ABSTRACT

Title of Document: MATHEMATICAL MODEL AND FRAMEWORK FOR MULTI-PHASE PROJECT OPTIMIZATION

Ali Shafahi, Ph.D., 2016

Directed By: Prof. Ali Haghani, Department of Civil and Environmental Engineering

This research aims to assist investors of "real" tangible assets such as construction projects in making an optimal portfolio of phased and regular projects which will yield the best financial outcome calculated in terms of discounted cash flow of future anticipated revenues and costs. We use optimization techniques to find the optimal timing and phasing of a single project that has the potential of being decomposed into smaller sequential phases.

Existence of uncertainties is inevitable especially in cases in which we are planning for long durations. In the presence of these uncertainties, full upfront commitment to large projects may jeopardize the rationality of investments and cause substantial economic risks. Breaking a big project into smaller stages (phases) and implementing a staged development is a potential mechanism to hedge the risk. Under this approach, by adding managerial flexibilities, we may choose to abandon a project at any time once the uncertain outcomes are not favorable. In addition to the benefits resulting from hedging

unfavorable risks, phasing a project can transform a financially infeasible project into a feasible one due to less load on capital budgets during each time.

Once some phases of a project are delayed and planned to be implemented sequentially, it is important to prepare the infrastructure required for their future development. Initially, we present a Mixed Integer Programming (MIP) model for the deterministic case with no uncertainties that considers interrelationships between phases of projects such as scheduling and costs (economy of scales) in addition to the initial infrastructural investment required for implementation of future phases. Pairing possible phases of a project and doing them in parallel is beneficial due to positive synergies between phases but on the downside requires larger capital investments. Unavailability of enough budgets to fully develop a profitable project will cause the investment to be carried out in different phases e.g. during times when the required capital for developing the next phase (or group of phases) is available.

After, presenting the model for the deterministic case, we present a scenario-based multi-stage MIP model for the stochastic case. The source of uncertainty considered is future demand that is modeled using a trinomial lattice. We then present two methods for solving the stochastic problem and finding the value of the here and now decision variable (the size of the infrastructure/foundation). Finding the value of the here and now decision variable for all scenarios using a novel technique that does not require solving all the scenarios is the first method. The second method combines simulation and optimization to find good solutions for the here and now decision variable.

Lastly, we present a MIP for the deterministic multi-project case. In this setting, projects could have multiple phases. The MIP will help the managers in making the project selection and scheduling decision simultaneously. It will also assist the managers in making appropriate decisions for the size of the infrastructure and the implementation schedule of the phases of each project. To solve this complex model, we present a pre-processing step that helps reduce the size of the problem and a heuristic that finds good solutions very fast.

MATHEMATICAL MODEL AND FRAMEWORK FOR MULTI-PHASE PROJECT OPTIMIZATION

By

Ali Shafahi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Professor Ali Haghani, Chair
Professor Michael Fu, Dean’s representative
Professor Paul Schonfeld
Associate Professor Lei Zhang
Associate Professor Qingbin Cui

Preface

This dissertation is composed as a requirement for the degree of Doctorate of Philosophy in Civil and Environmental Engineering.

Dedication

To my great Family. My mother, father, brother, sister in-law and my lovely nephew, Sepehr.

Acknowledgements

First and foremost, I would like to greatly thank my advisor, Prof. Ali Haghani for his great guidance and the advices he has given me throughout my graduate studies at UMD. He has been a wonderful source of ideas and motivation. My development as a scholar is due to his helpful comments and the opportunities provided by him. I have had the honor to learn from him on how to prepare material for courses, lecture, and give presentations. He has involved me in many projects that has boosted my confidence and helped me have a better understanding of practical research. I look forward to our continued collaboration. In addition to all the advice in the research, he has also been helpful in making me feel comfortable during my graduate life. He has always been available for advices on all topics. I can see the imprints of the lessons I have learned from Prof. Haghani in all aspects of my life.

During my graduate studies, I have had the pleasure of being a teaching assistant for many faculty in addition to my advisor. I would like to thank Prof. Paul Schonfeld, Prof. Lei Zhang, Prof. Cinzia Cirillo, Prof. Hakob Avetisyan, and Dr. Rafel Olarte for the great assistance and opportunity provided to me to work with them. Being a TA, I have also had great experience with working with the students. There were many students that I have had pleasure interacting with them weekly. I hope to see all of them excel in their professional life.

I would also greatly thank all my friends for their support. The friends that I have made during these times are really like my second family. Thank you Vahid Liaghat, Mohammad Reza Khani, Kiana Roshan Zamir, Faezeh Dorri, Hamid Mahini, Hossein Esfandiari, Saeed Seddighin, Sina Dehghani, Milad Gholami, Mohammad Alizadeh, Majeed Kazemitabaar, Melika Abolhassani, Soheil Ehsani, Mahyar Najibi, Ehsan Taheri, Omid Ghaffaripour, Mersedeh Tariverdi and Saba Ahmadi.

I am also grateful to all of the members of my dissertation committee and Prof. MohammadTaghi Hajiaghayi for their helpful comments and suggestions.

Lastly, but definitely most importantly, I would like to thank my parents and my brother. I can’t really find words grand enough to say thank you…

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Investments in real tangible assets such as projects often are assumed to be somewhat irreversible. An irreversible investment is an investment that its cost cannot be recovered once it is installed (Eberly, 2008). Irreversibility of investments increase the importance of planning. Projects are undertaken because of a demand. Therefore a project's importance and profitability is influenced directly by the demand. There are many methods for executing a project. This research focuses on phased implementation of projects.

