of replenishment (see Figures 3.4b and 3.4c). Based on Figure 3.4d, the price of both, old and new products increases to compensate for the inventory losses. The highest profits in the supply chain are achieved when the product does not deteriorate, $\rho_{t,t+1} = 0$.



(a) No markdown in the supply chain



(b) Price markdown in period 3



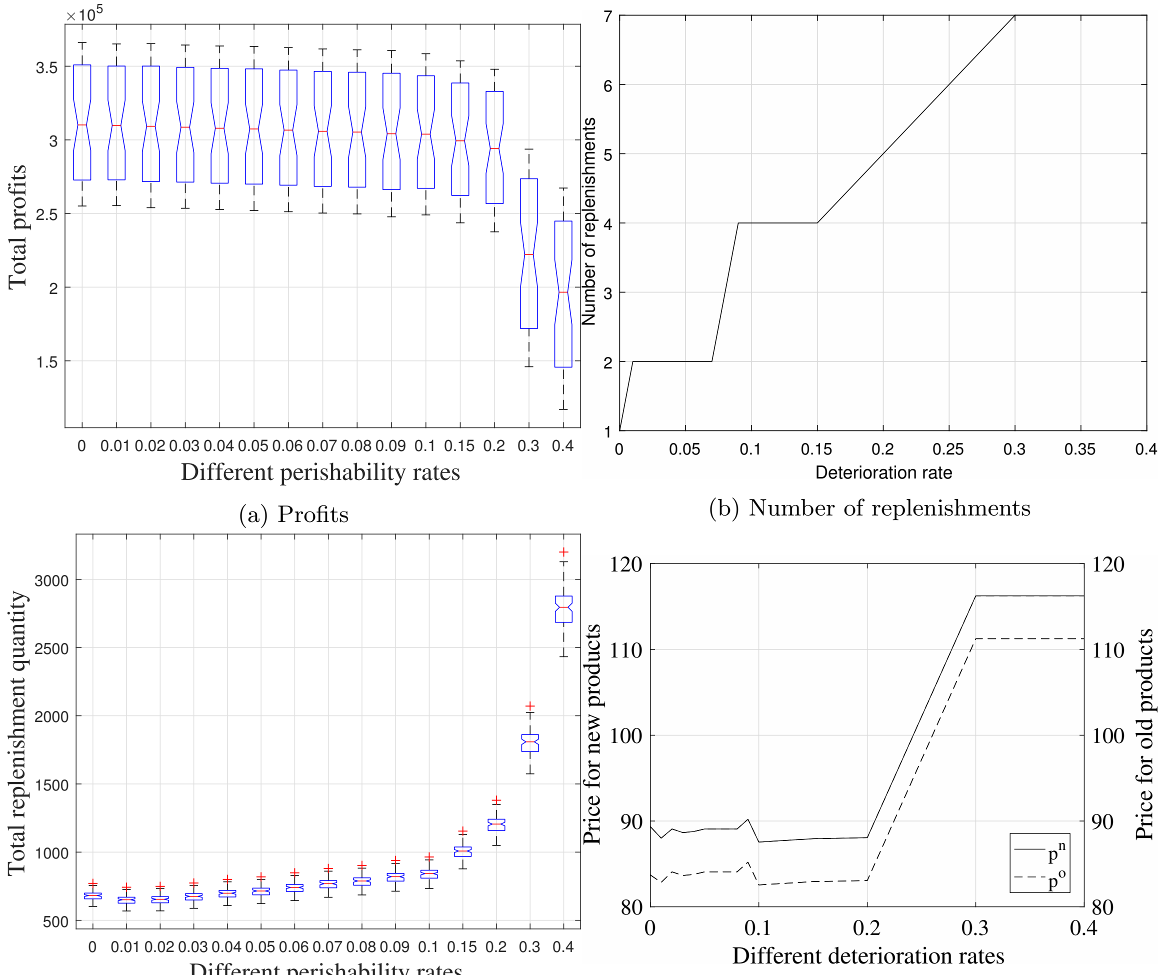
(c) Relative profits

Figure 3.3: Evaluating the impact of markdown on retailer's profits.

3.6.6.2 Analyzing the Impact of Service Level:

Service level is defined as the probability of meeting demand via on-hand inventory during a replenishment cycle. We adjust the service level of the supply chain by changing the value of the shortage cost g_t. Figure 3.5a indicates that increasing shortage cost from \$30 to \$40 improves the service level from 90% to 100%. This improvement in service level is achieved by increasing replenishment quantity (Figure 3.5c).

Increasing replenishment quantity results in additional loss of inventory due to product perishability, which results in loss of profit (Figure 3.5b). Increasing replenishment quantity results in a decrease of unit replenishment cost since the fixed ordering cost is distributed among more



(a) Profits

(b) Number of replenishments

(c) Replenishment decisions

(d) Price

Figure 3.4: Impact of deterioration rate.

products (Figure 3.5e). At the same time, the price of new products decreases. The accumulated effect of these changes is a decrease in profits in the supply chain.

3.6.6.3 Analyzing the Impact of Markdown Period:

Figure 3.6a indicates that the timing of markdown impacts the profits in the supply chain. When price is marked down in periods 5 and 6, retailer's profits are lowest. The average profits do not change when the price is marked down in periods 1 through 4, however, the deviations from this profit are smallest when price is marked down in period 4. Figure 3.6c indicates that service level decreases with markdown period. This is mainly because it becomes difficult to meet demand for old products. This is the reason why replenishment quantity increases, and consequently wastage due to product perishability. Figure 3.6d indicates a decrease in price for new and old products due to these changes of markdown period.

3.6.6.4 Analyzing the Impact of Demand Pattern:

We evaluate the performance of the supply chain under different patterns of demand which are presented in Figure 3.7. Patterns A, B and C represent non-stationary demands. To accomplish this we change the mean of ω^n ,ω^o from one period to the next. However, the mean value of these random variables overall planning horizon is the same. In pattern D both mean and standard deviation of ω^n ,ω^o change from one period to the next.

Patterns A and B differ by the standard deviation of ω^n ,ω^o. Both standard deviations are higher for pattern B. This is the reason why the average profit for both patterns is the same, but the variations around this profit is higher in pattern B. Pattern D represents erratic demand, and this is why its profits are lowest.

3.6.6.5 Analyzing the Impact of Unit Inventory Holding Cost:

Increasing the unit inventory holding cost reduces inventory in the supply chain. This impacts the frequency of orders, thus profits in a supply chain (see Figure 3.9).

3.6.6.6 Analyzing the Impact of Price Sensitivities:

Recall that $\beta^n(\beta^o)$ represents the sensitivity of demand for new (old) products to the price of new (old) products. In our experiments we maintained $\beta^n \geq \beta^o$ to ensure that an increase in price will result in an overall decrease of demand.

Increasing β^n represents an increase in price elasticity. Typically, demand for a product is elastic when there are close substitutes of that product. Revenues from these products are smaller than revenues from inelastic products. This relationship is captured in Figure 3.10a where the profits decrease with β^n.

3.6.7 Analyzing the Impact of Dual Sourcing:

We consider two suppliers, one that is reliable and expensive; and another who is unreliable and less expensive. The unreliable supplier has random capacity, thus, the amount supplied is represented via a random variable which follows the Normal distribution.

Figure 3.11a and Figure 3.11b summarize the impact of markdown period on dual sourcing. Mean supply is equal for both suppliers. Postponing product markdown to a later time results in lower prices for new and old products (see Figure 3.6d). A decrease in price reduces profit margins which leads to favoring the unreliable supplier because of his lower unit cost.

Figures 3.12a to 3.12d summarize the impact of supply availability on dual sourcing. An increase of supply available from the reliable supplier results in a decrease of price for new and old products (see Figure 3.12b), and consequently in an increase of demand which is satisfied via shipments from the reliable supplier (see Figures 3.12c and 3.12d). Based on Figure 3.12a, the number of replenishments from the reliable supplier decreases with an increase of supply which results in reduced replenishment costs.

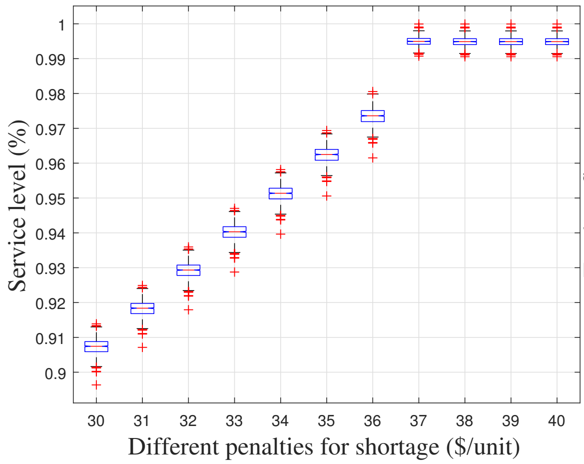
3.7 Summary of Results

This study proposes a two-stage stochastic optimization model that identifies a replenishment schedule for a periodic-review inventory system with non-stationary demand and dual sourcing. The model captures the relationship between price and stochastic demands via a linear function. The model considers a price markdown as the means to stimulate demand and minimize waste of this perishable product.

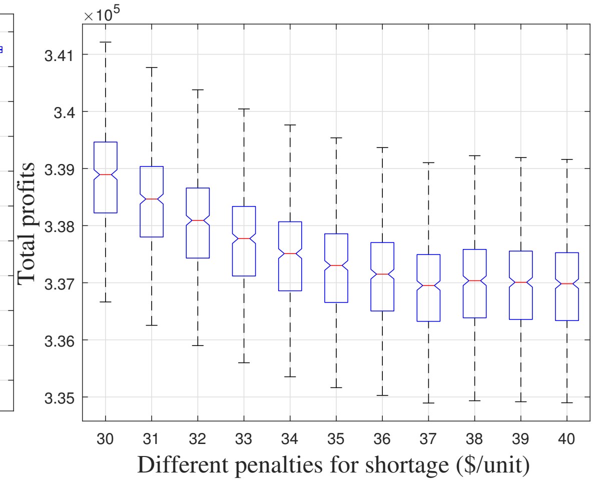
In the proposed model, the first-stage problem is bilinear. Thus, we develop a solution approach which extends the Benders' decomposition algorithm by employing a piecewise linear approximation method to solve the first-stage problem.

We develop a case study in order to validate the model and evaluate its performance. We conducted a thorough sensitivity analysis to observe the impact that timing and size of a price markdown has on inventory replenishment decisions and retailer's profits. We analyzed the impact of deterioration rate, inventory holding cost, and service level on inventory replenishment decisions and retailer's profits. Via this model we also evaluated the impact of dual sourcing in replenishment decisions. While the relationships identified via this analysis are intuitive, quantifying these relationships cannot easily be accomplished without the aid of models similar to the one proposed here.

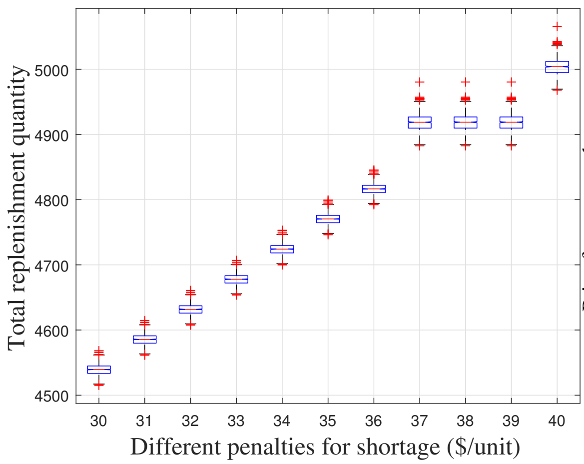
Current technological developments (such as RFID tags, point-of-sale data, loyalty programs) allow businesses to track the age of products; collect ample amount of data to estimate the impact of pricing on sales; etc. Businesses can take advantage of this data and the models proposed in here to develop decision support tools to aid pricing decisions which can potentially lead to a better align demand and supply, and consequently, higher retailer's profits.



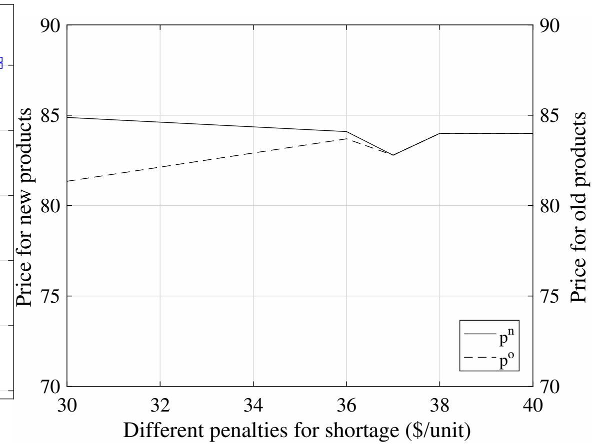
(a) Service level



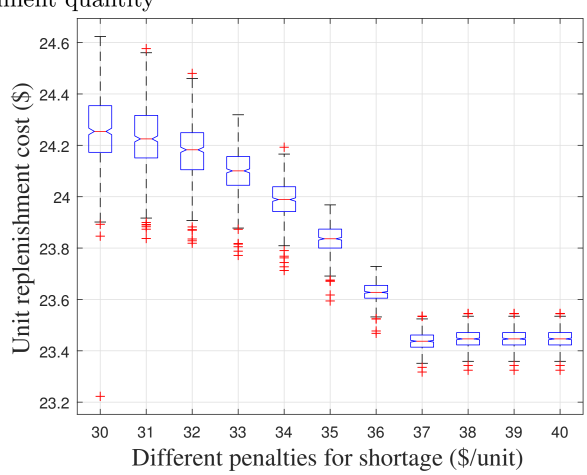
(b) Profits



(c) Replenishment quantity

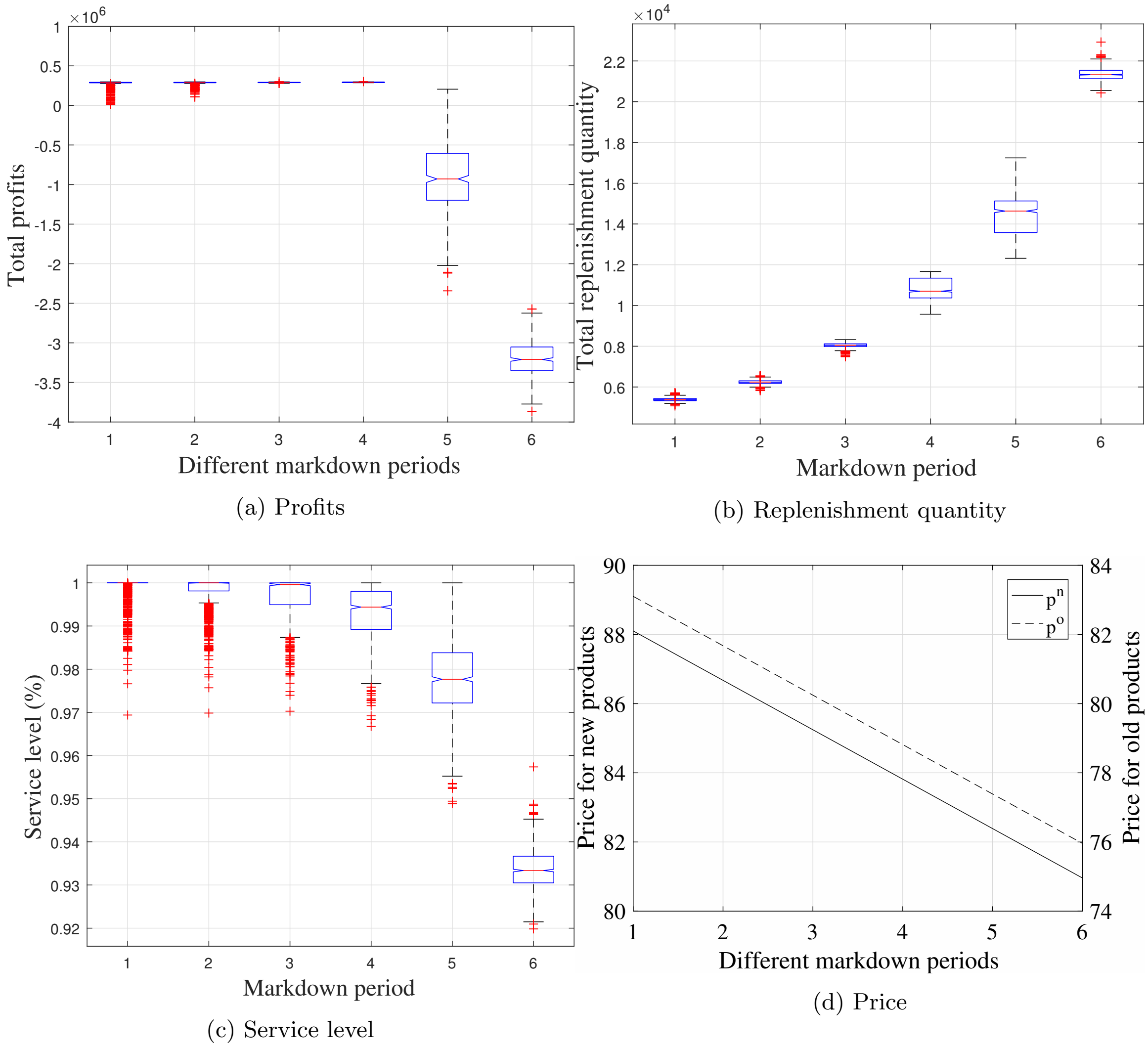


(d) Price



(e) Unit cost

Figure 3.5: Sensitivity of the optimal solution to the penalty of shortage.