Phased implementation of projects relies upon the ability of breaking a huge project into smaller pieces. The smaller pieces of the project can be implemented in a stage wise manner. In this research the terms "phases" and "stages" have been used interchangeably. The phases of projects considered in this research have some attributes which are listed below:

- Phases have precedence and successor relationships: The precedence relationship is sequential. Meaning that a successor phase, j, can start no sooner than its predecessor phase $j-1$. If each phase is thought as an activity, this relationship is analogous to the start to start dependency among activity j (phase j) and its prior activity ($j-1$). Figure 1 illustrates this dependency relationship among phases.
- Phases produce profits: Each phase by itself can potentially produce a positive utility (revenue) for the firm implementing that phase. The additional revenue

of a phase is assumed to be independent from the previous phases because from the previous assumption, the implementation of that phase is only possible if its prior phases were all started (implemented). In other words, the potential revenue at any time is independent of the number of phases implemented until that time.

- Phases can be implemented one by one and in sequences or some phases can be grouped together and done in parallel.

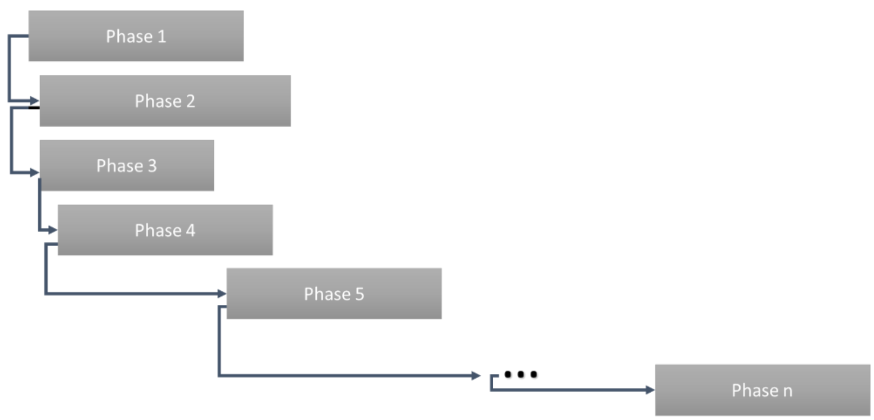

Figure 1 Start to Start Relationship between possible phases of a project

1.1 Examples of projects and potential sequential phases

Overall many projects can be broken down into smaller phases where some of the phases can be done sequentially. Sequence is defined as a particular order in which related things follow each other. Based on this definition a sequence can be based on time. The potential sequential phases do not necessarily have to be done sequentially but can be done sequentially based on management's preference and decision. For example expansion of an infrastructure can be done sequentially. Let this infrastructure be a highway. The investors can decide to expand a two lane highway into a 4 lane highway either right-away or by a sequence of first expanding it into a 3 lane highway and then doing another expansion and expanding it into a 4 lane highway. In this case the potential phases would be: 1) adding one lane to a two lane highway and making it a 3 lane highway, and 2) adding another lane to a 3 lane highway and making it a 4 lane highway. Note that phases 1 and 2 can be done either sequentially or in parallel. If done in parallel, we expand the highway from 2 lanes to 4 lanes at once.

Another example of an expansion is expanding a 2 story building into a 5 story building. The potential phases in this case are the levels: 1) construct level 3, 2) construct level 4, and 3) build level 5. Note that similar to the highway expansion example, the expansion of the building can be done in one shot and we can add the three stories at once. Or we can first finish phase 1 and then do phase 2 and afterwards do phase 3. The first method is a purely parallel expansion and the second method is a purely sequential expansion. In this example, there are many other alternatives for completing the 5 story building. These other “options” are not fully parallel or fully sequential. The possible expansion methods based on the phases are summarized in Figure 2. Due to ordering and sequential relationships it is not possible to do phases 2 and 3 in parallel and after that finish phase 1.

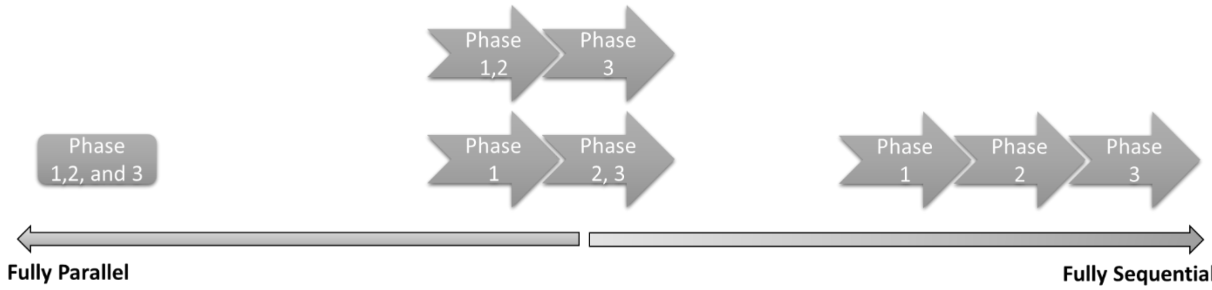

Figure 2 Spectrum of possible methods for undergoing a project with three phases

Construction of a residential complex is another example. Assuming that the residential complex has *b* apartment buildings, each building can be thought of as a phase.

In general when we have *n* potentially sequential phases, we would have $\sum_{i=0}^{n-1}\binom{n-1}{i}$ possible options to complete the project. For example, if $n=8$, the number of possible options to complete all phases will be 128.

$$\binom{7}{0}+\binom{7}{1}+\binom{7}{2}+\binom{7}{3}+\binom{7}{4}+\binom{7}{5}+\binom{7}{6}+\binom{7}{7}=128$$

1.2 Benefits and disadvantages of phasing projects

Phasing projects can be to the benefit of management if planned wisely. Some potential benefits are mentioned below:

a) Completion of costly and large projects by breaking them down into smaller pieces with less costs and performing the smaller pieces at different times upon availability of funds;

b) Breaking large projects into smaller manageable pieces;

c) Reducing operating costs required early due to having a smaller portion of the entire project completed at sooner times;

d) Less initial capital costs open the door for investing remaining funds in other projects;

e) Increasing resiliency and allowing for more learning by allowing partial completion at different times; and,

f) Preventing underutilization of a project by hedging risks.

g) Financing later phases with revenues from early phases.

Some of the disadvantages of phasing a project and not performing it all together (performing the phases in parallel) is listed in the following:

a) Loss of economy of scale;

b) Loss of revenue due to loss of unsatisfied demand;

To eliminate budget limitations and infeasibilities caused by budget limitation at a certain time, in addition to phasing a project, we can delay and postpone the project until we have enough budget to undertake the project.

Some of the advantages and disadvantages are summarized Table 1.

Table 1 Some advantages and disadvantages of phasing (staging) a project

Advantages of doing a project in one stage	Advantage of phasing and performing it in several stages
• Economy of scale • Gaining more revenue in case of a profitable and favorable project • Less unsatisfied demand (in case of public projects)	• Reduction in uncertainty (more resilient solutions). Learning! • Reduction in operating costs for the initial years (maintenance) • Less initial capital costs open the possibility for investment in other projects • Preventing under utilization

1.3 Scope of dissertation

In this dissertation we intend to tackle the problem of optimizing sequential multi-phase project investments. We have developed models that can assist decision makers in making optimal decisions with respect to potential multi-phase projects. For a single project, we determine what phases should be done in parallel and what phases should be done sequentially. Then we apply the multi-phase framework to a portfolio of projects and find the optimal portfolio of multi-phase projects and single phase projects.

In the remaining of the dissertation first we review the literature related to phased investment and sequential investments. Then we present a mathematical model for deterministic cases of multi-phase projects. Then in the next section we consider uncertainties and present a stochastic optimization framework for multi-phase projects that require initial investments. A deterministic multi-project MIP formulation is then presented. Finally, a summary is provided and conclusions are drawn and some directions for future research is provided.

Chapter 2: Literature Review

The scope of this dissertation covers problems related to phased investment and sequential investments. Phased investment framework has been applied to many practices and industries. Information Technology (IT) projects (Miller, et al., 2004), Flexible Manufacturing Systems (FMS) (Lamar & Lee, 1999), and industry plant expansion (Lieberman, 1987) are among the applications of phased investment and implementation of a project which have been studied in the literature. Other areas of studies are, R&D projects (Herath & Park, 2002), market entry decisions (Pennings & Lint, 2000), residential development (Ott, et al., 2012), commercial building energy retrofits (Lee, et al., 2014), distribution network expansion planning (Carvalho, et al., 1998), container port expansions (Dekker & Verhaeghe, 2008), parking garage construction (De Neufville, et al., 2006), and highway development and expansion (Zhao, et al., 2004).

2.1 Real Options for Evaluating Phased Projects

Many of the literature related to phased investment focuses on evaluation of such type of investments. The general objective of these studies is to evaluate real world cases of phased investments in which the investees have majorly decided to invest and justify this decision. Real options is the most common method used for investment evaluation. (Miller, et al., 2004) use real options to evaluate the Korean information technology infrastructure. They utilize the deferral real options framework to evaluate the investment decision for investing in the information technology infrastructure of South

Korea. They conduct a brief literature review on research that apply real options to IT investment decisions.

They model the three phased investment model using two methods: growth options, and compound options. For the growth option, they use the Black and Scholes equation and model the IT investment option as a standard European call option assuming that investing in phases 1 and 2 will buy us the option to invest in phase 3. They also model the investment as compound option such that investing in phase 1 will give us the option to invest in phase 2 and investing in phase 2 will give us the option of investing in phase 3. The compound option’s value is calculated using Geske’s compound option model (Geske, 1979). They state that estimating the volatility parameter is generally done using four methods: a) Twin Security Argument; b) Implied Volatility; c) Modified Scenario Analysis; and d) Monte Carlo Simulation.

Herath and Park study a multi-stage project setting in which each investment opportunity derives revenues from different markets but shares common technological resources (Herath & Park, 2002). They use the binomial lattice framework to model the multi-stage investment as a compound real option when uncorrelated underlying variables exist. They use Monte Carlo simulation to estimate the volatility parameter. They state that the famous Black and Scholes option pricing model should not be used for valuing sequential compound options.

(Herath & Park, 2002) assume that the time to maturity of options are known beforehand and also the time which the investments should take place are given. The decision which should be made is whether to invest at that time in the next option or not. To show the performance of their model, they present a hypothetical example of an R&D investment in a manufacturing environment. They assume that there are two sources of uncertainty. One of them being technological and the other being demand for the new product. They evaluate the project investment opportunity both using real options and DCF analysis and illustrate how DCF fails to capture the flexibilities in future investments and approaches the investment as an all or nothing investment.

(Pennings & Lint, 2000) use the Black and Scholes option pricing formula to price the call and put options in phased market introduction. They find optimal timing and optimal rollout area for a phased market introduction of a new product. They solve the modified Black-Scholes equations to find the optimal solutions. Their model assumes that rollout area is continuous and a firm can decide to enter the market for any portion of the entire area. In addition, their model inherits the assumptions and limitation of the Black and Scholes model because being built on the base of that model.

(Ott, et al., 2012) consider economy of scale in construction in their research and illustrate how the real options framework can be used in estimating the impacts of different economic variables such as economy of scale and inventory costs on optimal phasing and inventory decisions. They state that it is to the utmost interest of residential developers to sell the developed unit as soon as possible to possibly decrease the cost of inventory. However, since the developers actions might affect the prices, they might decide to develop in smaller phases and therefore increase the price per unit. They assume that each developer has the potential to build up to a certain number of units and does not need to acquire any land as it is assumed that they have the land available for the development.