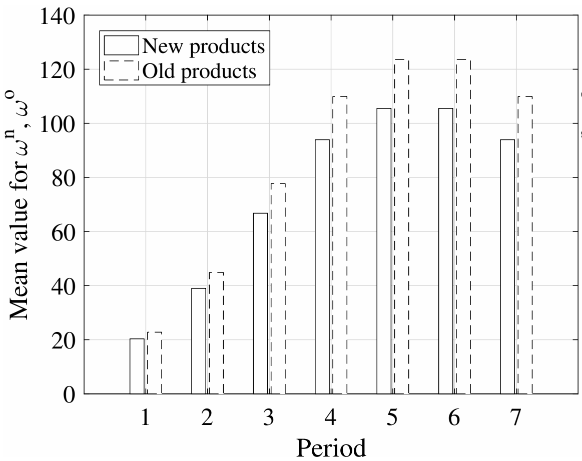
(a) Profits

(b) Replenishment quantity

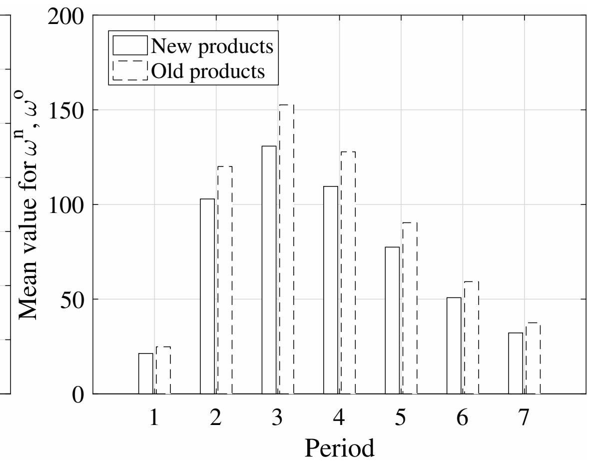
(c) Service level

(d) Price

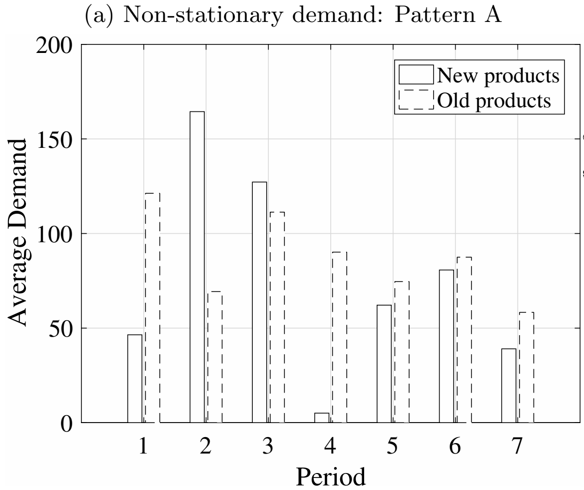
Figure 3.6: Evaluating the sensitivity of solution to markdown period.



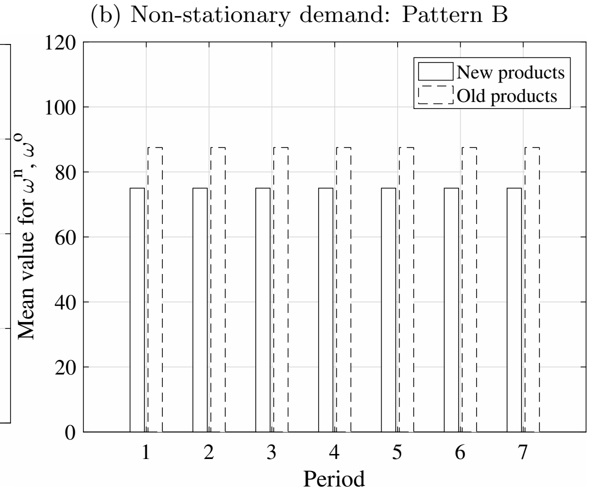
(a) Non-stationary demand: Pattern A



(b) Non-stationary demand: Pattern B

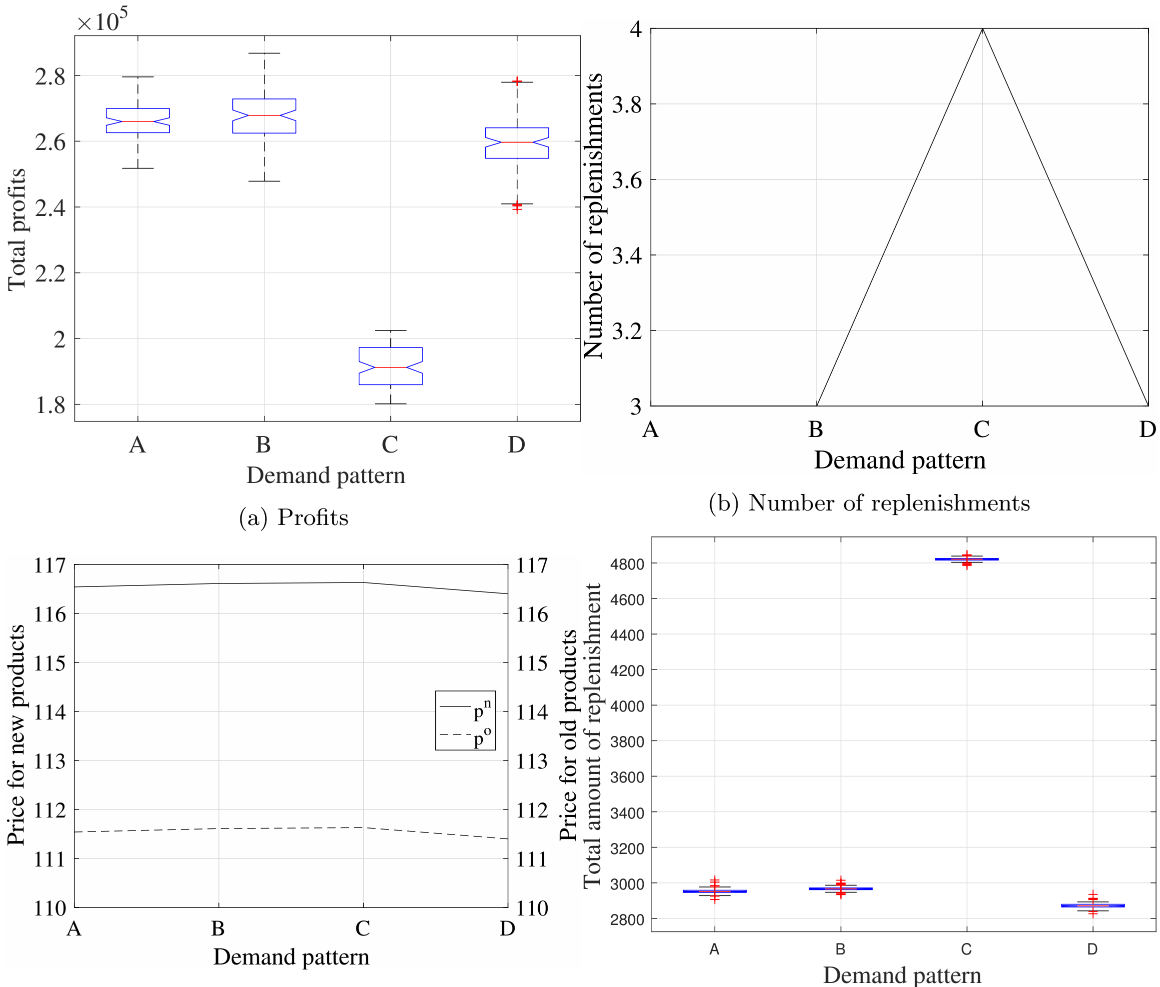


(c) Non-stationary demand: Pattern C



(d) Stationary demand: Pattern D

Figure 3.7: A summary of demand patterns tested.



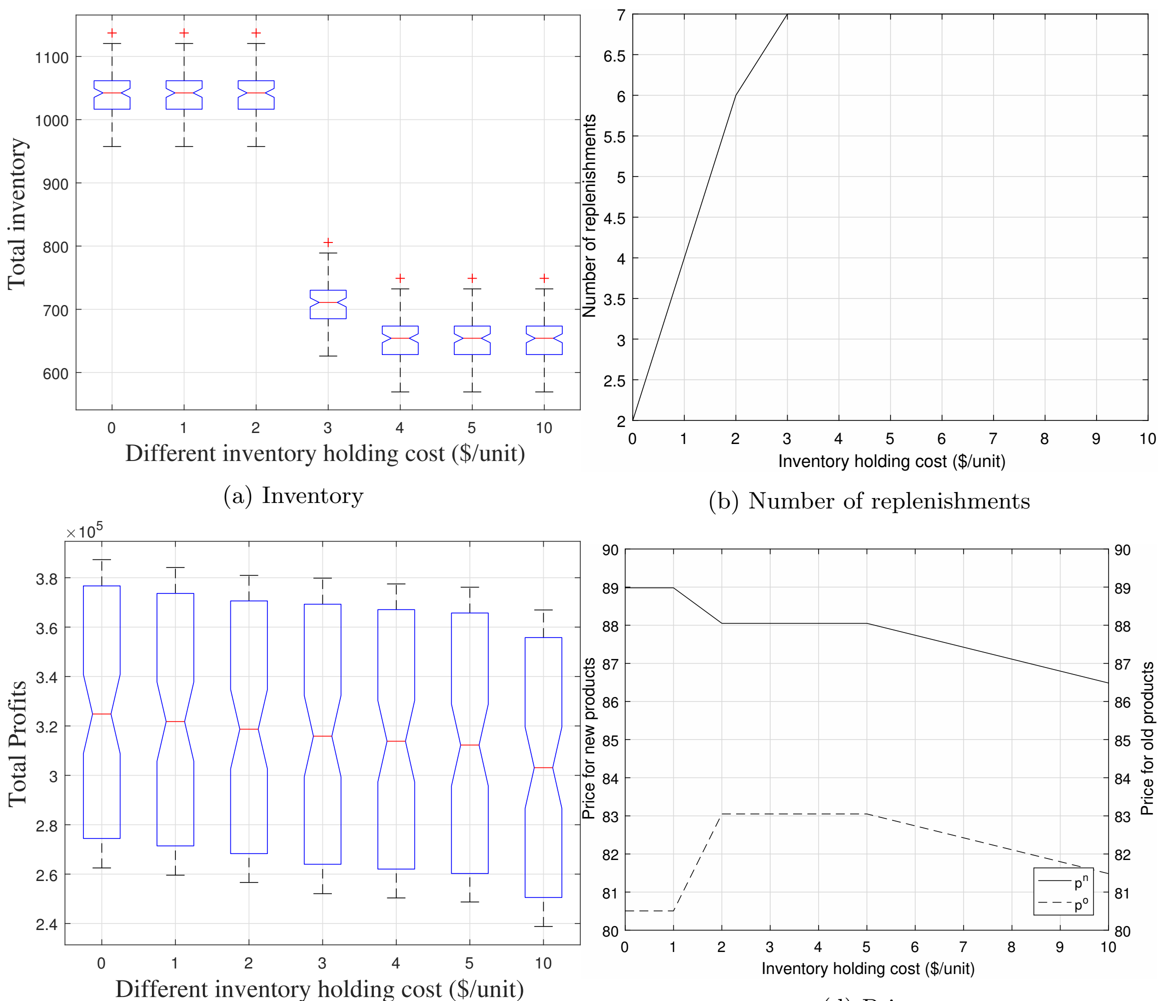
(a) Profits

(b) Number of replenishments

(c) Price

(d) Replenishment quantity

Figure 3.8: Evaluating the impact of demand patterns.



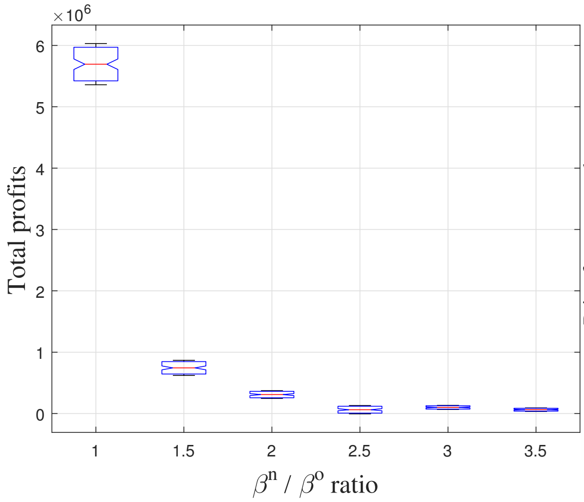
(a) Inventory

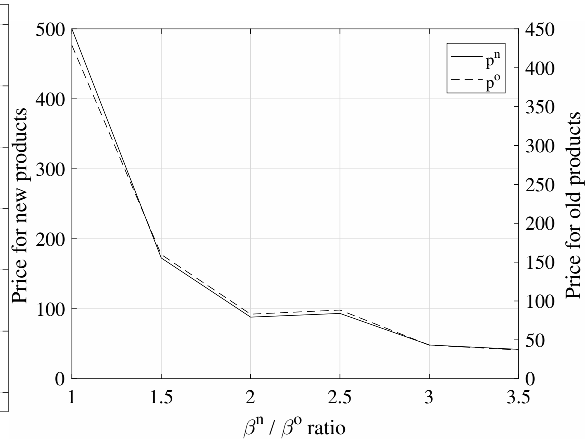
(b) Number of replenishments

(c) Profits

(d) Price

Figure 3.9: Sensitivity of the optimal solution to the unit inventory holding cost.

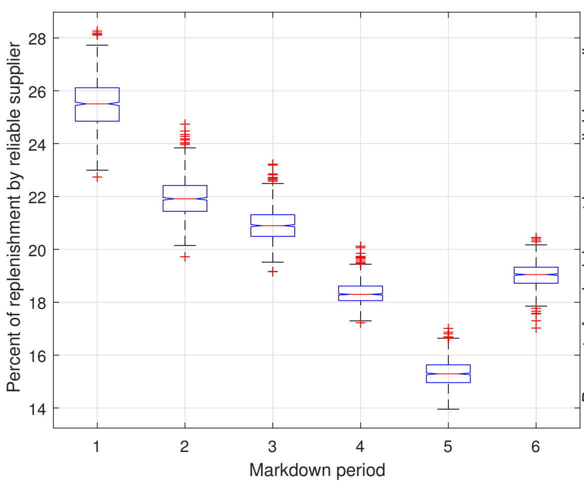


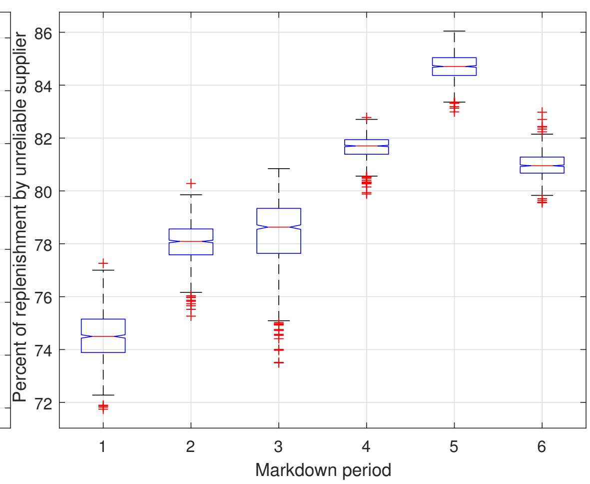


(b) Price

(a) Profits

Figure 3.10: Sensitivity of the optimal solution to the ratio β^n/β^o.

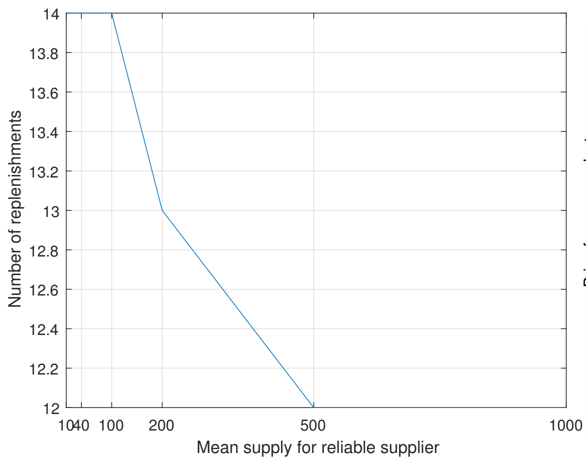




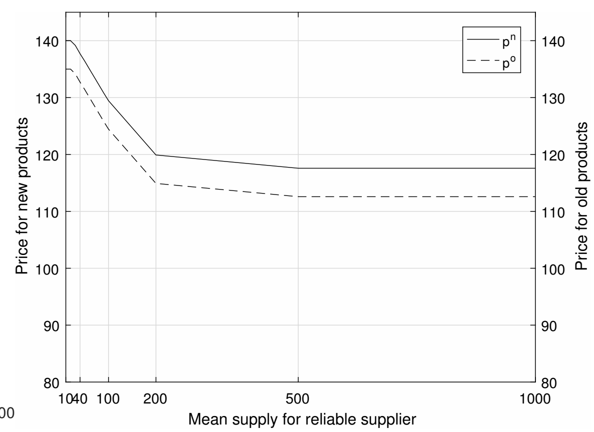
(a) Reliable supplier.

(b) Unreliable supplier.

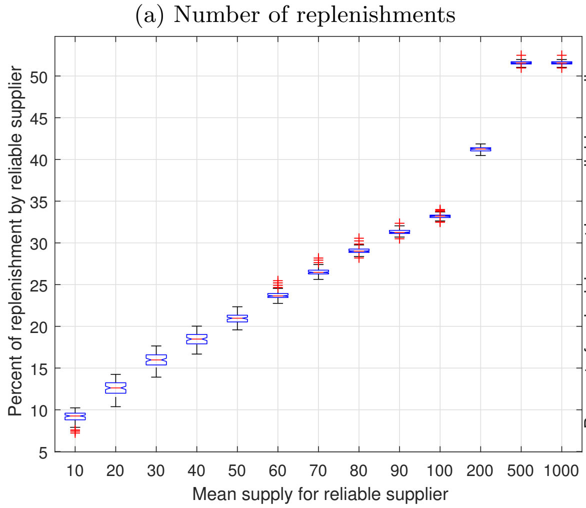
Figure 3.11: Impact of markdown period when the supplier is unreliable.



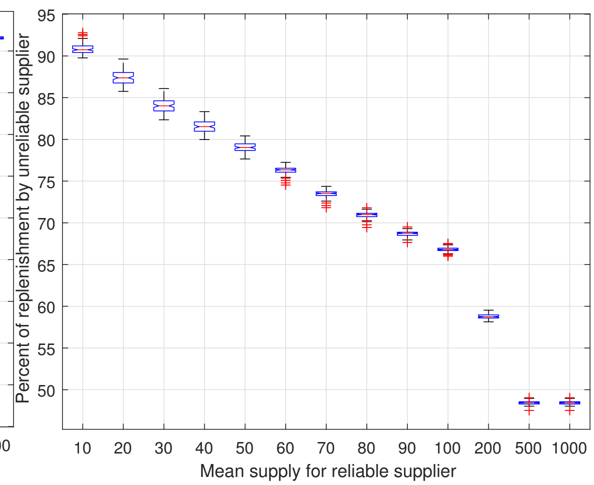
(a) Number of replenishments



(b) Price



(c) Reliable supplier.



(d) Unreliable supplier.

Chapter 4

Forecasting Childhood Routine Immunization Vaccine Demand: A Case Study in Niger

4.1 Introduction

Over the years, several deadly diseases such as Polio and Diphtheria have been prevented via vaccines. Every country has developed and implements a childhood routine vaccination program. This has resulted in a large decrease in the global mortality rate associated with unvaccinated children. However, the death rate of children under age five is still high in low-income countries. In 2016, the death rate was 73.1 deaths per 1000 live births in low-income countries which is almost 14 times greater than the average mortality rate in high-income countries [109].

In 1974, the WHO initiated the EPI with the goal of providing vaccination to every child around the world [107]. Later in 1999, the WHO with the help of UNICEF, the World Bank and other institutions extended the EPI to the GAVI with the objective of saving children's lives in lower-income countries [106]. The original standardized vaccination schedule includes 1 dose of BCG after birth, 3 doses of DTP in primary series within 12-23 months after birth, 3 doses of oral polio vaccine in primary series within 6-14 weeks after birth, and 1 dose of measles vaccine within 6-9 months after birth.

The success of EPI program depends on the implementation of policies and the structure of the distribution network design. Distribution network design is an important and difficult task to accomplish since many new vaccines, such as, Hepatitis B (HepB), Yellow Fever, and Haemophilus Influenza Meningitis (Hib) have been introduced into the routine immunization since the EPI has started [69].

The objective of a vaccine distribution network is to ensure an on-time delivery of vaccines to clinics. In many countries, the vaccine supply chain includes multiple levels. A typical distribution network consists of four levels. The first layer connects a central storage to the district level warehouses. The second layer connects the district level warehouses to the regional stores. The third layer connects the district level warehouses to the clinics where vaccines are administered [14].

For many products, the inventory replenishment and transportation schedule decisions rely on data forecasts. Demand is typically the most important data that needs to be forecast. The demand pattern for different types of products are different. For example, the demand for luxurious products increases with an increase in family income level. However, the demand is not a function of salary for the basic products such as bread, milk, and egg. Ozawa et al. indicate that the expected demand for vaccines in northern Nigeria is influenced by factors such as vaccination cost, lack of vaccine knowledge, media coverage, social norms and religious views, and care seeking [110]. Analyzing the impact of these factors on the expected demand for vaccines is challenging since these are subjective measures.

Identifying factors which impact demand, and can be measured, is important in developing functional relationships to predict expected demand for vaccines. Accurate estimates of the demand pattern for vaccines is critical for the design and implementation of inventory replenishment and transportation plans. Inventory replenishment and transportation decisions are important because they impact vaccine wastage and shortage. Vaccine wastage is a result of different factors, such as, expiration, vial breakage, continuous heat exposure, etc [108]. Well-designed transportation policies result in lower amount of wastage. For example, decreasing vaccine handling reduces the chance of vial breakages. Vaccines can be quite expensive and expired vaccines cannot be used [89]. Therefore, reducing wastage is crucial for cost containment. Moreover, high quality demand forecasts have a great impact on the development of high quality replenishment schedules since these forecasts prevent shortages.

Various studies have assessed the factors which impact expected demand for vaccines [110, 121, 99, 54]. However, they do not prescribe any prediction model for vaccination demand. The analysis of demand is vital when designing a vaccine distribution network. To do that, the authors in [14] estimated the demand at Integrated Health Centers (IHC) using the birth registry at the district level. In another study of Benin's vaccine supply chain, the authors modeled the expected demand for vaccines using Poisson distribution [24]. However, these papers ignore the monthly demand variations as well as the difference in monthly demand across various regions in the country. Therefore, analysis of expected demand for vaccine represents a research need. In this regard, we contribute to the literature by developing a regression-based estimate of the future demand for CIV. This model is characterized by different vaccine types as well as region.

The models developed are used to analyzes the demand for CIV in Niger. We used data from DHS to develop these models. The current vaccine supply chain in Niger comprises of four tiers which consists of 1 central store, 8 regional stores, 42 district stores, and over 600 IHC. UNICEF replenishes the inventory of the central store every two month via air transportation. Vaccines are shipped from the central store to the regional stores every three month via cold trucks. The district stores pick up the vaccines from the regional stores on a monthly basis via 4×4 trucks. The IHC picks up the vaccines from the district stores monthly using motorcycles, bikes, or private cars. We study the monthly demand pattern for CIV in 8 different region of Niger.

4.2 Methods

4.2.1 Data input

This study used data from DHS and also the census data for Niger. The DHS population-based and nationally representative surveys from large samples of sizes between 5,000 and 30,000 households.[5]. A household survey provides information related to women and children in a household, such as, weight, height, vaccination dates within the last 5 years, etc. The vaccination data is provided by the survey respondents for the EPI vaccines for children younger than 5 years old. The household survey for Niger was collected in 2012 by interviewing each household. The census data provides the demographical information including population density, poverty level, and education level. The census data in Niger is updated every 5 years, and it is only available at the national level [46]. The data about population density, as well as, poverty and education level at the regional level is only provided for certain years [16, 42]. We used the world bank data portal to extrapolate population size for the remaining years. Population size is estimated based on the annual birth and death rate [8]. The data about the number of integrated health centers in each region was collected from [33].

Using the DHS data, we enumerated the number of children per household. We used each child vaccination date to estimate the number of vaccinations in a month for different types of vaccines and for each region in Niger. The vaccination dates range from 2007 to 2011. The number of vaccinations was divided by the total number of children (provided by the survey) to estimate the vaccination rate. It should be noted that, we excluded any household for which the date of vaccination was unknown.

In order to estimate the total number of vaccines, we used the most recent Niger census records. The regional level records are only available for 2012. We used the 2012 records to estimate the distribution of population by region [42]. This distribution for 2012 was: 2.85% of total population lived in Agadez, 3.46% of total population lived in Diffa, 11.89% of total population lived in Dosso, 19.85% of total population lived in Maradi, 19.42% of total population lived in Tahoua, 15.89% of total population lived in Tillaberi, 20.65% of total population lived in Zinder, and 5.99% of total population lived in Niamey. The total population of Niger during 2002 to 2011 is estimated as follows: 10,639,740; 11,058,590; 11,360,540; 11,665,940; 12,525,090; 12,894,870; 13,272,680; 15,306,250; 15,878,270; 16,468,890; 16,344,690. Using the population distribution by region, we estimated the regional population during 2002 to 2011.

The birth rate of Niger during 2002 to 2011 is the following: 49.95, 49.54, 48.91, 48.3, 50.73, 50.16, 49.62, 51.6, 51.08, 50.54, 50.06 [8]. These rates indicate the number of births per 1000 people. To estimate the number of children younger than 5 years old in a particular year, we summed the number of birth in the corresponding year and over the last 5 years. The total number of vaccines per region was calculated by multiplying the regional population of children under 5 years old by the vaccination rate. In all regions there are some data point for which the number of vaccines administered were not available.

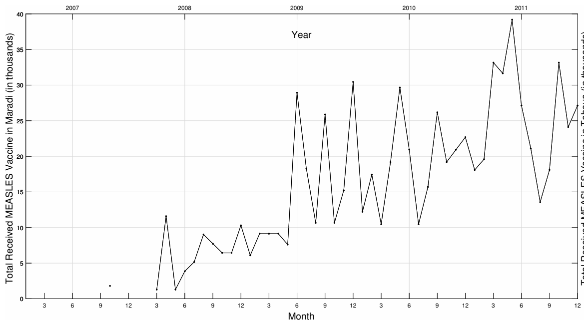
4.2.2 Data analysis

The data analysis presented in this study is conducted using the statistical software SAS. In all the analysis, the significance level is set at $p = 0.05$. We decided to discard missing data since they induce bias, and this bias negatively impacts the data analysis. Therefore, we used imputation to deal with missing data points. Imputation substitutes a missing value with a predicted value. There are different methods to predict the missing value, such as, hot-deck, cold-deck, mean substitution, and regression. Repeating imputation is called multiple imputation, which results in reducing the standard error of prediction outcome. In multiple imputation, the missing values are filled by the mean of results [148]. In this study we used multiple imputation of regression to predict the missing data points.

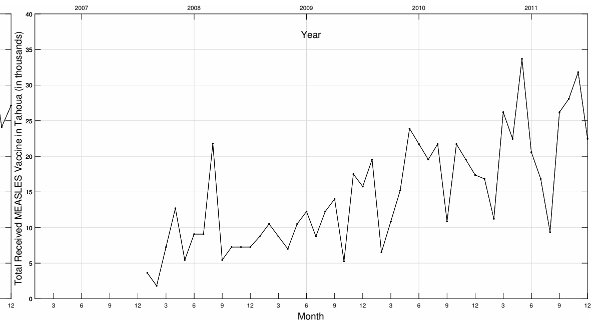
In order to detect any non-randomness in the number of vaccines administered during 2007 to 2011, autocorrelation tests were conducted. A single autocorrelation test was conducted for each vaccine type and each region using the Durbin-Watson test in the “Autoreg” procedure in SAS.

In order to address the effect of population size on the expected demand for vaccines, a linear regression model is developed. In this model, the independent variable is the population size and the dependent variable is expected demand for vaccines. Each vaccine type in each region was examined via a single regression model. The intercept of these regression models was forced to zero since population of size zero would require zero vaccines. For this analysis, we used the data about monthly number of vaccines administered as the dependent variable. Since the population size per region in each month of a year is not available, the yearly population size was divided using a mid-year split that used the difference in annual population estimates and split the difference at the mid year.

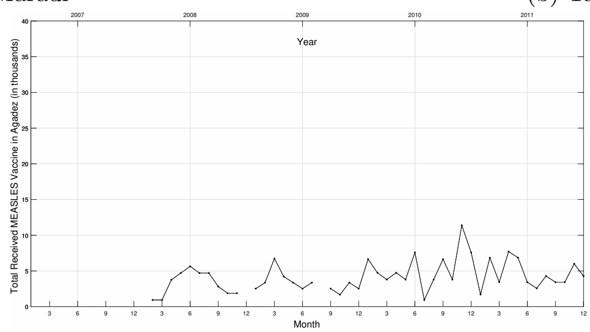
There are other variables that have previously been identified as factors influencing expected demand for vaccines [121]. These factors include maternal education level, wealth index, and clinic availability. A linear regression model at the national level was conducted to analyze the impact of these factors as well. In this model, the dependent variable is expected demand for vaccines, and the independent variables are percentage of population below the national poverty line, adult literacy rate, and the number of clinics. The data about the percentage of population below the national poverty line and adult literacy rate is available only for 2008. The adult literacy rate is defined as “the percentage of population aged 15 years and over who cannot both read and write with understanding a short simple statement on their everyday life” [6].



(a) Maradi



(b) Tahoua



(c) Agadez

Figure 4.1: Total administered Measles vaccine through 2007-2011 in three different regions of Niger

4.3 Results

The total number of Measles vaccine administered in Maradi, Tahoua, and Agadez from 2007 to 2011 increases with the time as it is shown in figures (4.1a) and (4.1b). However, the results of autocorrelation analysis indicate that the Durbin-Watson test with the hypothesis of no autocorrelation is not significant ($P > 0.05$). That is, the expected demand for vaccines is not autocorrelated when the lag is 1. These results are consistent for different vaccine types and different regions. It is observed that the demand pattern differs for the three regions even though Maradi and Tahoua have pretty close population sizes. For example, in the mid year of 2009, there is a big upward shift in the number of vaccines administered in Maradi, but this is not observed in Tahoua and Agadez. This confirms the differences between the magnitude of parameter estimates for individual simple linear regression models predicting the expected demand for vaccines based on the regional population size per month in Table (4.1).

The impact of population size on the expected demand for vaccines in all regions is significant ($p < 0.001$). The parameter estimates reveal the contrast between different vaccine types and different regions (Table (4.1)). The parameter estimates are different for Agadez and Diffa although these

Table 4.1: Parameter estimate (standard error, and R-square) for all individual simple linear regression models predicting the vaccine demand based on the monthly population size in different regions.

Regions	Agadez		Diffa		Dosso		Maradi	
Vaccines	Parameter Estimate (%) (Standard Error)	R-squared	Parameter Estimate (%) (Standard Error)	R-squared	Parameter Estimate (%) (Standard Error)	R-squared	Parameter Estimate (%) (Standard Error)	R-squared
BCG	1.176 (0.0009)	0.75	0.463 (0.0004)	0.71	0.838 (0.0006)	0.77	0.450 (0.0708)	0.43
DPT1	1.151 (0.0008)	0.79	0.447 (0.0004)	0.72	0.952 (0.0006)	0.80	0.513 (0.0788)	0.44
POLIO1	1.12 (0.0008)	0.79	0.447 (0.0004)	0.73	0.918 (0.0006)	0.79	0.500 (0.0774)	0.44
DPT2	1.138 (0.0007)	0.84	0.424 (0.0003)	0.73	0.888 (0.0005)	0.83	0.468 (0.0707)	0.45
POLIO2	1.112 (0.0007)	0.83	0.423 (0.0003)	0.73	0.879 (0.0005)	0.83	0.461 (0.0691)	0.46
DPT3	1.094 (0.0006)	0.85	0.413 (0.0003)	0.73	0.796 (0.0006)	0.78	0.4058 (0.0612)	0.45
POLIO3	1.062 (0.0006)	0.86	0.413 (0.0003)	0.73	0.788 (0.0005)	0.79	0.390 (0.0595)	0.45
MEASLES	0.956 (0.0006)	0.80	0.397 (0.0003)	0.72	0.672 (0.0006)	0.71	0.326 (0.0508)	0.44
Regions	Tahoua		Tillaberi		Zinder		Niamey	
Vaccines	Parameter Estimate (%) (Standard Error)	R-squared	Parameter Estimate (%) (Standard Error)	R-squared	Parameter Estimate (%) (Standard Error)	R-squared	Parameter Estimate (%) (Standard Error)	R-squared
BCG	0.636 (0.0005)	0.78	0.971 (0.0006)	0.81	0.565 (0.0004)	0.80	1.378 (0.0007)	0.89
DPT1	0.701 (0.0005)	0.78	1.037 (0.0006)	0.87	0.632 (0.0005)	0.74	1.306 (0.0007)	0.87
POLIO1	0.686 (0.0004)	0.81	0.998 (0.0005)	0.87	0.605 (0.0005)	0.75	1.298 (0.0006)	0.88
DPT2	0.642 (0.0004)	0.79	0.974 (0.0006)	0.84	0.564 (0.0004)	0.74	1.219 (0.0006)	0.87
POLIO2	0.635 (0.0004)	0.81	0.938 (0.0005)	0.85	0.557 (0.0004)	0.74	1.209 (0.0006)	0.88
DPT3	0.577 (0.0004)	0.76	0.906 (0.0005)	0.84	0.556 (0.0005)	0.73	1.16 (0.0006)	0.86
POLIO3	0.578 (0.0004)	0.77	0.886 (0.0005)	0.85	0.546 (0.0004)	0.74	1.145 (0.0006)	0.87
MEASLES	0.469 (0.0003)	0.81	0.796 (0.0005)	0.80	0.383 (0.0004)	0.65	0.957 (0.0005)	0.85

regions have similar population sizes. For example, if the population size of both regions is equal to 1000, the region of Agadez would require around 12 BCG vaccines while the region of Diffa would require only 5. In contrast, the parameter estimates for the population size within the region of Diffa does not substantially change for different vaccine types. For example, the parameter estimate of populations size for DPT3 and POLIO3 vaccines equals 0.413%. Additionally, in all regions, the impact of population size decreases for the final dose in the childhood immunization schedule. For example, the parameter estimates of the population size for the first dose of DPT (DPT1) is lower than the second dose (DPT2) and the third dose (DPT3). The values of $R-squared$ indicate that the models developed (except the ones for the region of Maradi) provide a pretty good fit to the data.