(De Neufville, et al., 2006) illustrate how practitioners can simply perform real option analysis on flexible projects using spreadsheets. They consider three cases: one deterministic case, one case which allows for uncertainty, and one case which flexibility in design is taken into consideration. The last case is valued using real options. They propose a risk preference based decision making method. They use binomial lattice tree to model the only source of uncertainty they considered, (parking) demand. The decision makers risk preference is measured by the certainty equivalent (CE) of a random wealth variable.

2.2 Mathematical Models for Phased Investments and Expansions

Even-though the majority of the work done related to phased investment is based on the real options framework, some of the existing literature has modelled phased investments using mathematical optimization models.

(Lim & Kim, 1998) present a Mixed Integer Programming (MIP) model for gradually replacing conventional dedicated machines with Flexible Manufacturing Modules (FMM) under budget restrictions. The objective of their mathematical model is to minimize the discounted costs of acquisition and operation of FMMs and the operational costs of the conventional machines. Under the phased implementation assumption, FMMs are not acquired all at once and are usually acquired at different years.

(Carvalho, et al., 1998) formulate the problem of finding optimal decisions for the distribution network expansion planning under uncertainties. They model the uncertainties using a set of scenarios.

(Zhao, et al., 2004) develop a multistage stochastic model which assists in making highway development, operation, expansion, and rehabilitation decisions. Three sources of uncertainties and their interdependencies are considered. Namely, traffic demand, land price, and highway deterioration. The solution algorithm used is a combination of Monte-Carlo simulation and least-squares regression. They consider traffic demand to follow a wiener process and use Markov processes to model infrastructure deterioration. They also consider rehabilitation and land acquisition in their research.

(Lieberman, 1987) discusses the assumptions and general findings of two models which are developed for finding industrial plant sizes and capacity expansion based on economy of scale and demand growth parameters.

Some other research which falls close to phased investment is related to a problem called the capacity planning problem. In this problem, we are interested in finding the optimal timing and size to increase capacity such that future demand is met. Generally, the objective is to satisfy all future demand and deficits in demand are not allowed.

It is worth noting that if we relax the assumption of no unsatisfied demand, the phased investment problem is very similar to the capacity planning problem. Note that if we assume the possible increases in capacities to be discrete and each capacity increase to be implemented in each phase, the phased investment model can be used to model capacity planning problems. Flexible Manufacturing Systems (FMS) is an area in which many optimization models have been developed to find optimal investments to meet demands (Fine & Freund, 1990).

2.3 Sequential Investments

Another area of research which is close to phased investment is sequential investment. If we assume that each investment in a sequential investment is a phase, the two are more or less the same.

(Bar-Ilan & Strange, 1998) consider sequential investments in two phases which its phases take time to complete. The benefits are only assumed to be gathered after the entire investment is finished. They assume that the price of the output follow a geometric Brownian motion. After the first stage, the investors have the option to abandon the project or to pursue by investing in the second stage. They use dynamic programming for evaluating the investments and model it using recursive equations. They do not consider any budget limitations.

(Baldwin, 1982) looks into the case that investment opportunities arrive sequentially thus should be reviewed sequentially. She models firms' investment decisions as Markov reward processes. Another research which assumes that investment opportunities arrive in sequences is (Prastacos, 1983).

(Gupt & Rosenhead, 1968) state that for long-range investment plans which last for number of years one possible way of guarding against the danger of committing to a decision based on current knowledge which might turn out to be a bad decision is to keep as many options as open as possible.

(Rocha, et al., 2007) compare the simultaneous and sequential investment strategies in real estate development using real options. They conclude that sequential investment can help reduce the risks.

The most common assumption among the majority of studies is that we have the opportunity and infrastructure of investing in as many phases as we want. However in reality, this is not always the case. This might be the case for some cases of lane expansions performed by the government who is the owner of the land required for expansion. Or the case that no infrastructure is needed for future phases. These are very rare situations in which the infrastructure was acquired and prepared in the past or no infrastructure is needed. However, in a majority of cases the investors have to invest in purchasing and preparing the required infrastructure for future phases at the time they decide to implement a phased investment plan for a project. In (De Neufville, et al., 2006) since the number of levels of the parking garage was assumed to be flexible, the columns were assumed to be "strong columns" and not specifically optimized and designed for a certain number of levels. This assumption which is very simplistic was done since the proposed method for valuing the real options was intended to be very simple to implement. However, having the size of the column following an either or paradigm might significantly increase the cost and difficulty of construction. In addition, due to simplicity, they have not taken into consideration the economies of scale and budget constraints which make the proposed framework further away from real world applications.

The work done by (Zhao & Tseng, 2003) provides a model that addresses the limitations of the aforementioned research. It allows for constructing foundations which can support different levels and also allows the expansion of different levels at different times. The case study they use is the construction of a multi-level parking garage which has fixed and variable costs. They present three “models”. The differences in the models are in how they model demand. In the first model, they assume that demand is constant over time. In their second model, demand varies over time but is deterministic. The third model assumes that the demand is uncertain and a trinomial lattice framework is used for modelling different outcomes of the demand. Dynamic Programming (DP) is used for finding optimal expansion decisions. However, they do not consider synergies between different levels and any benefits for constructing different levels in parallel.

Another limitation which many of the previous research including (Zhao & Tseng, 2003) have is not considering budget limitations at different times. In addition they have not considered time value of money in terms of having the ability to invest the remaining capital in another investment. Considering these and using real options framework is very complicated. However if these limitations are neglected, the real options framework will only be useful to a small audience of investors.

This dissertation aims to fill the gaps and overcome the limitations of existing research. In this dissertation we provide Mixed Integer Programming (MIP) models that can be used for finding optimal phasing and expansion decisions. The models consider synergies among phases and find an optimal strategy for performing some phases in parallel and some other in sequence subjected to budget constraints.