In order to understand the factors that impact expected demand for vaccines at the national level, we developed regression models with parameter estimates and standard errors presented in Table (4.2). These regression models were built using the national level data available in 2008. It is observed that the impact of population size, as well as, the number of clinics are not significant for all vaccine types ($p > 0.05$). Note that $*$ indicates that the impact of a parameter is not statistically significant ($p > 0.05$). The results show that, the adult literacy rate does not influence the demand for DPT3, and POLIO3. The impact the adult literacy rate on BCG, DPT1, POLIO1, DPT2, and POLIO2 vaccines is similar. In contrast, the impact of the percentage of population under poverty line is different for various vaccine types. Also, MEASLES vaccine is not affected by any parameter in the model. The R-squared values indicate that, the models provide a good fit to the data for all vaccine types.

Table 4.2: Parameter estimate (standard error, and R-squared) for the national level linear regression models predicting the vaccine demand based on the regional population size, percentage of people under poverty line, adult literacy rate in 2008.

Vaccines	Intercept		Percentage of people under poverty line		Adult literacy rate		R-square
	Parameter Estimate (thousands)	Standard Error	Parameter Estimate (thousands)	Standard Error	Parameter Estimate (thousands)	Standard Error	
BCG	-106.505	40.333	2.695	0.625	2.241	0.611	0.91
DPT1	-110.182	39.294	2.821	0.609	2.192	0.596	0.93
POLIO1	-111.421	44.099	2.754	0.684	2.202	0.668	0.91
DPT2	-103.684	41.249	2.549	0.641	2.103	0.625	0.91
POLIO2	-104.484	46.411	2.586	0.720	2.107	0.703	0.90
DPT3	*	*	1.813	0.774	*	*	0.84
POLIO3	*	*	1.831	0.819	*	*	0.82
MEASLES	*	*	*	*	*	*	0.79

Note: * indicates that the parameter is not significant in the model.

4.4 Discussion

The main objective of this study is to develop regression models to predict the demand for CIV in different regions based on population size. The results indicate that, population size significantly impacts the expected demand for vaccines at the regional level. Our analysis suggests that, this impact differs by region and vaccine type. It was expected that the regression models developed for Maradi and Tahoua would be similar since their population size is very close. However, the results do not support our expectations. This is mainly because, on addition to population size, there are other factors which impact expected demand for vaccines.

The difference between parameter estimates for Agadez and Diffa are mainly due to differences in population size and education level. While the adult literacy rate in Agadez was 47.7% in 2008, this level was 36.9% in Diffa. Additionally, Diffa has the lowest percentage of population below the national poverty line. These estimates indicate that, families in Diffa are more likely to fully immunize their children as compared to the other regions in Niger. Within Diffa, the parameter estimates are similar for different vaccine types. However, we observe a decreasing trend in the parameter estimate for the series of a vaccine type (such as, DPT1, DPT2 and DPT3; or POLIO1, POLIO2 and POLIO3). This trend may be the result of some new societal issues (as discussed in [110]) and should be further investigated.

It is well recognized that, expected demand for vaccines in developing countries is not only influenced by population size. To account for this, we used the national level data of 2008 to evaluate the impact of a number of factors, including the total population size, the percentage of population below the national poverty line, the adult literacy rate, and the number of clinics. The study highlights the difference between the regions. The results suggest that, at the national level, the population size and the number of clinics are not significant. This highlights the importance of developing a single, region-based regression model as presented in Table (4.1). The percentage of people under poverty line and the adult literacy rate are significant factors that impact the demand for several vaccines.

In 2008, the immunization coverage for MEASLES in Niger is highest among CIV [7]. Therefore, MEASLES might not be affected by the percentage of people under poverty line and adult literacy rate. The results suggest that, regardless of vaccine type, the adult literacy rate has the same impact on the expected demand for vaccines. However, this impact is not significant for the final series of DPT and POLIO. This outcome may be due to other societal issues which are not discussed in this study.

The results of our data analysis provide insights about the influence of population size, poverty and education level in the expected demand for vaccines. That means, the availability of region-based data about population size is important to estimate the expected demand for vaccines, which in turn, is important information to design a vaccine distribution network and shipping polices. There are a number of limitations associated with this analysis, such as, (i) we used the DHS data which is a representative sample and may not have accounted for some of the variability in vaccine usage; (ii) we estimated population size using a mid-year split; (iii) the low R-squared value in regression models for Maradi suggests that there might be other factors which impact expected demand for vaccines at the regional level. Future analysis of more comprehensive data may help us understand the regional differences in expected demand for vaccines.

Chapter 5

Stochastic Optimization Models For Childhood Vaccine Distribution Network Design: A Case Study In Niger

Vaccines have been used for more than 50 years to prevent childhood diseases. In 1974, The World Health Organization (WHO) launched the Expanded Program on Immunization (EPI) with the goal of providing vaccines to every child and pregnant woman around the world [142]. Nevertheless, today, many children do not receive the necessary vaccinations required by the routine immunization schedule as defined by EPI. In 2015, 56.4% of deaths in Africa were due to infectious diseases [144].

EPI members are non-profit organizations such as the World Bank, WHO, UNICEF, and public health department of several countries and vaccine manufacturers. Health departments in developing countries are facing a number of challenges managing their vaccine supply chain. These challenges include introduction of a new vaccine to the routine immunization schedule, changing the vial size of a vaccine, and global vaccine shortage. Introducing a new vaccine to an existing immunization schedule requires additional capacity in the cold chain. However, acquiring the necessary investments to increase capacity in developing countries is challenging. When a single-dose vaccine is replaced with a multi-dose vaccine, the OVW increases. Furthermore, when a global vaccine shortage happens, an effective use of resources is required. These challenges affect their effectiveness in managing EPI. This negatively impacts immunization coverage levels in these countries.

The goal of EPI is to achieve the predefined targeted immunization coverage rate. For example, the Global Vaccine Action Plan signed by the World Health Assembly in 2012, established a target immunization coverage of 90% for EPI vaccines by 2020 [143]. To achieve this objective, efficient vaccine supply chains should be established to guarantee a timely delivery of vaccines to clinics and eliminate wastage. Efficient supply chains eliminate missed opportunities due to vaccine unavailability.

The typical EPI vaccine supply chain consists of procurement, storage, and distribution activities. This supply chain usually comprises of four tiers, which are, a central store, regional stores, district stores and clinics. UNICEF replenishes the inventory of the central store annually. Vaccines are shipped from the central store to the regional stores using refrigerated trucks. Vaccines are shipped to district stores using 4×4 trucks. The district stores ship vaccines to clinics using cars, and motorcycle. Vaccination is done only at these clinics. The number of tiers in the supply chain might be more or less than four depending on the country. For example, Vietnam has one additional tier between the regional and district stores. However, the typical supply chain structure is composed of one central or national level store where vaccines are inventoried and delivered to downwards members of the supply chain. Including more tiers, while decreases transportation costs, increases the costs of collecting and storing vaccines. Removing a tier in the structure of this supply impacts on the system-wide costs and reduces vial breakage due to fewer times vaccines are touched.

Designing an effective and robust vaccine supply chain requires accurate prediction of patient arrivals and of the corresponding variations. This is because underestimating vaccination needs results in shortages and overestimating vaccination needs results in unused inventory and wastage. Therefore, accounting for stochastic patient arrivals is necessary to develop an accurate inventory replenishment schedules in the supply chain.

This study develops a stochastic optimization model to design a vaccine supply chain which maximizes the number of fully immunized children. The model determines vial distribution strategy and storage capacity in each location. The goal is to design an efficient cold chain to deliver vaccines from manufactures to clinics. This is a data driven model which is developed using real life data from Niger. We use our data-driven model to evaluate different supply chain designs and vaccine administration policies.

This study is organized as follows. In §(5.1), a review of the relevant literature on vaccine supply chain design is discussed, and a research gap is identified. The proposed model is presented in §(5.2). In §(5.4), data and computational results are presented.

5.1 Literature Review

Prior research about vaccine supply chain design proposed simulation models which were build to evaluate the supply chain performance of a specific country. For example, [13] developed a custom designed discrete event simulation model, namely Highly Extensible Resource for Modelling Supply Chains (HERMES), to simulate the the vaccine distribution network for EPI vaccines in Niger. The authors used HERMES to analyze the impact of changing the Measles vaccine vial size on the performance of Niger's vaccine supply chain. Similarly, [58] utilized HERMES to compare the impact of adding stationary storage capacity with increasing transport capacity on the vaccine availability at clinics in Niger. Research conducted by [14], [79], and [24] used HERMES to explore the impact of different vaccine distribution strategies on the supply chain costs and vaccine availability in clinics. A An important observation made by these researchers is that changing a four-tier to a three-tier supply chain increases the vaccine availability in clinics.

Optimization tools for vaccine supply chain design have rarely been used in the literature. Most of the optimization models proposed have focused on a single developing country and for one specific disease. For example, [39] proposed an integrated supply chain and health economic model to improve the supply of Influenza vaccine. The authors point to the need for providing incentives to help governments improve the distribution network. In the context of developing countries, [34] proposed a deterministic model model to analyze the impact of several distribution strategies and vaccination policies on Niger's vaccine supply chain. One of the findings is that the performance of a three-tier supply chain is almost identical to the performance of a four-tier supply chain. This observation is different from the finding of HERMES.

Models that use the expected value to approximate the demand for vaccines do overestimate/underestimate the actual demand in a time period. These estimations result in vaccine wastage or vaccine shortage. Additionally, simulation-based approaches provide approximate policies which are often different from exact strategies. Therefore, there is a need for stochastic optimization models to help decision makers identify efficient supply chain strategies. In this regard, we contribute to the literature by proposing a chance constrained programming (CCP) model for vaccine supply chain design. To the best of our knowledge, stochastic optimization models in vaccine distribution network design has not been addressed in the literature.

CCP models have been developed to optimize problems with several uncertainties. The CPP modeling approach ensures that the probability of meeting a specific constraint is above a predefined level. This approach is proven to be robust, but, difficult to solve [132]. CPP have been used in the literature to model problems in various fields such as renewable energy generation, unmanned autonomous vehicle navigation, and financial risk management [53].

When the random variable can be decoupled, the constraint is relaxed to deterministic constraints via probability density functions. One strategy used in the literature is to obtain the probability from the true distribution and substitute the chance constraint with a deterministic expression [126]. However, when calculating probabilities is not straightforward, converting the chance constraints to the corresponding deterministic constraints is challenging [61]. Therefore, approximate solution approaches like SAA have been developed to address these computational challenges. Via SAA the stochastic optimization model can be formulated as a deterministic model [111]. This DEF is obtained via replacing the original distribution of the random variable with an empirical distribution. The empirical distribution is found by generating realizations for the random variables in the chance constraints via random samples. The obtained DEF guarantees that the number of failures in the independent trials of random sampling is below a certain level. [111] have studied the theoretical properties of SAA. The authors presented methods to find statistical lower and upper bounds to the optimal value of CCP problem.

5.2 Problem Statement

We consider a vaccine supply chain via which EPI vaccines are distributed to clinics for administration. The planning horizon considered is one year, and the planning period is one month. We use a network design model to represent this supply chain. We define the vaccine supply chain at time period $t \in \mathcal{T}$ on a directed graph $G_t = (J, A)$, where J is the set of nodes and A is the set of arcs. This model assumes that vaccine vial size (dose/vial) is predefined and there is no manufacturing capacity limitation. Vaccines have different characteristics, and therefore, some vaccines can be stored in both, refrigerators and freezers, others must only be stored in refrigerators. Some vaccines must be restored using diluent before injection. The diluents are also refrigerated.

Some EPI vaccines require one dose for a complete immunization, others require multiple doses at different stages of childhood. For example, to be fully immunized against Measles two doses are used. The first dose is administered when a child is 12 to 14 months old, and the second dose is administered at least 4 months later. The number of doses needed to fully immunize children is termed "vaccine regimen".

It is assumed that a fraction of vaccines is lost during transportation due to breakage. Also, OVW in clinics depends on the mean patient arrivals and vial size. The OVW is zero for single-dose vials and is typically positive for multi-dose vials. In the model presented here, the OVW rates are calculated offline. Finally, vaccines are only administered in clinics. It is assumed that demand for vaccines is zero for the rest of facilities in the supply chain.

In Table 1 we summarize the decision variables and parameters used in our mathematical model. Note that we use similar notation as in [34] our model since we are extending this model to capture additional characteristics of the supply chain. Particularly, in our model, demand for vaccine $i \in I$ at clinic $j \in J$ in time period $t \in T$ is represented via $\mu_{ijt} + \tilde{\omega_{ijt}}$. μ_{ijt} is the meand and $\tilde{\omega_{ijt}}$ is the random variable with mean 0 and constant standard deviation. This relationship is observed from the real data in Niger and obtained via the regression models proposed in Chapter 4. This allows us to capture the uncertainty in the demand for vaccines. The authors in [34] consider the randomness of demand for vaccines via a simulation process. This provides approximated vaccine distribution strategies. Therefore, there is a need to capture the uncertainty via exact methods. In this study we use the chance constraint method to represent the uncertainty. This method is used to ensure that the probability of satisfying a certain constraint is beyond a certain level[132]. By using the chance constraints, this method enables us to ensure that the predefined targeted immunization coverage (e.g. 90%) by WHO is met.

Figure 5.1 provides a schematic representation of the modeling approach used for this problem. This network represents a three tier supply chain with 1 central store, 2 district stores, and 1 clinic which only use refrigerators for storing the vaccines. This network is shown for one vaccine type over two time periods. The grey, green, and yellow shaded circles represent the central sore, district stores, and clinics respectively. These circles in successive time periods are connected to

Table 5.1: Decision variables and parameters.

Set of Indices	
Set	**Description**
I	Set of vaccine types
I^R	Set of vaccines that can only be stored in a refrigerator, $I^R \subset I$
J	Set of the nodes in the network
T	Set of periods in the planning horizon
Parameters	
Parameter	**Description**
μ_{ijt}	Mean demand for doses of vaccine type i at location j in time period t
$\tilde{\omega}_{ijt}$	Random variable which represents the demand for doses of vaccine type i at location j in time period t
C_j^R	Effective refrigerator capacity at location j
C_j^F	Effective freezer capacity at location j
C_{kj}^V	Effective transport capacity from location k to location j
q_i	Effective packed volume of one dose of vaccine i
r_i	Diluent volume for vaccine i
$\imath$	Number of doses administered of vaccine i within the vaccine regimen
β_{ij}	Minimum fraction of demand for vaccine i at location j that must be met each period
w_{ijt}^R	Fraction of vaccine i inventory in refrigerators lost at location j in period t
w_{ijt}^F	Fraction of vaccine i inventory in freezers lost at location j in period t
w_{ikjt}^{RR}	Fraction of vaccine i going from a refrigerator at location k to a refrigerator at location j in time period t that is lost
w_{ikjt}^{RF}	Fraction of vaccine i going from a refrigerator at location k to a freezer at location j in time period t that is lost
w_{ikjt}^{FR}	Fraction of vaccine i going from a freezer at location k to a refrigerator at location j in time period t that is lost
w_{ikjt}^{FF}	Fraction of vaccine i going from a freezer at location k to a freezer at location j in time period t that is lost
w_{ijt}^O	Fraction of open vial wastage for vaccine i at location j in time period t
ϵ	Positive constant used to ensure that maximizing the number of fully immunized children (FIC) is preferred to partial immunization
Decision variables	
Variable	**Description**
x_{ijt}^R	Units of vaccine i used from a refrigerator to satisfy demand at location j in period t
x_{ijt}^F	Units of vaccine i used from a freezer to satisfy demand at location j in period t
n_j	Number of fully immunized children (FIC) at location j
I_{ijt}^R	Inventory of vaccine i in a refrigerator at location j at end of time period t
I_{ijt}^F	Inventory of vaccine i in a freezer at location j at end of time period t
S_{ikjt}^{RR}	Units of vaccine i shipped from a refrigerator at location k to a refrigerator at location j in time period t
S_{ikjt}^{RF}	Units of vaccine i shipped from a refrigerator at location k to a freezer at location j in time period t
S_{ikjt}^{FR}	Units of vaccine i shipped from a freezer at location k to a refrigerator at location j in time period t
S_{ikjt}^{FF}	Units of vaccine i shipped from a freezer at location k to a freezer at location j in time period t

represent the inventory flow which are shown by red arcs. The black arcs represent the shipment decisions.

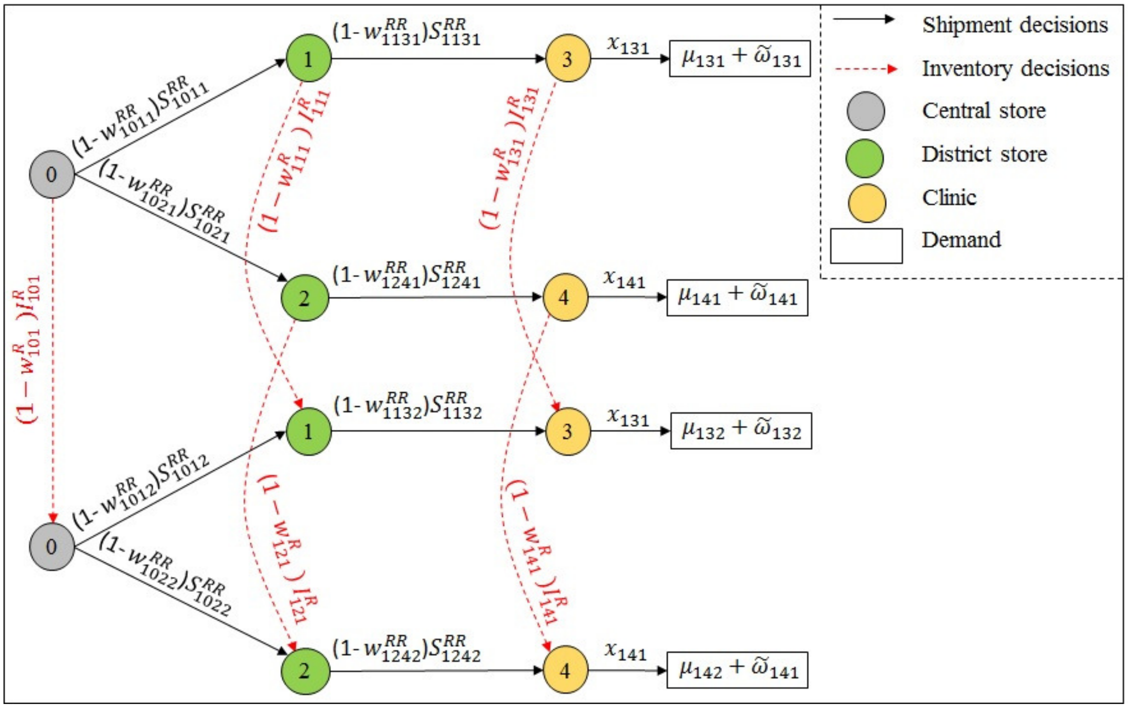


Figure 5.1: Vaccine supply chain network with 3 tiers: 1 central store, 2 district stores, and 1 clinic having only refrigerators and for 2 time periods and 1 vaccine type.

The chance-constrained programming model is defined as follows:

$$\max \sum_{j\in\mathcal{J}} n_j + \epsilon \sum_{i\in I}\sum_{j\in\mathcal{J}}\sum_{t\in T}(x^R_{ijt} + x^F_{ijt}) \tag{5.1a}$$

$s.t.$

$$I^R_{ijt} = (1 - w^R_{ijt-1})I^R_{ijt-1} + \sum_{k\in J,k\neq j}(1 - w^{FR}_{ikjt-1})S^{FR}_{ikjt-1}$$
$$+ \sum_{k\in J,k\neq j}(1 - w^{RR}_{ikjt-1})S^{RR}_{ikjt-1} - \sum_{k\in J,k\neq j} S^{RF}_{ijkt} - \sum_{k\in J,k\neq j} S^{RR}_{ijkt} - x^R_{ijt}/(1 - w^O_{ijt}) \qquad \forall i\in I, j\in J, t\in T\setminus\{0\} \tag{5.1b}$$

$$I^F_{ijt} = (1 - w^F_{ijt-1})I^F_{ijt-1} + \sum_{k\in J,k\neq j}(1 - w^{RF}_{ikjt-1})S^{RF}_{ikjt-1}$$
$$+ \sum_{k\in J,k\neq j}(1 - w^{FF}_{ikjt-1})S^{FF}_{ikjt-1} - \sum_{k\in J,k\neq j} S^{FR}_{ijkt} - \sum_{k\in J,k\neq j} S^{FF}_{ijkt} - x^F_{ijt}/(1 - w^O_{ijt}) \qquad \forall i\in I, j\in J, t\in T\setminus\{0\} \tag{5.1c}$$

$$\sum_{i\in I} q_i\left(I^R_{ijt} + \sum_{k\in J,k\neq j}(1 - w^{FR}_{ikjt})S^{FR}_{ikjt} + \sum_{k\in J,k\neq j}(1 - w^{RR}_{ikjt})S^{RR}_{ikjt}\right) \leq C^R_j \qquad \forall j\in J, t\in T \tag{5.1d}$$

$$\sum_{i\in I} q_i\left(I^F_{ij} + \sum_{k\in J,k\neq j}(1 - w^{RF}_{ikjt})S^{RF}_{ikjt} + \sum_{k\in J,k\neq j}(1 - w^{FF}_{ikjt})S^{FF}_{ikjt}\right) \leq C^F_j \qquad \forall j\in J, t\in T \tag{5.1e}$$

$$I^R_{ij0} = 0 \qquad \forall i\in I, j\in J \tag{5.1f}$$

$$I^F_{ij0} = 0 \qquad \forall i\in I, j\in J \tag{5.1g}$$

$$I^F_{ij|T|} = 0 \qquad \forall i\in I^R, j\in J \tag{5.1h}$$

$$\sum_{i\in I} q_i(S^{RR}_{ikjt} + S^{RF}_{ikjt} + S^{FR}_{ikjt} + S^{FF}_{ikjt}) \leq C^V_{kj} \qquad \forall j,k\in J; j\neq k, t\in T \tag{5.1i}$$

$$n_j \leq \sum_{t\in T}(x^R_{ijt} + x^F_{ijt})/a_i \qquad \forall i\in I, j\in J \tag{5.1j}$$

$$P\left(x^R_{ijt} + x^F_{ijt} \geq \mu_{ijt} + \tilde{\omega}_{ijt}\right) \geq \beta_{ij} \qquad \forall i\in I, j\in J, t\in T \tag{5.1k}$$

$$\text{All Variables} \geq 0 \tag{5.1l}$$

In the above model, the first term in the objective function maximizes the number of FIC in each clinic. This results in maximizing the number of required doses in the vaccine regimen that are sent downward to the clinics. The second term maximizes the total number of doses used for partial immunization. The coefficient ϵ is used to ensure that maximizing the number of FIC is preferred to partial immunization. In other words, the second term represents the flow of vaccines in the supply chain.

Equations (5.1b) and (5.1c) are the inventory balance constraints for refrigerators and freezers at each facility in the supply chain. It is assumed that the lead time for shipping is one week. The model accounts for specific shipment schedules between locations j and k by fixing variables S_{ikjt}^{RR}, S_{ikjt}^{RF}, S_{ikjt}^{FR}, and S_{ikjt}^{FF} to zero during those time period when there are no shipments. Doing this, allows us to model situtatios, such as, Niger's vaccine supply chain where shipments between the central store and regional stores are currently scheduled every three months.

Constraints (5.1d) and (5.1e) limit the storage capacity of refrigerators and freezers, respectively. Constraints (5.1f) and (5.1g) initialize the inventory of vaccines in refrigerator and freezer, respectively. Constraint (5.1h) guarantees that vaccines that are not stable at freezer temperature must only be stored in refrigerators. Constraint (5.1i) limits the transportation capacity between two locations. Note that the model assumes that both freezable and non-freezable vaccines are shipped via the same vehicle type. Constraint (5.1j) determines FIC for each location. The right hand side of this inequality represents the average number of individuals to whom a full set of doses of vaccine i is administered at clinic j over the planning horizon. This is found by dividing the total number of vaccine i doses administered at clinic j by the number of required doses defined in the vaccine regimen. Therefore, FIC is bounded by the smallest value among all vaccine types.

Finally, (5.1k) is a chance constraint which indicates that the total number of vaccine type i delivered to clinic j in period t should be greater than the total demand for vaccines at least β_{ij} of the time. This constraint captures the random nature of patient arrivals. All of the children would not be vaccinated if the available inventory is less than the total demand. The goal is to to maintain the necessary inventory to vaccinate at least β_{ij} (i.e. 90%) of the patients.

5.3 Approximating The Chance Constraints

The proposed model is a CPP Due to constraints (5.1k). CPP models are challenging to solve due to the non-convexity of the feasible region [111]. One of the methods to which is frequently used to solve CPPs is the SAA. This method represents the uncertain parameter $\tilde{\omega}_{ijt}$ via a finite number of realizations (scenarios) [111]. Let Ω denote this finite set of scenarios.

Using SAA, we approximate model (5.1) with the corresponding DEF. This DEF is a linear program. To develop DEF, we introduce additional variables and constraints. We introduce the slack variables and replace (5.1k) with (5.2b) and (5.2c). In these new equations, the stochastic parameter $\tilde{\omega}_{ijt}$ is substituted with the corresponding realization ω_{ijt} obtained from the random sample $s \in \Omega$. One could use Monte Carlo simulation to randomly generate the observations. The additional term in the objective function minimizes the cost of violating constraints (5.1k) [88, 111].

We use π_{ijt} to denote the penalty for not serving demand of vaccine i at clinic j in time period t. Thus, we can build the DEF of model (5.1) as following.

$$\max \qquad \sum_{j\in\mathcal{J}} n_j + \epsilon \sum_{i\in I} \sum_{j\in\mathcal{J}} \sum_{t\in T} \Bigg((x^R_{ijt} + x^F_{ijt}) - \sum_{s\in\Omega} \pi_{ijt} \mathcal{V}^s_{ijt} \Bigg) \tag{5.2a}$$

$$s.t. (5.1b) - (5.1g), (5.1i) - (5.1l)$$

$$x^R_{ijt} + x^F_{ijt} + \mathcal{V}^s_{ijt} - \mathcal{W}^s_{ijt} = \mu_{ijt} + \omega^s_{ijt} \qquad \forall i \in I, j \in J, t \in T, s \in \Omega \tag{5.2b}$$

$$\mathcal{V}^s_{ijt} \mathcal{W}^s_{ijt} \geq 0 \qquad \forall i \in I, j \in J, t \in T, s \in \Omega \tag{5.2c}$$

Identifying the values for the penalty cost π is challenging. $\pi_{ijt} = 0$ results in $V^s_{ijt} > 0$ for all $s \in \Omega$, and consequently $x^R_{ijt} + x^F_{ijt} \leq \mu_{ijt} + \omega^s_{ijt}$ for all $s \in \Omega$. $\pi_i jt = \inf$ results in $V^s_{ijt} = 0$ for all $s \in \Omega$, and consequently $x^R_{ijt} + x^F_{ijt} \geq \mu_{ijt} + \omega^s_{ijt}$ for all $s \in \Omega$. Thus, the value of π_{ijt} should be determined so that constraints (5.2b) are satisfied at least in (β)% of the scenarios generated. To achieve this, we propose a binary search algorithm which identifies the value of π_{ijt} such that the number of FIC is maximized and constraints (5.2b) are violated in β_{ij}% of the scenarios generated. In this algorithm we solve problem (5.2) iteratively until $M_{ijt} = \alpha|\Omega| \pm \sigma$, $\forall i \in I, j \in J, t \in T$.

Proposition 2 *Algorithm* (4) *generates a lower bound for problem* (5.1).

Proof: Algorithm (4) iteratively changes the values of π_{ijt} so that $P\Bigg(x^R_{ijt} + x^F_{ijt} \geq \mu_{ijt} + \tilde{\omega}_{ijt} \Bigg) \geq \beta_{ij}$ for $\tilde{\omega}_{ijt} \in \Omega$. If Ω were to represent all the potential realizations of random variable $\tilde{\omega}_{ijt}$, then, the solution obtained by the Algorithm (4) would be optimal for (5.1) as σ goes to zero. Therefore, Algorithm (4) provides only a lower bound to the problem (5.1). ■

5.4 Numerical Experiments

The stochastic model presented in §(5.2) is used to analyze the current vaccine distribution network in Niger. The vaccine supply chain in Niger comprises of four tiers. The central government

Algorithm 4 Binary Search Algorithm

Notation: Let π^L_{ijt} and π^U_{ijt} denote the lower and upper bounds of the penalty term π_{ijt}. Also, let σ and δ be small positive constants.

while $|\pi_{ijt} - \frac{\pi^L_{ijt}+\pi^U_{ijt}}{2}| > \delta \quad \forall i \in I, j \in J, t \in T$ **do**
  $\pi_{ijt} \leftarrow \frac{\pi^L_{ijt}+\pi^U_{ijt}}{2}$
  Solve (5.2) and let $\bar{\mathcal{W}}^s_{ijt}$, $\forall i \in I, j \in J, t \in T$ be the corresponding incumbent solution.
  $M \leftarrow 0$, $N \leftarrow 0$
  for $i \in I, j \in J, t \in T$ **do**
    for $s \in \Omega$ **do**
      if $\bar{\mathcal{W}}^s_{ijt} > 0$ **then**
        $M_{ijt} \leftarrow M_{ijt} + 1$
      end if
    end for
    if $M_{ijt} \geq \alpha|\Omega| + \sigma$ **then**
      $\pi^L_{ijt} \leftarrow \frac{\pi^L_{ijt}+\pi^U_{ijt}}{2}$
    else if $M_{ijt} \leq \alpha|\Omega| - \sigma$ **then**
      $\pi^U_{ijt} \leftarrow \frac{\pi^L_{ijt}+\pi^U_{ijt}}{2}$
    end if
  end for
end while
Return π_{ijt}, and the solution to (5.2)

of Niger purchases vaccines from UNICEF headquarters in Denmark. UNICEF replenishes the inventory of the central store every two month via plane. Vaccines are shipped from the central store to the regional stores every three month via cold trucks. The district stores pick up the vaccines from the regional stores on a monthly basis via 4×4 trucks. The Integrated Health Centers (IHC) or clinics then pick up the vaccines from the district stores monthly using motorcycle, bikes, or private cars. This network comprises of 8 regional stores, 42 district stores, and 642 clinics. The central and regional stores are equipped with cold rooms. While the district stores have chest refrigerators and freezers, the clinics have smaller refrigerators and/or freezers.

Public health authorities in Niger are interested in evaluating the impact of converting the current four-tier supply chain to a three-tier one by removing the regional stores, and changing Measles's vial size from multi-dose to single-dose on the vaccine availability and immunization coverage. We use the proposed model in (5.2) to address these questions. To this end, we use two performance measures: (a) Supply Ratio (SR) for each clinic, and (b) Average percentage of fully immunized children (FIC) for each region. These measure are calculated using the following equations.

$$SR_i^s = \left(\sum_{j \in J} \sum_{t \in T} (x_{ijt}^R + x_{ijt}^F) \right) \Big/ \sum_{j \in J} \sum_{t \in T} (\mu_{ijt} + \tilde{\omega}_{ijt}^s), \forall i \in I, s \in \Omega, \tag{5.3a}$$

$$FIC_j^s = 100 n_j \Big/ \phi_j^s, \forall j \in J, s \in \Omega, \tag{5.3b}$$

where ϕ_j^s is the total number of children vaccinated at clinic j and is equal to $x_{ijt}^R + x_{ijt}^F - W_{ijt}^s$.

Furthermore, we illustrate the value of stochastic solution (VSS) by comparing with the expected value solution.

5.4.1 Input Data

In our numerical experiments, we consider EPI vaccines which consist of Bacillus Calmette-Gurin (BCG), Tetanus, Measles, Oral Polio, Yellow Fever, and Diphtheria-Tetanus-Pertussis (DTP). The characteristics of each vaccine type is summarized in Table (1). These inputs are obtained from [34].

Table 5.2: Vaccine characteristics.

Vaccine type	Num. of doses per vial	Volume (*c.c./dose*)	Dilunet volume (*c.c./dose*)	Num. of doses in regimen	Storage
BCG	20	1.2	0.7	1	Refrigerator/freezer
Tetanus	10	3.0		3	Refrigerator
Measles	10	2.1	0.5	2	Refrigerator
Oral polio	20	1.0		4	Freezer
Yellow fever	10	2.5	6.0	1	Refrigerator/freezer
DTP	1	16.8		3	Refrigerator

To estimate the monthly demand for vaccines in different regions of Niger, we conducted a comprehensive data analysis which is presented in Chapter 4. Via this analyses, we fitted regression models to the available data about monthly demand for vaccines to identify factors that impact the demand and represent the demand as a function of known parameters. The regression indicated that demand is a function of population size and as the value of the independent variable, population size, increases, the mean of the dependent variable, demand for vaccine, also tends to increase. This allows us to estimate the mean demand for vaccine, μ_{ijt} given the population size. Moreover, $\tilde{\omega}_{ijt}$ represents the error term in the regression function. Our analysis indicates that this error term follows the Normal distribution with mean 0 and a constant standard deviation. These values differ by regions and vaccine types. Note that, population data is available only at the regional level, thus

Table 5.3: Relative distance between objective function values for different number of scenarios.

S;S'	Relative distance between objective function values
30;40	0.010
40;50	0.008

we assume that the clinics in a district serve equal proportions of the demand. The storage capacity at each location is estimated based on an study by [13].

5.4.2 Experimental Setup

Problem formulation (5.2) is developed using JuMP programming language and solved using Gurobi 7.0.2. The experiments are implemented using high-performance computing (HPC) resources at Clemson University. We consider a total horizon of 12 months for the analyses of Niger's supply chain. Furthermore, to implement Algorithm (4), both δ and σ are set to 0.01.

5.4.2.1 Scenario Selection:

To identify an appropriate number of scenarios for our experimentations we use the following method. This methods analyze the difference between the optimal solutions obtained when using different number of scenarios S and S'. This approach finds the relative distance between the two objective function values. A small value implies that the solutions are statistically indistinguishable. Table 5.3 presents the results.

We observe that increasing the number of scenarios from 40 to 50 results in a relative distance smaller than 0.01 . Thus, it is not necessary to increase the number of scenarios used from 40 to 50. This is the reason why in our experiments we use 40 scenarios.