The more detailed contributions of each chapter is presented within the chapters themselves. Nonetheless, some of the major contributions of our research are:

- Developing a Mixed Integer Programming (MIP) mathematical model which can provide optimal decisions in terms of phasing projects and timing of execution,
- Accounting for infrastructural preparations and requirements for implementation of future phases,
- Considering budget constraints and limitations at different times,
- Accompanying synergies among phases such as cost efficiencies,
- Considering uncertainties using the binomial lattice framework of real options analysis,
- Presenting a method for finding the optimal here and now decision variable values for scenario based problems that the scenarios are generated using the lattice framework,
- Solving the multi-stage stochastic optimization problem using simulation,
- Performing portfolio optimization and selecting the optimal mix of phased projects while considering the interdependencies among phases, and
- Presenting a pre-processing step and a heuristic to deal with large cases of multi-project phased investment problems.

Chapter 3: Deterministic Multi-phase Project Optimization

Under certainty assumption, risk is no longer a concern. Therefore the main driving factor for phasing a project and sequentially carrying the phases could be budgetary constraints or better financially viable options in which the firm can invest its remaining capital or preventing underutilization of the entire project at early stages when the demand for the project is limited. As a result, phases of a project might be carried in a sequential order to save on the funds required at sooner times.

In this chapter of the dissertation we assist investors in making three decisions. The first decision is with regard to the quantity and size of investment. The second decision is with regard to pairing of phases and what phases should be done in parallel and what phases should be done sequentially. The last decision is about the optimal timing for investing in each sequential phase or set of parallel phases which are supposed to be done together.

3.1 Modeling Assumptions

The modeling assumptions used in this research are with regard to financials and different interdependencies among project phases.

3.1.1 Financial Assumptions

The financial assumptions of this study are about budget limitations and financial opportunities (time value of money). We assume that the firm has a limited capital (budget) at time zero. At future times the amount of capital they have depends on the amount they had at its previous time step minus the costs incurred at the previous time steps plus whatever is the positive income during the previous time and any additional infused capital from outside. An assumption regarding timing of the costs and profits is that the costs are all collected at the beginning of a time step but the positive income cash flow is rewarded at the end of each time step. To accommodate opportunity losses and time value of money we include a risk free interest rate which is used to update the available budget at future times based on the remaining budget at its previous time step. In addition we include inflation in all types of costs.

3.1.2 Different Types of Costs

Two type of costs are considered in this model. Initial fixed costs which are representative of the required cost for constructing and installing phases. The other cost is operating/maintenance costs (OM). OM costs are ongoing costs and depend on how many of the phases are implemented during each time. In general, the more number of phases implemented, the more the OM costs are likely to be.

3.1.3 Profits from Phases

We assume that each phase potentially generates a profit (revenue) depending on the time it is implemented. These profits depend on the demand for the project. The profit of each phase is expressed in terms of the surplus profit brought by implementing that phase and increasing capacity (supply). Consider an arbitrary supply demand curve such as what is depicted in Figure 3. If we increase the supply capacity (shift out supply) by adding a new phase, the additional revenue which is yielded by serving more demands is the additional profit from that phase. Figure 3 illustrates how the served demand quantity will increase. In this Supply-Demand curve, the hashed area is the profit.

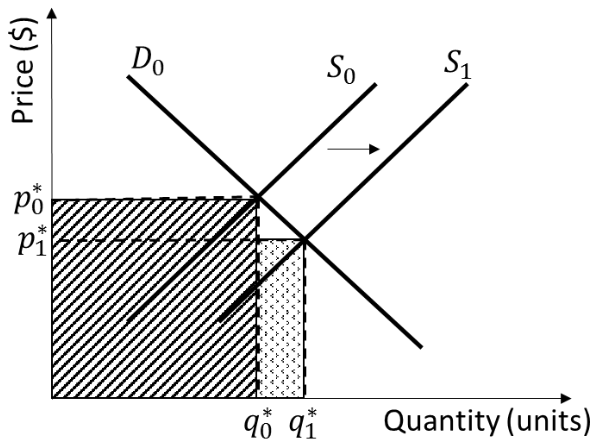

Figure 3 Supply and Demand analysis in case of increasing supply

Note that in Figure 3 p_0^* is the equilibrium price before adding a phase and increasing capacity. The total revenue prior to adding a phase is therefore captured as (3-1):

$$S_0 = p_0^* \times q_0^* \quad \rightarrow \quad R_0 = p_0^* \times q_0^* \tag{3-1}$$

where q_0^* is the equilibrium quantity prior to adding the phase.

After adding the phase, the equilibrium price and quantity are subject to change depending on the shape of the supply-demand curve. The additional revenue however is captured as follows (3-2):

$$R_1 = p_1^* \times q_1^* - p_0^* \times q_0^* \tag{3-2}$$

where p_1^* and q_1^* are the equilibrium price and quantity after adding the phase, respectively.

For the cases of infrastructure developments, the price does not usually fluctuate a lot depending on the demand. For example if the price of a parking space is set to *$p* at a year, if we increase the number of parking spaces (until a certain threshold) the price will not change and therefore can be considered constant. Here, the supply-demand analysis can be done under the assumption of existence of a price ceiling. Figure 4 illustrates the shifting of supply in the presence of a price ceiling and under the assumption of *ceteris paribus*. The hashed area, illustrates the additional revenue brought from implementing the phase and increasing the supply capacity.