5.4.3 VSS

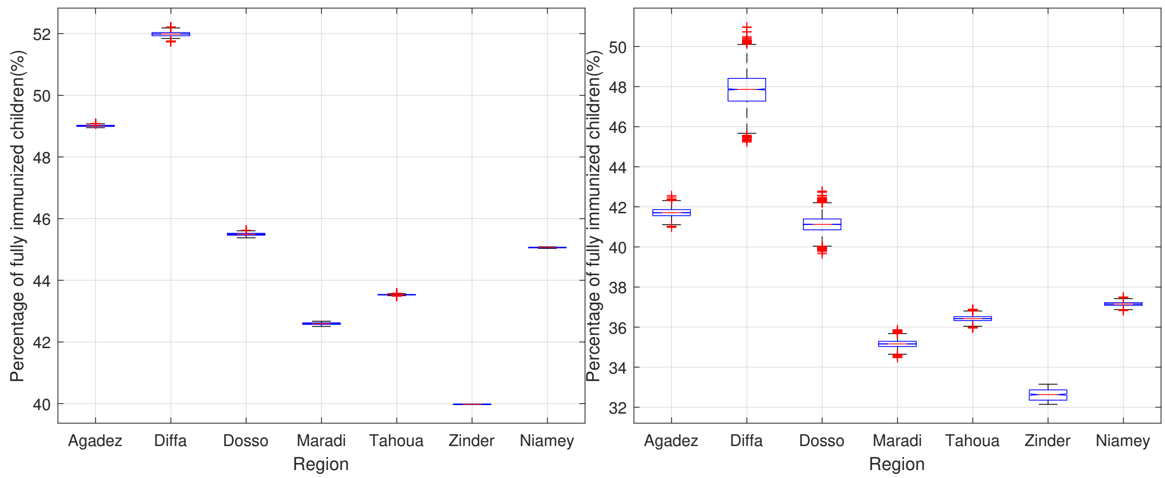
In this section we evaluate the performance of stochastic solutions obtained from solving problem (5.2). To this goal, we compute the VSS for the current vaccine supply chain in Niger. VSS measures the impact of random demand for vaccine to the performance of the system [20]. VSS is is calculated as the difference between the objective function value of (5.2) and the expected value solution (EEV). To calculate EEV we solve the deterministic version of our problem which uses the expected value of demand for vaccines in a time period. Then, we use the total number of vaccine type i delivered to clinic j in period t obtained from the deterministic solution to resolve the stochastic model. The corresponding objective function value is EEV. The objective function value of (5.2) equals $11,126,993$ and EEV equals $-33,458,721$. Therefore, VSS is equal to $(-33,458,721) - (11,126,993) = -44,585,714$. VSS indicates that, on the average, the deterministic solution overestimates the number of children fully immunized by as much as 40%. This indicates that there is value in solving the proposed stochastic model formulation rather than the corresponding mean value problem.

5.4.4 Analyzing The Current System

To analyze the capabilities of the current system, we used FIC across all regions summarized in Figure 5.2a for low standard deviation of demand for vaccines and Figure 5.2b for high standard deviation. The results indicate that the average FIC does not exceed 52%, which implies low vaccination coverage in the country. Due to the stochastic nature of patient arrivals to the clinics, as well as, storage capacity limitations, fully immunizing children, as recommended by the vaccine regimen, is challenging. This analyses provides a tool to aid public health authorities estimate the number of additional cold rooms, refrigerators and freezers needed to achieve higher immunization coverage. Moreover, when the standard deviation of patient arrivals to the clinics is high the vaccination coverage decreases. This motivates policy makers to develop policies which encourage high participation in vaccination.

5.4.5 Removing Regional Level

In this section we explore the effects of removing the regional stores from the current four-tier Niger's vaccine supply chain on FIC and SR. In the three-tier hierarchy, the district stores receive shipments of vaccines directly from the central store. Also, the cold rooms at the regional stores are eliminated to save costs. Figure (5.3) shows the countrywide percentage of FIC for the four-tier and the three-tier supply chains. The results indicate that the three-tier supply chain design increases the FIC from 44.28% to 44.41%. A paired t-test at $p - value = 0.05$ was conducted to determine if there is a statistically significance difference between the two designs. The results indicated that, the difference in the percentage of FIC in these supply chains is statistically insignificant. This observation is similar to the results presented in [34].



(a) Low standard deviation. (b) High standard deviation.

Figure 5.2: Regional FIC for the current supply chain for low and high standard deviations of demand for vaccine.

5.4.6 Changing the Vial Size of Measles Vaccine

Health care administrators suggest replacing multi-dose vaccine vials with single-dose vials to improve safety and eliminate OVW. Since single-dose vials are accessed only once, the chance that needles are contaminated is minimized. In this section, we analyze the impact of replacing ten-dose Measles vials with single-dose vials on FIC and SR. The results are illustrated in Figure (5.5). Figure 5.4 indicates that the average countrywide FIC percentage decreases by about 41%, from 43.76% to 30.85% with single-dose vials. Figure 5.5b indicates that that SR of every vaccine is reduced, including Measles. In particular, SR decreases by about 13% for Oral polio, 21% for BCG, 23% for Yellow fever, 23% for Measles, and 24% for DTP when single-dose vials of Measles vaccine are used. One might expect that eliminating OVW would increase vaccine availability. However, single-dose vaccine vials are presented in larger volumes and require additional storage space. This increase in space requirement for Measles impacts the availability of storage space for other vaccines in the case when storage space is limited.

5.5 Conclusion and Future Research

This paper presents a stochastic optimization model that identifies distribution strategies for vaccine vials in developing countries. This is a data driven model which is developed using real

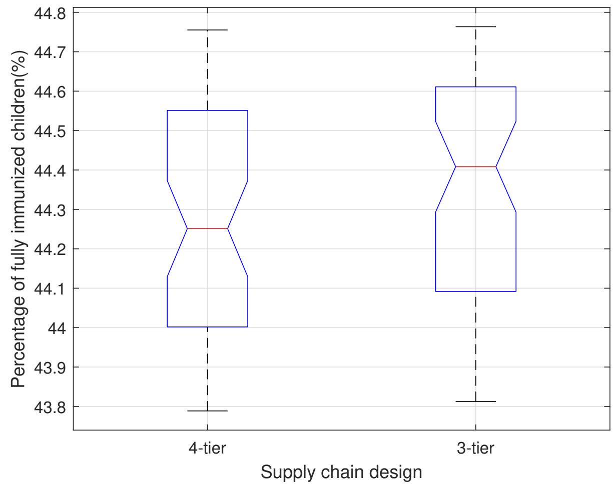


Figure 5.3: Countrywide FIC for the four- and three-tier vaccine supply chains.

life data from Niger. To the best of our knowledge, this is the first stochastic optimization model used in the literature to capture vaccine distribution decisions in the supply chain. The model captures uncertainties of demand for vaccination via chance constraints. Public health authorities can use this model to evaluate different supply chain designs and vaccine administration policies.

After developing a case study using real-world data from Niger, a sensitivity analysis is conducted to evaluate the impact of converting the current four-tier to a three-tier supply chain by removing regional stores, and changing Measles's vial size from multi-dose to single-dose on FIC and SR. Our observation can be summarized as follows:

1. Changing vaccine supply chain hierarchy dose not impact immunization rates and vaccine availability.

2. Replacing ten-dose vials of Measles vaccine with single-dose vials reduces immunization coverage rate and supply ratio for all vaccine types including Measles itself.

Removing regional stores decreases the inventory holding holding costs, but increases transportation costs due to longer trips. Therefore, an economic analysis is required to identify the trade-offs between inventory holding and transportation costs. We plan to extend the model to evaluate the economical impact of eliminating regional stores. Moreover, the proposed model only captures the randomness in patient arrivals. We plan to extend the model to evaluate the impact of other stochastic problem parameters like capacity utilization on vaccination coverage.

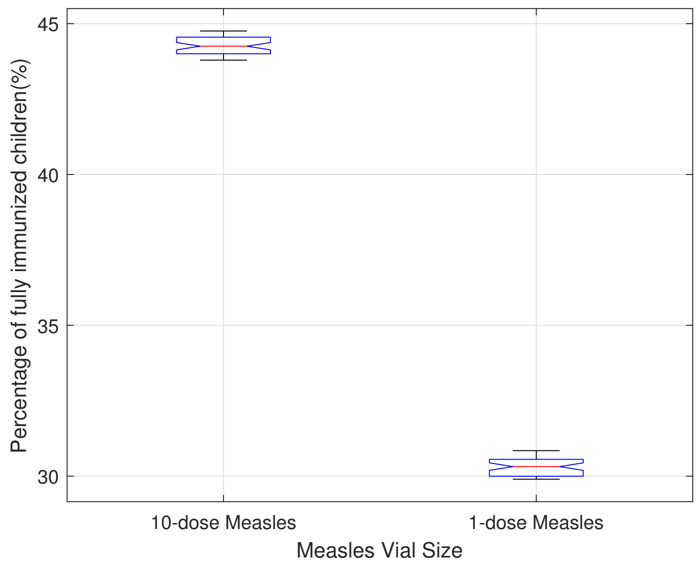
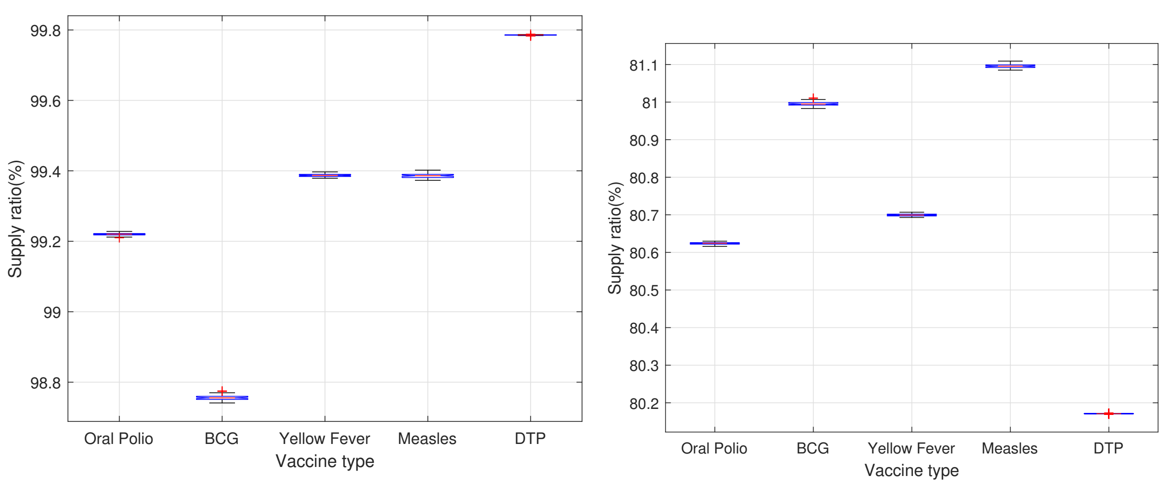


Figure 5.4: Countrywide FIC for ten-dose Measles vials and single-dose Measles vials.



(a) SR (ten-dose Measles).

(b) SR (single-dose Measles).

Figure 5.5: SR for all vaccine types when either ten-dose Measles vials are used or single-dose vials.

Chapter 6

Conclusions and Future Research

The cost, environmental and social impacts of perishable products wastage have drawn global attention. While a number of efficient models and algorithms to manage these products exist, shortage of supply and product's wastage are still challenging issues. This study contributes to various research areas, such as, supply chain design, inventory management, pricing, and stochastic optimization.

6.1 Research Summary

In my dissertation I focus on three major areas of research. The goal of these studies is to optimize inventory replenishment decisions for perishable products.

First, we presented a 2-SIP model for inventory replenishment and the administration of childhood vaccines in targeted outreach immunization sessions. To our knowledge, this is the first stochastic optimization model which captures the relationships that exist among these decisions. The proposed model minimizes replenishment and OVW costs. Different from the current practice which relies on the use of a single multi-dose vial, this study models the performance of an inventory replenishment policy that allows the use of mix of multi-dose vials for vaccination. Motivated by the solutions obtained from the proposed model we developed simple-to-implement vaccine administration policies. Statistical analysis indicated that the performance of some policies is not different than the optimal policy and others outperform the optimal policy. In order to solve the proposed 2-SIP, the LS method was extended by incorporating GMI and MRI cuts in the first-stage problem.

Via an extensive numerical study we showed that the proposed algorithm is scalable; it outperforms the LS method by providing high quality solutions in a much shorter CPU time.

Second, we proposed a two-stage stochastic optimization model that identifies a replenishment schedule for a periodic-review inventory system for perishable products with non-stationary demand and dual sourcing. The model captures the relationship between price and stochastic demands via a linear function. The model considers a price markdown as the means to stimulate demand and minimize waste of this perishable product. In the proposed model, the first-stage problem is bilinear. Thus, we developed a solution approach which extends the LS method by employing a piecewise linear approximation of the bilinear term in order to solve the first-stage problem. We developed a case study in order to validate the model and evaluate its performance. We conducted a thorough sensitivity analysis to observe the impact that timing and size of a price markdown has on inventory replenishment decisions and retailer's profits. We analyzed the impact of deterioration rate, inventory holding cost, and service level on inventory replenishment decisions and retailer's profits. Via this model we also evaluated the impact of dual sourcing in replenishment decisions. While the relationships identified via this analysis are intuitive, quantifying these relationships cannot easily be accomplished without the aid of models similar to the one proposed in this research.

Finally, we proposed a data-driven stochastic optimization model that identifies distribution strategies for vaccine vials. To the best of our knowledge, this is the first stochastic optimization model that integrates transportation, inventory and replenishment decisions of vaccine supply chain. The model captures uncertainties of demand for vaccination via chance constraints. The objective is to maximize the number of fully immunized children in developing countries. Public health authorities can use this model to evaluate the impact of different supply chain designs on immunization coverage; evaluate the impact of introducing a new vaccine on vaccine inventory at a clinic; and develop vaccine administration policies which reduce OVW. We developed a case study using real-life data from Niger. We conducted an extensive statistical analysis of the data in order to identify the factors which impact immunization in different regions of Niger. Population size, poverty and education levels do impact expected demand for vaccination. Analyzing the national level data suggested that the country-wide population size does not significantly impact the expected demand for vaccination. This highlights the importance of developing a single, region-based regression model as presented in this study. The results of region-based regression models were incorporated on the optimization model. This model allowed us to evaluate the impact of converting the current four-tier supply chain to a three-tier one by removing the regional stores , and changing Measles's vial size from multi-dose to single-dose on FIC and SR.

Overall, the main contributions of this dissertation can be summarized as follows:

1. Development of a two-stage SP model for integrating vial replenishment and vaccine administration which captures (a) the order frequency for respective quantities of different-sized vials, (b) the opening schedule for these vials, and (c) the administration of available doses to patients.

2. Development of simple to use and economic vaccine administration policies for outreach sessions. The performance of these policies is evaluated and benchmarked with existing practice via an extensive simulation analysis.

3. Development of a new solution approach for two-stage stochastic integer programs (2-SIP) with continuous recourse by using GMI and MIR cuts to address the non-convexity of the first-stage problem.

4. Development of a two-stage stochastic optimization model that integrates inventory replenishment and pricing decisions for age-dependent perishable products in a periodic-review inventory system.

5. Development of a solution approach for a two-stage stochastic, bilinear model with linear recourse by using extensions of McCormick relaxation to approximate the non-linear first stage problem.

6. Development of a regression-based estimate of the future demand for CIV for each vaccine type and region in Niger.

7. Development of a data-driven chance constrained programming model for the vaccine supply chain in developing countries.

8. Application of the data-driven chance constrained programming model to manage transportation, inventory and OVW of vaccine supply chain in Niger.

6.2 Future Research

In the future we plan to extend this research in two directions. First, we plan to investigate methods which will lead to an improvement of the algorithms proposed for solving 2-SIPs. Second, we plan to model and solve other supply chains of perishable products.

Algorithm: Over the past several decades, different algorithms have been proposed to solve two-stage stochastic programs. To achieve computational tractability, many of these methods represent uncertainty through a finite number of realizations or scenarios. In a scenario-based approach, different realizations of uncertainty are associated with a probability of occurrence. However, the probability of strategic uncertainties, such as earthquakes, wars, and pandemic outbreaks, are not easy to determine. Moreover, no clear guidelines determine the number of scenarios used. Thus, we plan to further the study of sequential sampling algorithms, such as the two-stage and multiple-stage stochastic decomposition (SD) method. Current methods assume linearity in order to preserve the convexity of the problem. However, only a few works assume integer and binary first-stage variables, and, to the best of our knowledge, no algorithms exist for problems with integer and binary second-stage problems. We plan to tackle these problems and contribute to the literature by incorporating cuts such as, GMI, MIR, etc., which linearize these nonconvexities.

Application: We plan to extend our modeling framework to consider pharmaceutical logistics. Pharmaceutical products are perishable, and their demand is stochastic. Logistic costs have been identified as one of the largest expense for hospitals. These costs constitute up to 40% of the total operating budget because of the high prices of prescription drugs [120]. In the existing distribution networks, each hospital pharmacy typically uses its own wholesaler to replenish its inventory. This decentralized decision-making approach, while giving the pharmacies flexibility to select suppliers and determine inventory-replenishment schedules, is costly. Ample literature exists on supply chain optimization and several successful implementations of these strategies in the retail sector, such as at Walmart and Amazon. However, implementing these models in healthcare should be done cautiously because the objective of retailers is to maximize their profits. However, a public, not-for-profit academic healthcare delivery systems strategy is to provide medical excellence through clinical care, education, and research. This strategy will impact the supply chain decisions. We plan to extend the current work by building stochastic optimization models for pharmaceuticals.