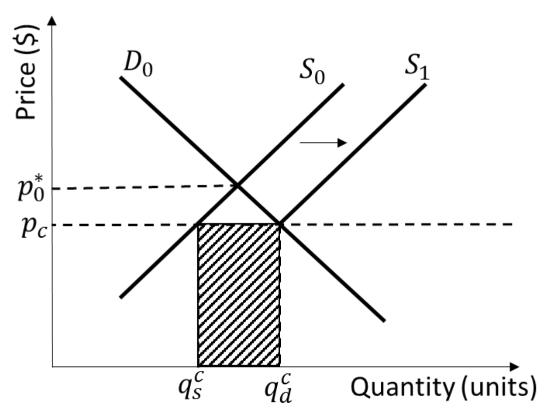

Figure 4 Supply and Demand analysis in presence of price ceiling and increasing supply

In the presence of a price ceiling, by adding more supply (implementing a new phase), the new price will remain the same and it will not be equal to the new equilibrium price.

Therefore, the increase in revenue is solely proportional to the increase in demand served. The additional revenue from adding a phase is therefore as follows:

$$R_1 = p_c \times (q_d^c - q_s^c) \tag{3-3}$$

3.1.4 Resource and Budget Assumptions

We assume that during each time, *t*, the investing firm has a certain budget available to itself and the operating and investment cost at that time cannot exceed this budget. The budget at each time by itself is a function of some variables of the previous time frame. These variables are:

- Budget of the previous time frame,
- Money invested for development of phases in the previous time frame,
- Money consumed for head over costs (operation and maintenance costs) in the previous time frame,
- Revenue (Profit) received from previous investments in the previous time frame,
- Risk-free interest rate, and
- Inflation.

3.1.5 Time-related Relationships between Phases

It is assumed that the phases have sequential relationships such that we cannot invest in a phase, *p*, prior to investing in all its previous phases. However, we can decide to invest in a group of phases together in parallel due to many reasons such as economy of scale.

3.1.6 Initial Investment and Phase-Infrastructure Investment

Investments in phases at future times depend on the initial investment infrastructural investment decision at time zero. We can only invest in a phase at any future time if we have invested in acquiring the required infrastructure for that future phase. Figure 5 depicts two different cases of a two phase investment. The left picture is different cost profiles and paths if we only invest in the infrastructure of one phase. In this case the dotted line represents the cost profile if we do not invest in the phase and abandon the project. The solid line represents the cost profile if we invest in the phase.

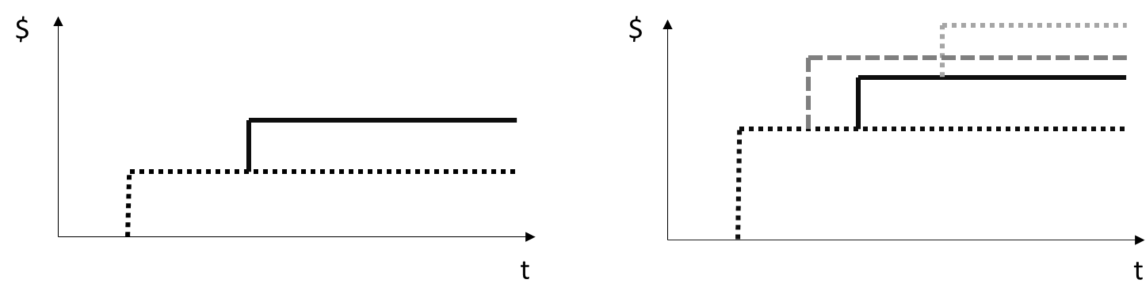

Figure 5 Different Cost Profiles based on potential number of phases. The left figure the potential is only 1 phase. The figure at right can have up to 2 phases

The picture at right is the cost profile if we invest in the infrastructure of both phases. The darkest dotted line is the cost profile of abandoning the remainder of the investment. The solid line represents the costs if we invest in only phase one. The dotted light line represents the cost profile of investing in both phases at two different times (sequential investment). The semi dotted/solid light line represent a parallel investment in two phases.

3.2 Mathematical Model

Based on the above-mentioned assumptions we present a novel mathematical model. The mathematical model provided in this research is a deterministic Mixed Integer Program (MIP). It is deterministic because we assume that all of the model parameters and inputs are known a priori without much uncertainty. The Deterministic Single-project Phased-investment Problem (DSPP) is presented below.

3.2.1 Model Parameters and Variables

The input/parameters of the model are mainly monetary. The list of all parameters and variables for DSPP are as follows. For convenience and better representation of the mathematical model, the parameters all start with Uppercase letters and variables all start with lowercase letters. All model parameters and variables are summarized in Table 2.

Table 2 Deterministic single-project phased investment problem parameters and variables

Variables	
ub	Number of phases selected for implementation
Variables regarding infrastructure	
it_{ub}	Construction duration for infrastructure based on actual number of phases to be implemented
$cost_{ub}$	Construction cost for the infrastructure required for implementing ub phases.
Variables regarding phases	
x_{ij}^{t}	Binary variable that equals 1 if phases i through j start their implementation at time t
Other main variables	
b_t	Available budget at the beginning of each time period t

n_t	Number of phases that have already been implemented or are being implemented at time t
Variables used for linearization	
$nt_{t,i}$	Binary variable for linearization
$cn_{i,t}$	Binary representation of number of phases that have already been implemented or are being implemented at time t for linearization
nc_i	Binary variable for linearization of infrastructure cost
Parameters	
T	Planning period
UB	Maximum number of phases of project
$B0$	Initial available budget
R	Risk free interest rate
I	Inflation rate
M_T	Big-M value used for timing
Times	Set of time periods = 0 .. T
Phases	Set of phases = 1 .. UB
$P_{i,t}$	Profit gained from first i phases at time t
IT_{ub}	Preparation (construction) time for infrastructure of ub phases
$ICOST_{ub}$	Preparation (construction) cost for infrastructure of ub phases
$PT_{i,j}$	Duration required for implementation of each phase $i-j$ when phases $i-j$ are implemented together
$CCPA_{i,j}$	Construction cost of each phases between i,j if phases $i-j$ are being done together
$OCPA_{i,t}$	Operation cost at time t when Number of phases that have already been implemented or are being implemented at time t is i

The main decision variables in our model are related to the initial investment in infrastructures for preparation of future phases, and the time at which each (grouped) potential phase is implemented.