Appendices

Appendix A Two-stage stochastic programming model

Table 1: Decision variables and parameters.

Set of Indices	
Set	Description
$\mathcal{T}$	Set of ordering decision epochs
$\mathcal{N}$	Set of consumption decision epochs
$\mathcal{V}$	Set of vials of different sizes
Parameters	
Parameter	Description
N	Number of consumption decision epochs within an ordering period
T	Number of ordering decision epochs
f_t	Ordering fixed cost at time period t
$c_{\nu t}$	Variable purchase cost of vial size ν at time period t
$d_{\nu t}$	Unit inventory holding cost of vial size ν at time period t
q_ν	Number of doses in vial ν
$\tilde{\omega}$	Random patient arrival
M_t	Limit on the number of vials ordered in time period t
τ	Safe open vial use time in periods
g	Unit wastage cost
p	Unit penalty cost of an unserved patient
Decision variables	
Variable	Description
z_t	A binary decision variable which takes 1 if an order is placed at time period t and takes 0 otherwise
$r_{\nu t}$	Replenishment quantity for vial size ν at time period t
$u_{\nu n}$	Number of vials of size ν to open at time period n
$s_{\nu n}$	Number of vials of size ν in the inventory at time period n
y_{nm}	Number of doses obtained from vials opened in period n, and used in period m

$$\min \sum_{t\in\mathcal{T}}(f_t z_t + \sum_{\nu\in\mathcal{V}} c_\nu r_{\nu t}) + \sum_{\nu\in\mathcal{V}}\sum_{n\in\mathcal{N}} d_\nu s_{\nu n} + \mathbb{E}\{H(x,\omega)\}, \tag{1a}$$

$s.t.$

$$s_{\nu Nt} = s_{\nu(Nt-1)} + r_{\nu t} - u_{\nu Nt} \qquad \forall t \in \mathcal{T}, \tag{1b}$$

$$s_{\nu n} = s_{\nu n-1} - u_{\nu n} \qquad \forall n \in \mathcal{N} \setminus \{N, 2N, \dots, TN\}, \tag{1c}$$

$$\sum_{\nu\in\mathcal{V}} r_{\nu t} \le M_t z_t \qquad \forall t \in \mathcal{T}, \tag{1d}$$

$$z_t \in \{0,1\}; s_{\nu n}, u_{\nu n}, r_{\nu t} \in \mathbb{Z}^+ \qquad \forall t \in \mathcal{T}, \nu \in \mathcal{V}, n \in \mathcal{N}, \tag{1e}$$

where,

$$H(x,\omega) = \min \sum_{n\in\mathcal{N}} (gy_{n(n+\tau)} + p\ell_n) \tag{2a}$$

$$s.t.$$

$$\sum_{m=n}^{n+\tau-1} y_{nm} + y_{n(n+\tau)} = \sum_{\nu\in\mathcal{V}} q_\nu u_{\nu n} \qquad \forall n \in \mathcal{N}, \tag{2b}$$

$$\sum_{m=n-\tau+1}^{n} y_{mn} + \ell_n = \omega \qquad \forall n \in \mathcal{N}, \tag{2c}$$

$$y_{mn}, \ell n \in \mathbb{Z}^+ \qquad \forall m, n \in \mathcal{N}. \tag{2d}$$

Appendix B Performance evaluation of stochastic solutions

In this section we evaluate the performance of stochastic solutions obtained from solving the 2-SIP with a first-stage problem (2.6) and a second-stage problem (2.7). To this goal, we compute metrics such as expected value of perfect information (EVPI) and the value of stochastic solution (VSS). We obtain EVPI and VSS for a problem instance created by using data from the Barisal region. The number of scenarios generated for these experiments is $S = 1000$.

EVPI measures the price one is willing to pay to gain access to perfect information. EVPI is the difference between the objective function value of the wait-and-see and here-and-now problems. To calculate the objective function value of the wait-and-see problem, we solve the SAA problem in (2.9) for each single scenario separately. Next, we calculate the corresponding expected value over the scenarios generated. The here-and-now problem is indeed the SAA problem in (2.9). We find $EVPI = \$3300 - \$2870 = \$428$. This means, if the number of patients arriving in a session is known, the total costs would only be \$2870. Thus, the cost of not knowing the future is $EVPI = \$428$.

VSS measures the impact of random patient arrivals to the performance of the system [20]. VSS is the difference between the objective function value of SAA problem in (2.9) and the deterministic mean value problem. To compare the objective function values of SAA and deterministic mean value problem, we initially solve both problems for a given set of scenarios. Next, we fix the values of the first-stage solutions, and simulating the second-stage problem (2.7) in a different set of scenarios. The corresponding objective function values are reported in Figure 1. We observe less variations in the objective function values obtained from the stochastic solution. Also, the average objective function value of stochastic solution is lower than solutions to the mean value problem. These results indicate that there is value in solving the proposed stochastic model formulation (rather than the corresponding mean value problem).

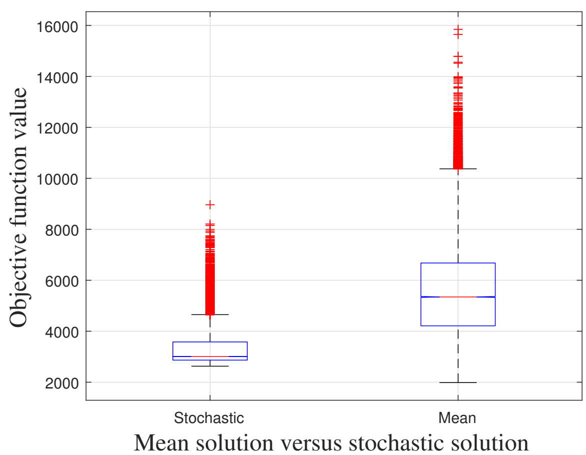


Figure 1: The value of stochastic solution.

Appendix C McCormick relaxations

The following are the McCormick relaxations of the bilinear terms $W = p^n p^o, Y_{it} = p^n y_{it}, V_{it} = p^o y_{it}$.

$$W \geq \underline{p}^o p^n + \underline{p}^n p^o - \underline{p}^o \underline{p}^n \tag{3a}$$

$$W \geq \overline{p}^o p^n + \overline{p}^n p^o - \overline{p}^o \overline{p}^n \tag{3b}$$

$$W \leq \underline{p}^o p^n + \overline{p}^n p^o - \underline{p}^o \overline{p}^n \tag{3c}$$

$$W \leq \overline{p}^o p^n + \underline{p}^n p^o - \overline{p}^o \underline{p}^n \tag{3d}$$

Piecewise linear approximation via bivariate partitioning

$$p^n = \textstyle\sum_{i=1}^{N^n} d^n B_i^n, \tag{4a}$$

$$p^o = \textstyle\sum_{j=1}^{N^o} d^o B_j^o, \tag{4b}$$

$$W = \textstyle\sum_{i=1}^{N^n} \sum_{j=1}^{N^o} d^n d^o \Upsilon_{ij}, \tag{4c}$$

$$\textstyle\sum_{j=1}^{N^o} \Upsilon_{ij} = B_i^n, \qquad \forall i = 1, \ldots, N^n, \tag{4d}$$

$$\textstyle\sum_{i=1}^{N^n} \Upsilon_{ij} = B_j^o, \qquad \forall j = 1, \ldots, N^o, \tag{4e}$$

$$\textstyle\sum_{i=1}^{N^n-1} E_i^n = 1, \tag{4f}$$

$$\textstyle\sum_{j=1}^{N^o-1} E_j^o = 1, \tag{4g}$$

$$B_1^n \leq E_1^n, \tag{4h}$$

$$B_i^n \leq E_{i-1}^n + E_i^n, \qquad 2 \leq i \leq N^n - 1, \tag{4i}$$

$$B_{N^n}^n \leq E_{N^n-1}^n, \tag{4j}$$

$$B_1^o \leq E_1^o, \tag{4k}$$

$$B_i^o \leq E_{i-1}^o + E_i^o, \qquad 2 \leq i \leq N^o - 1, \tag{4l}$$

$$B_{N^o}^o \leq E_{N^o-1}^o, \tag{4m}$$

$$E_i^n, E_j^o, B_i^n, B_j^o \in \{0, 1\}, \quad \forall i = 1, \ldots, N^n, j = 1, \ldots, N^o \tag{4n}$$

(4o)

Bibliography

[1] Monitoring vaccine wastage at country level: guidelines for programme managers. 2005.

[2] Vaccine volume calculator; 2009. Technical report, WHO World Health Organization Geneva., 2012.

[3] Global immunization data. Technical report, WHO world Health Organization and others, February 2014.

[4] United states 2030 food loss and waste reduction goal. `https://www.epa.gov/sustainable-management-food/united-states-2030-food-loss-and-waste-reduction-goal#main-content`, 2015. Accessed: 2017-02-13.

[5] Demographic health survey. 2017.

[6] Unesco institute for statistics, glossary, 2018.

[7] Who and unicef estimates of immunization coverage: 2016 revision, 2018.

[8] World bank, population, total, 2018.

[9] Claire S Adjiman, Ioannis P Androulakis, and Christodoulos A Floudas. Global optimization of mixed-integer nonlinear problems. *AIChE Journal*, 46(9):1769–1797, 2000.

[10] Claire S Adjiman, Stefan Dallwig, Christodoulos A Floudas, and Arnold Neumaier. A global optimization method, αbb, for general twice-differentiable constrained nlpsi. theoretical advances. *Computers & Chemical Engineering*, 22(9):1137–1158, 1998.

[11] Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. Network flows. Technical report, DTIC Document, 1988.

[12] Pedro Amorim, Eduardo Curcio, Bernardo Almada-Lobo, Ana PFD Barbosa-Póvoa, and Ignacio E Grossmann. Supplier selection in the processed food industry under uncertainty. *European Journal of Operational Research*, 252(3):801–814, 2016.

[13] Tina-Marie Assi, Shawn T Brown, Ali Djibo, Bryan A Norman, Jayant Rajgopal, Joel S Welling, Sheng-I Chen, Rachel R Bailey, Souleymane Kone, Hailu Kenea, et al. Impact of changing the measles vaccine vial size on niger's vaccine supply chain: a computational model. *BMC Public Health*, 11(1):425, 2011.

[14] Tina-Marie Assi, Shawn T Brown, Souleymane Kone, Bryan A Norman, Ali Djibo, Diana L Connor, Angela R Wateska, Jayant Rajgopal, Rachel B Slayton, and Bruce Y Lee. Removing the regional level from the niger vaccine supply chain. *Vaccine*, 31(26):2828–2834, 2013.

[15] Baris Ata, Deishin Lee, and Mustafa H Tongarlak. *Got local food?* Harvard Business School, 2012.

[16] World Data Atlas. Niger data portal, 2018.

[17] Monique Bakker, Jan Riezebos, and Ruud H Teunter. Review of inventory systems with deterioration since 2001. *European Journal of Operational Research*, 221(2):275–284, 2012.

[18] Pietro Belotti, Pierre Bonami, Stefan Vigerske, and Andreas Waechter. Couenne an exact solver for nonconvex minlps. 2006.

[19] Jacques F Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik*, 4(1):238–252, 1962.

[20] John R Birge and Francois Louveaux. *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[21] Asit K. Biswas. India must tackle food waste. Technical report, World Economic Forum, 2014.

[22] Merve Bodur, Sanjeeb Dash, Oktay Günlük, and James Luedtke. Strengthened benders cuts for stochastic integer programs with continuous recourse. *INFORMS Journal on Computing*, 29(1):77–91, 2016.

[23] William K Bosu, Doris Ahelegbe, Emmanuel Edum-Fotwe, Kobina A Bainson, and Paa Kobina Turkson. Factors influencing attendance to immunization sessions for children in a rural district of ghana. *Acta tropica*, 68(3):259–267, 1997.

[24] Shawn T Brown, Benjamin Schreiber, Brigid E Cakouros, Angela R Wateska, Hamadou M Dicko, Diana L Connor, Philippe Jaillard, Mercy Mvundura, Bryan A Norman, Carol Levin, et al. The benefits of redesigning benin's vaccine supply chain. *Vaccine*, 32(32):4097–4103, 2014.

[25] Jean C Buzby, Hodan Farah-Wells, and Jeffrey Hyman. The estimated amount, value, and calories of postharvest food losses at the retail and consumer levels in the united states. *USDA-ERS Economic Information Bulletin*, (121), 2014.

[26] Lap MA Chan, ZJ Max Shen, David Simchi-Levi, and Julie L Swann. Coordination of pricing and inventory decisions: A survey and classification. In *Handbook of quantitative supply chain analysis*, pages 335–392. Springer, 2004.

[27] Horng-Jinh Chang, Jinn-Tsair Teng, Liang-Yuh Ouyang, and Chung-Yuan Dye. Retailers optimal pricing and lot-sizing policies for deteriorating items with partial backlogging. *European Journal of Operational Research*, 168(1):51–64, 2006.

[28] Xiuli Chao, Hong Chen, and Shaohui Zheng. Joint replenishment and pricing decisions in inventory systems with stochastically dependent supply capacity. *European Journal of Operational Research*, 191(1):142–155, 2008.

[29] Xiuli Chao, Xiting Gong, and Shaohui Zheng. Optimal pricing and inventory policies with reliable and random-yield suppliers: characterization and comparison. *Annals of operations Research*, 241(1-2):35–51, 2016.

[30] Richard E Chatwin. Optimal dynamic pricing of perishable products with stochastic demand and a finite set of prices. *European Journal of Operational Research*, 125(1):149–174, 2000.

[31] Jen-Ming Chen and Liang-Tu Chen. Periodic pricing and replenishment policy for continuously decaying inventory with multivariate demand. *Applied Mathematical Modelling*, 31(9):1819–1828, 2007.

[32] Li-Ming Chen and Amar Sapra. Joint inventory and pricing decisions for perishable products with two-period lifetime. *Naval Research Logistics (NRL)*, 60(5):343–366, 2013.

[33] Shen-I Chen. *Modeling the WHO-EPI vaccine supply chain in low and middle income countries.* PhD thesis, University of Pittsburgh, 2012.

[34] Sheng-I Chen, Bryan A Norman, Jayant Rajgopal, Tina M Assi, Bruce Y Lee, and Shawn T Brown. A planning model for the who-epi vaccine distribution network in developing countries. *IIE Transactions*, 46(8):853–865, 2014.

[35] Wen Chen, Qi Feng, and Sridhar Seshadri. Sourcing from suppliers with random yield for price-dependent demand. *Annals of Operations Research*, 208(1):557–579, 2013.

[36] Xin Chen, Zhan Pang, and Limeng Pan. Coordinating inventory control and pricing strategies for perishable products. *Operations Research*, 62(2):284–300, 2014.

[37] Yiwei Chen and Cong Shi. Joint pricing and inventory management with strategic customers. 2017.

[38] Ek Peng Chew, Chulung Lee, and Rujing Liu. Joint inventory allocation and pricing decisions for perishable products. *International Journal of Production Economics*, 120(1):139–150, 2009.

[39] Stephen E Chick, Hamed Mamani, and David Simchi-Levi. Supply chain coordination and influenza vaccination. *Operations Research*, 56(6):1493–1506, 2008.

[40] Prashant Chintapalli. Simultaneous pricing and inventory management of deteriorating perishable products. *Annals of Operations Research*, 229(1):287–301, 2015.

[41] George B Dantzig and Philip Wolfe. Decomposition principle for linear programs. *Operations Research*, 8(1):101–111, 1960.

[42] Institut National de la Statistique du Niger. du recensement, ins rappor, 2012.

[43] Aswin Dhamodharan and Ruben A Proano. Determining the optimal vaccine vial size in developing countries: a monte carlo simulation approach. *Healthcare Management Science*, 15(3):188–196, 2012.

[44] J Donn, Martha Mendoza, and Justin Pritchard. Pharmaceuticals found in drinking water, affecting wildlife, maybe humans. *An AP Investigation: Pharmaceuticals Found in Drinking Water*, 2008.

[45] Paul K Drain, Carib M Nelson, and John S Lloyd. Single-dose versus multi-dose vaccine vials for immunization programmes in developing countries. *Bulletin of the World Health Organization*, 81(10):726–731, 2003.

[46] INS Rapport du Recensement. Institut national de la statistique du niger, 2011.

[47] Sandra Duni Ekşioğlu and Mingzhou Jin. Cross-facility production and transportation planning problem with perishable inventory. In *Computational Science and Its Applications-ICCSA 2006*, pages 708–717. Springer, 2006.

[48] Wedad Elmaghraby and Pınar Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management science*, 49(10):1287–1309, 2003.

[49] Lin Feng, Ya-Lan Chan, and Leopoldo Eduardo Cárdenas-Barrón. Pricing and lot-sizing polices for perishable goods when the demand depends on selling price, displayed stocks, and expiration date. *International Journal of Production Economics*, 185:11–20, 2017.

[50] Qi Feng. Integrating dynamic pricing and replenishment decisions under supply capacity uncertainty. *Management Science*, 56(12):2154–2172, 2010.

[51] Brant E Fries. Optimal ordering policy for a perishable commodity with fixed lifetime. *Operations Research*, 23(1):46–61, 1975.

[52] GAVI. Comprehensive multi-year plan of the national immunization program of bangladesh 2011 2016. Technical report, GAVI, 2011.

[53] Abebe Geletu. Chance constrained optimization-applications, properties and numerical issues. 2012.