3.2.2 MIP Formulation (DSPP)

3.2.2.1 Objective Function

In our formulation we only consider monetary benefits. Therefore the objective function maximizes the Net Present Value (NPV) of our budget at the end of the planning period (equation (3-4)).

$$\max z = \frac{B_T + \sum_i P_{i,T} \times cn_{i,T} - \left(\sum_i \sum_j (j-i+1) \times x_{i,j}^t + \sum_i OCPA_{i,T} \times cn_{i,T}\right) \times (1+I)^T}{(1+R)^T} \tag{3-4}$$

The first part of the objective function is the remaining money at the beginning of the last year of the planning period, *T*. The next component is the profit gained at the end of the planning period and thereafter. If the project still yields profit after the planning period, the future net profit is discounted back to the year of the planning period, *T* and added to the profit of the last year of the planning period (*T*). The third component of the objective function is all the costs at time *T* and thereafter. All components discounted back to time 0 using the risk free interest rate.

3.2.2.2 Constraints related to the size of investment (DSPP)

The majority of the constraints of our MIP formulation are either related to the initial infrastructural investment, or phases of the investment, or the scheduling, or budget constraints. The constraints for the initial infrastructural investments are for calculating the size of the initial infrastructural investment. The size of the investment is based on the number of potential phases that can be built in the future (ub). Depending on the size, the cost of such an investment and required time for completion of the infrastructure can be calculated.

$$n_t \leq ub \qquad \forall t \tag{3-5}$$

$$cost_{ub} = \sum_i ICOST_i \times nc_i \tag{3-6}$$

$$it_{ub} = \sum_i IT_i \times nc_i \tag{3-7}$$

$$ub = \sum_i i \times nc_i \tag{3-8}$$

$$\sum_i nc_i \leq 1 \tag{3-9}$$

Constraints (3-5) limit the total number of phases that are implemented at each time to the maximum invested infrastructure. Constraints (3-6) and (3-7) are for calculating the infrastructure cost and install duration. For them being linear, we need to express the number of phases selected for implementation, ub, using binary variables. This is done using constraints (3-8) and (3-9).

3.2.2.3 Constraints related to phases (DSPP)

Constraint (3-10) – (3-15) are the phase related constraints:

$$n_0 = \sum_i \sum_j (j - i + 1) \times x_{i,j}^0 \tag{3-10}$$

$$n_t = n_{t-1} \sum_i \sum_j (j - i + 1) \times x_{i,j}^t \qquad \forall t \geq 1 \tag{3-11}$$

$$n_t = \sum_i i \times cn_{i,t} \qquad \forall t \tag{3-12}$$

$$\sum_i cn_{i,t} \leq 1 \qquad \forall t \tag{3-13}$$

$$\sum_{j \geq i} x_{i,j}^t \leq \sum_{l \leq i-1} \sum_{t' \leq t} x_{l,i-1}^{t\prime} \qquad \forall i \in \{2, \dots, UB\}, t \tag{3-14}$$

$$\sum_t \sum_{i \leq l} \sum_{j \geq l} x_{i,j}^t \leq 1 \qquad \forall l \in \{1, \dots, UB\} \tag{3-15}$$

Constraints (3-10) and (3-11) are for calculating the number of phases implemented/being implemented at different times. Constraints (3-12) and (3-13) are for representing the number of phases that are implemented/are being implemented using binary variables so that we would have linear constraints when calculating the different costs and times for phases. Constraints (3-14) prevent implementation of succeeding phases prior to the implementation of phases that are preceding them. Constraints (3-15) prevent the assignment of a phase to two different groups of phases.

3.2.2.4 Scheduling constraints of phases (DSPP)

These group of constraints are for timing of each phase.

$$it_{ub} \leq \sum_j \sum_t t \times x_{1,j}^t + M_T \times \left(1 - \sum_j \sum_t x_{1,j}^t\right) \tag{3-16}$$

$$t \times x_{i,j}^t + (j - i + 1) \times PT_{i,j} \times x_{i,j}^t \leq \sum_{l \geq j+1} \sum_{t' \geq t} t' \times x_{j+1,l}^{t'} + M_T \times \left(1 - \sum_{l \geq j+1} \sum_{t' \geq t} x_{j+1,l}^{t'}\right) \qquad \forall i,j \in \{2, \dots, UB\} \mid j \geq i,\ t \tag{3-17}$$

$$\sum_i \sum_{j \geq i} x_{i,j}^t \leq 1 \qquad \forall t \tag{3-18}$$

Constraint (3-16) assures that the first group of phases are implemented after the infrastructure is completed and successfully implemented. Constraints (3-17) ensure that each phase is implemented after the completion of implementation of its preceding phases. Constraints (3-18) prevent multiple groups of phases to start their implementation together.