[54] Saheed Gidado, Patrick Nguku, Oladayo Biya, Ndadilnasiya Endie Waziri, Abdulaziz Mohammed, Peter Nsubuga, Henry Akpan, Akin Oyemakinde, Abdulsalami Nasidi, Idris Suleman, et al. Determinants of routine immunization coverage in bungudu, zamfara state, northern nigeria, may 2010. *The Pan African medical journal*, 18(Suppl 1), 2014.

[55] Stephane Guichard, Karen Hymbaugh, Brent Burkholder, Serguei Diorditsa, Christine Navarro, Selina Ahmed, and Mohd Mahbubur Rahman. Vaccine wastage in bangladesh. *Vaccine*, 28(3):858–863, 2010.

[56] Dana Gunders. Wasted: How america is losing up to 40 percent of its food from farm to fork to landfill. *Natural Resources Defense Council*, pages 1–26, 2012.

[57] Cong Guo and Xueping Li. A multi-echelon inventory system with supplier selection and order allocation under stochastic demand. *International Journal of Production Economics*, 151:37–47, 2014.

[58] Leila A Haidari, Diana L Connor, Angela R Wateska, Shawn T Brown, Leslie E Mueller, Bryan A Norman, Michelle M Schmitz, Proma Paul, Jayant Rajgopal, Joel S Welling, et al. Augmenting transport versus increasing cold storage to improve vaccine supply chains. *PLoS One*, 8(5):e64303, 2013.

[59] René Haijema, Nico van Dijk, Jan van der Wal, and Cees Smit Sibinga. Blood platelet production with breaks: optimization by sdp and simulation. *International Journal of Production Economics*, 121(2):464–473, 2009.

[60] MM Hasan and IA Karimi. Piecewise linear relaxation of bilinear programs using bivariate partitioning. *AIChE journal*, 56(7):1880–1893, 2010.

[61] René Henrion. Introduction to chance-constrained programming. *Tutorial paper for the Stochastic Programming Community home page*, 2004.

[62] Julia L Higle and Suvrajeet Sen. Stochastic decomposition: An algorithm for two-stage linear programs with recourse. *Mathematics of Operations Research*, 16(3):650–669, 1991.

[63] Julia L Higle and Suvrajeet Sen. *Stochastic decomposition: a statistical method for large scale stochastic linear programming*, volume 1. John Wiley & Sons, Inc., 1999.

[64] Curt Hjorring and John Holt. New optimality cuts for a single-vehicle stochastic routing problem. *Annals of Operations Research*, 86:569–584, 1999.

[65] Dorit S Hochbaum and Arie Segev. Analysis of a flow problem with fixed charges. *Networks*, 19(3):291–312, 1989.

[66] Vernon Ning Hsu. Dynamic economic lot size model with perishable inventory. *Management Science*, 46(8):1159–1169, 2000.

[67] Peng Hu, Stephen Shum, and Man Yu. Joint inventory and markdown management for perishable goods with strategic consumer behavior. *Operations Research*, 64(1):118–134, 2015.

[68] Md Rahedul Islam and Md Zahidul Hassan. Land use changing pattern and challenges for agricultural land: A study on rajshahi district. *Journal of Life and Earth Science*, 6:69–74, 2011.

[69] Dean T Jamison, Joel G Breman, Anthony R Measham, George Alleyne, Mariam Claeson, David B Evans, Prabhat Jha, Anne Mills, and Philip Musgrove. *Disease control priorities in developing countries*. World Bank Publications, 2006.

[70] Larissa Janssen, Thorsten Claus, and Jürgen Sauer. Literature review of deteriorating inventory models by key topics from 2012 to 2015. *International Journal of Production Economics*, 182:86–112, 2016.

[71] Itir Z Karaesmen, Alan Scheller-Wolf, and Borga Deniz. Managing perishable and aging inventories: review and future research directions. In *Planning production and inventories in the extended enterprise*, pages 393–436. Springer, 2011.

[72] Onur Kaya and Sajjad Rahimi Ghahroodi. Inventory control and pricing for perishable products under age and price dependent stochastic demand. *Mathematical Methods of Operations Research*, pages 1–35, 2018.

[73] Padmanaban Kesavan, Russell J Allgor, Edward P Gatzke, and Paul I Barton. Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. *Mathematical Programming*, 100(3):517–535, 2004.

[74] A Krishnamoorthy and TV Varghese. Inventory with disaster. *Optimization*, 35(1):85–93, 1995.

[75] Gilbert Laporte and François V Louveaux. The integer l-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13(3):133–142, 1993.

[76] Amy Hing-Ling Lau and Hon-Shiang Lau. The newsboy problem with price-dependent demand distribution. *IIE transactions*, 20(2):168–175, 1988.

[77] John Lavitt. Nearly $2 billion in prescription drugs wasted every year. Technical report, The fix addiction and recovery, straight up, 2015.

[78] Bruce Y Lee, Tina-Marie Assi, Korngamon Rookkapan, Diana L Connor, Jayant Rajgopal, Vorasith Sornsrivichai, Shawn T Brown, Joel S Welling, Bryan A Norman, Sheng-I Chen, et al. Replacing the measles ten-dose vaccine presentation with the single-dose presentation in thailand. *Vaccine*, 29(21):3811–3817, 2011.

[79] Bruce Y Lee, Diana L Connor, Angela R Wateska, Bryan A Norman, Jayant Rajgopal, Brigid E Cakouros, Sheng-I Chen, Erin G Claypool, Leila A Haidari, Veena Karir, et al. Landscaping the structures of gavi country vaccine supply chains and testing the effects of radical redesign. *Vaccine*, 33(36):4451–4458, 2015.

[80] Bruce Y Lee, Bryan A Norman, Tina-Marie Assi, Sheng-I Chen, Rachel R Bailey, Jayant Rajgopal, Shawn T Brown, Ann E Wiringa, and Donald S Burke. Single versus multi-dose vaccine vials: an economic computational model. *Vaccine*, 28(32):5292–5300, 2010.

[81] Qing Li, Peiwen Yu, and Xiaoli Wu. Managing perishable inventories in retailing: Replenishment, clearance sales, and segregation. *Operations Research*, 64(6):1270–1284, 2016.

[82] Qing Li and Shaohui Zheng. Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Operations Research*, 54(4):696–705, 2006.

[83] Shukai Li, Jianxiong Zhang, and Wansheng Tang. Joint dynamic pricing and inventory control policy for a stochastic inventory system with perishable products. *International Journal of Production Research*, 53(10):2937–2950, 2015.

[84] I-Chen Lin and Yu-Hsiang Hung. On a dual sourcing policy in two-sided disruptive supply chain. In Ershi Qi, Jiang Shen, and Runliang Dou, editors, *Proceedings of the 21st International Conference on Industrial Engineering and Engineering Management 2014*, pages 113–117, Paris, 2015. Atlantis Press.

[85] J. Linderoth, A. Shapiro, and S. Wright. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research*, 142(1):215–241, 2006.

[86] Liming Liu and Zhaotong Lian. (s, s) continuous review models for products with fixed lifetimes. *Operations Research*, 47(1):150–158, 1999.

[87] BENJAMIN P Loevinsohn and MICHAEL E Loevinsohn. Well child clinics and mass vaccination campaigns: an evaluation of strategies for improving the coverage of primary health care in a developing country. *American journal of public health*, 77(11):1407–1411, 1987.

[88] James Luedtke and Shabbir Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008.

[89] Richard T Mahoney and James E Maynard. The introduction of new vaccines into developing countries. *Vaccine*, 17(7-8):646–652, 1999.

[90] Reza Maihami and Isa Nakhai Kamalabadi. Joint pricing and inventory control for non-instantaneous deteriorating items with partial backlogging and time and price dependent demand. *International Journal of Production Economics*, 136(1):116–122, 2012.

[91] Reza Maihami and Behrooz Karimi. Optimizing the pricing and replenishment policy for non-instantaneous deteriorating items with stochastic demand and promotional efforts. *Computers & Operations Research*, 51:302–312, 2014.

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

[93] W.K. Mak, D. P. Morton, and K. Wood. Monte carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, 24(1):47 – 56, 1999.

[94] Garth P McCormick. Computability of global solutions to factorable nonconvex programs: Part iconvex underestimating problems. *Mathematical programming*, 10(1):147–175, 1976.

[95] Edwin S Mills. Uncertainty and price theory. *The Quarterly Journal of Economics*, 73(1):116–130, 1959.

[96] Maryam H Mofrad, Gian-Gabriel P Garcia, Lisa M Maillart, Bryan A Norman, and Jayant Rajgopal. Customizing immunization clinic operations to minimize open vial waste. *Socio-Economic Planning Sciences*, 54:1–17, 2016.

[97] Maryam H Mofrad, Lisa M Maillart, Bryan A Norman, and Jayant Rajgopal. Dynamically optimizing the administration of vaccines from multi-dose vials. *IIE Transactions*, 46(7):623–635, 2014.

[98] Dipana Jyoti Mohanty, Ravi Shankar Kumar, and A Goswami. An improved inventory model with random review period and temporary price discount for deteriorating items. *International Journal of System Assurance Engineering and Management*, 7(1):62–72, 2016.

[99] Bola Murele, Rui Vaz, Alex Gasasira, Pascal Mkanda, Tesfaye Erbeto, and Joseph Okeibunor. Vaccine perception among acceptors and non-acceptors in sokoto state, nigeria. *Vaccine*, 32(26):3323–3327, 2014.

[100] Steven Nahmias. Optimal ordering policies for perishable inventoryii. *Operations Research*, 23(4):735–749, 1975.

[101] Steven Nahmias. Perishable inventory theory: A review. *Operations research*, 30(4):680–708, 1982.

[102] Steven Nahmias. *Perishable Inventory Systems*. Springer, 2011.

[103] Steven Nahmias and William P Pierskalla. Optimal ordering policies for a product that perishes in two periods subject to stochastic demand. *Naval Research Logistics Quarterly*, 20(2):207–229, 1973.

[104] Steven Nahmias and Shan Shan Wang. A heuristic lot size reorder point model for decaying inventories. *Management Science*, 25(1):90–97, 1979.

[105] NIPORT. Bangladesh demographic and health survey 2011. Technical report, NIPORT National Institute of Population Research and Training, Associates for Community and Population Research (ACPR), and ICF International., 2011.

[106] World Health Organization. Global alliance for vaccines and immunization. geneva: World health organization. 2001.

[107] World Health Organization et al. Expanded programme on immunization. 1983.

[108] World Health Organization et al. Monitoring vaccine wastage at country level: guidelines for programme managers. 2005.

[109] World Health Organization et al. Global health observatory (gho) data. 2016. *Child mortality and causes of death. WHO, Geneva*, 2016.

[110] Sachiko Ozawa, Chizoba Wonodi, Olufemi Babalola, Tukur Ismail, and John Bridges. Using best-worst scaling to rank factors affecting vaccination demand in northern nigeria. *Vaccine*, 35(47):6429–6437, 2017.

[111] BK Pagnoncelli, Shapiro Ahmed, and Alexander Shapiro. Sample average approximation method for chance constrained programming: theory and applications. *Journal of optimization theory and applications*, 142(2):399–416, 2009.

[112] Julia Pahl and Stefan Voß. Integrating deterioration and lifetime constraints in production and supply chain planning: A survey. *European Journal of Operational Research*, 238(3):654–674, 2014.

[113] Gokce Palak. *Optimization models for cost efficient and environmentally friendly supply chain management*. Mississippi State University, 2013.

[114] Gökçe Palak, Sandra Duni Ekşioğlu, and Joseph Geunes. Analyzing the impacts of carbon regulatory mechanisms on supplier and mode selection decisions: An application to a biofuel supply chain. *International Journal of Production Economics*, 154:198–216, 2014.

[115] Divya Parmar, Elaine M Baruwa, Patrick Zuber, and Souleymane Kone. Impact of wastage on single and multi-dose vaccine vials: Implications for introducing pneumococcal vaccines in developing countries. *Human Vaccines*, 6(3):270–278, 2010.

[116] Karin GJ Pauls-Worm. Inventory control for a perishable product with non-stationary demand and service level constraints. Technical report, 2013.

[117] Stanley A Plotkin. Vaccines: past, present and future. *Nature medicine*, 11(4s):S5, 2005.

[118] E. L. Porteus. *Foundations of Stochastic Inventory Theory*. Stanford Business Books, Stanford, 2002.

[119] Fred Raafat. Survey of literature on continuously deteriorating inventory models. *Journal of the Operational Research society*, pages 27–37, 1991.

[120] Ilma Nurul Rachmania and Mursyid Hasan Basri. Pharmaceutical inventory management issues in hospital supply chains. *Management*, 3(1):1–5, 2013.

[121] Mosiur Rahman and Sarker Obaida-Nasrin. Factors affecting acceptance of complete immunization coverage of children under five years in rural bangladesh. *Salud pública de méxico*, 52(2):134–140, 2010.

[122] Walter Rei, Jean-François Cordeau, Michel Gendreau, and Patrick Soriano. Accelerating benders decomposition by local branching. *INFORMS Journal on Computing*, 21(2):333–345, 2009.

[123] R Tyrrell Rockafellar and Roger J-B Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147, 1991.

[124] Arvind Sainathan. Pricing and replenishment of competing perishable product variants under dynamic demand substitution. *Production and Operations Management*, 22(5):1157–1181, 2013.

[125] Suvrajeet Sen. Stochastic mixed-integer programming algorithms: Beyond benders' decomposition. *Wiley Encyclopedia of Operations Research and Management Science*, 2011.

[126] Jhih-Shyang Shih and H Christopher Frey. Coal blending optimization under uncertainty. *European Journal of Operational Research*, 83(3):452–465, 1995.

[127] Patrick Lydon Marta Gacic-Dobo Rudolf Eggers Laura Conklin Subaiya, Laure Dumolard. Global routine vaccination coverage, 2014. Technical report, Centers for Disease Control and Prevention, 11 2015.

[128] Chirag Surti, Elkafi Hassini, and Prakash Abad. Pricing and inventory decisions with uncertain supply and stochastic demand. *Asia-Pacific Journal of Operational Research*, 30(06):1350030, 2013.

[129] Mohit Tawarmalani and Nikolaos V Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Mathematical programming*, 99(3):563–591, 2004.

[130] Michael Tsiros and Carrie M Heilman. The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories. *Journal of marketing*, 69(2):114–129, 2005.

[131] Huseyin Tunc, Onur A Kilic, S Armagan Tarim, and Burak Eksioglu. A reformulation for the stochastic lot sizing problem with service-level constraints. *Operations Research Letters*, 42(2):161–165, 2014.

[132] Wim van Ackooij, Riadh Zorgati, René Henrion, and Andris Möller. Chance constrained programming and its applications to energy management. In *Stochastic Optimization-Seeing the Optimal for the Uncertain*. InTech, 2011.

[133] Karel H Van Donselaar and Rob ACM Broekmeulen. Approximations for the relative outdating of perishable products by combining stochastic modeling, simulation and regression modeling. *International Journal of Production Economics*, 140(2):660–669, 2012.

[134] Richard M Van Slyke and Roger Wets. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17(4):638–663, 1969.

[135] Gideon Johannes Jacobus Van Zyl. *Inventory control for perishable commodities*. PhD thesis, University of North Carolina at Chapel Hill, 1963.

[136] Jelena V Vlajic, Jack GAJ Van der Vorst, and René Haijema. A framework for designing robust food supply chains. *International Journal of Production Economics*, 137(1):176–189, 2012.

[137] Jean-Paul Watson and David L Woodruff. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science*, 8(4):355–370, 2011.

[138] Björn Weber, Sarah Herrlein, and Greg Hodge. The challenge of food waste. *Planet Retail*, 2011.

[139] Hui-Ming Wee. Joint pricing and replenishment policy for deteriorating inventory with declining market. *International journal of production economics*, 40(2-3):163–171, 1995.

[140] Howard J Weiss. Optimal ordering policies for continuous review perishable inventory models. *Operations Research*, 28(2):365–374, 1980.

[141] Thomson M Whitin. Inventory control and price theory. *Management science*, 2(1):61–68, 1955.

[142] WHO. The expanded programme on immunization. `http://www.who.int/immunization/programmes_systems/supply_chain/benefits_of_immunization/en/`, 1974. Online; accessed May 2018.

[143] WHO. Immunization coverage reaches 84%, still short of 90% goal . `http://www.who.int/immunization/newsroom/press/immunization_coverage_july2014/en/`, 2014. Online; accessed May 2018.

[144] WHO. WHO Mortality Database. `http://www.who.int/healthinfo/mortality_data/en`, 2016. Online; accessed May 2018.

[145] World Health Organization WHO. Who policy statement: Multi-dose vial policy (mdvp)-revision 2014. geneva, 2014. Technical report, WHO/IVB/14.07, 2014.

[146] Gino K Yang, Robert Lin, Jennifer Lin, Kuo-Chen Hung, Peter Chu, and Wayne Chouhuang. Note on inventory models with weibull distribution deterioration. *Production Planning and Control*, 22(4):437–444, 2011.

[147] Wanfei Yang, Monika Parisi, Betsy J Lahue, Md Jasim Uddin, and David Bishai. The budget impact of controlling wastage with smaller vials: A data driven model of session sizes in bangladesh, india (uttar pradesh), mozambique, and uganda. *Vaccine*, 32(49):6643–6648, 2014.

[148] Yang C Yuan. Multiple imputation for missing data: Concepts and new development (version 9.0). *SAS Institute Inc, Rockville, MD*, 49:1–11, 2010.

[149] P. Zipkin. *Foundations of Inventory Management.* McGraw-Hill, Columbus, 2000.