3.2.2.5 Budget constraints (DSPP)

At each time, the costs for investing in the phases and the operation and maintenance costs should not exceed the budget of that time. These budgetary constraints are for calculating the available budget at the beginning of each time period and limiting the expenses during a time period to the available budget at the beginning of that time period.

$$b_1 = B0 - cost_{ub} - \left(\sum_i \sum_{j\geq i}(j-i+1) \times CCPA(i,j) \times x_{i,j}^0 + \sum_i OCPA_{i,0} \times cn_{i,0}\right) \times (1+R) + \sum_i P_{i,0} \times cn_{i,0} \tag{3-19}$$

$$b_t = \left(b_{t-1} - \left(\sum_i \sum_{j\geq i}(j-i+1) \times CCPA(i,j) \times x_{i,j}^{t-1} + \sum_i OCPA_{i,t-1} \times cn_{i,t-1}\right) \times (1+I)^{t-1}\right) \times (1+R) + \sum_i P_{i,t-1} \times cn_{i,t-1} \qquad \forall t \in \{2, \ldots, T\} \tag{3-20}$$

$$cost_{ub} + \sum_i \sum_{j\geq i}(j-i+1) \times CCPA(i,j) \times x_{i,j}^0 + \sum_i OCPA_{i,0} \times cn_{i,0} \leq B0 \tag{3-21}$$

$$\left(\sum_i \sum_{j\geq i}(j-i+1) \times CCPA(i,j) \times x_{i,j}^{t-1} + \sum_i OCPA_{i,t-1} \times cn_{i,t-1}\right) \times (1+I)^{t-1} \leq b_t \qquad \forall t \geq 1 \tag{3-22}$$

Constraints (3-19) and (3-20), are the updates on the available budget at the beginning of each time period. The available budget at the beginning of each time period is equal to remaining budget from the previous time period in the current period's value (incorporating time value of money) plus the profits earned at the end of the last period as a result of implemented phases in the previous period. Constraints (3-21) and (3-22) are the budget limitations during different times.

3.2.2.6 Variable domain constraints (DSPP)

The domain of the variables used for this model are presented below:

$$x_{i,j}^{t}, cn_{i,t}, nc_i, nt_{t,i} \in \{0,1\} \qquad (3\text{-}23)$$

$$ub, it_{ub}, cost_{ub}, b_t, n_t \geq 0 \qquad (3\text{-}24)$$

3.2.2.7 The complete MIP model

The complete mixed integer programming model is summarized in the following:

$$\max z = \frac{B_T + \sum_i P_{i,T} \times cn_{i,T} - \left(\sum_i \sum_j (j-i+1) \times x_{i,j}^{t} + \sum_i OCPA_{i,T} \times cn_{i,T}\right) \times (1+I)^T}{(1+R)^T} \qquad (3\text{-}4)$$

Subject to:

$$n_t \leq ub \qquad \forall t \qquad (3\text{-}5)$$

$$cost_{ub} = \sum_i ICOST_i \times nc_i \qquad (3\text{-}6)$$

$$it_{ub} = \sum_i IT_i \times nc_i \qquad (3\text{-}7)$$

$$ub = \sum_i i \times nc_i \qquad (3\text{-}8)$$

$$\sum_i nc_i \leq 1 \qquad (3\text{-}9)$$

$$n_0 = \sum_i \sum_j (j - i + 1) \times x_{i,j}^0 \tag{3-10}$$

$$n_t = n_{t-1} \sum_i \sum_j (j - i + 1) \times x_{i,j}^t \qquad \forall t \geq 1 \tag{3-11}$$

$$n_t = \sum_i i \times cn_{i,t} \qquad \forall t \tag{3-12}$$

$$\sum_i cn_{i,t} \leq 1 \qquad \forall t \tag{3-13}$$

$$\sum_{j \geq i} x_{i,j}^t \leq \sum_{l \leq i-1} \sum_{t' \leq t} x_{l,i-1}^{t'} \qquad \forall i \in \{2, \dots, UB\}, t \tag{3-14}$$

$$\sum_t \sum_{i \leq l} \sum_{j \geq l} x_{i,j}^t \leq 1 \qquad \forall l \in \{1, \dots, UB\} \tag{3-15}$$

$$it_{ub} \leq \sum_j \sum_t t \times x_{1,j}^t + M_T \times \left(1 - \sum_j \sum_t x_{1,j}^t\right) \tag{3-16}$$

$$t \times x_{i,j}^t + (j - i + 1) \times PT_{i,j} \times x_{i,j}^t \leq \sum_{l \geq j+1} \sum_{t' \geq t} t' \times x_{j+1,l}^{t'} + M_T \times \Big(1 - \sum_{l \geq j+1} \sum_{t' \geq t} x_{j+1,l}^{t'}\Big) \qquad \forall i, j \in \{2, \dots, UB\} \mid j \geq i,\ t \tag{3-17}$$

$$\sum_i \sum_{j \geq i} x_{i,j}^t \leq 1 \qquad \forall t \tag{3-18}$$

$$b_1 = B0 - cost_{ub} - \Big(\sum_i \sum_{j \geq i} (j - i + 1) \times CCPA(i,j) \times x_{i,j}^0 + \sum_i OCPA_{i,0} \times cn_{i,0}\Big) \times (1 + R) + \sum_i P_{i,0} \times cn_{i,0} \tag{3-19}$$

$$b_t = \Big(b_{t-1} - \Big(\sum_i \sum_{j \geq i} (j - i + 1) \times CCPA(i,j) \times x_{i,j}^{t-1} + \sum_i OCPA_{i,t-1} \times cn_{i,t-1}\Big) \times (1 + I)^{t-1}\Big) \times (1 + R) + \sum_i P_{i,t-1} \times cn_{i,t-1} \qquad \forall t \in \{2, \dots, T\} \tag{3-20}$$

$$cost_{ub} + \sum_i \sum_{j \geq i} (j - i + 1) \times CCPA(i,j) \times x_{i,j}^0 + \sum_i OCPA_{i,0} \times cn_{i,0} \leq B0 \tag{3-21}$$

$$\left(\sum_i \sum_{j\geq i}(j-i+1)\times CCPA(i,j)\times x_{i,j}^{t-1}+\sum_i OCPA_{i,t-1}\times cn_{i,t-1}\right)\times(1+I)^{t-1}\leq b_t \qquad \forall t\geq 1 \tag{3-22}$$

$$x_{i,j}^{t}, cn_{i,t}, nc_i, nt_{t,i}\in\{0,1\} \tag{3-23}$$

$$ub, it_{ub}, cost_{ub}, b_t, n_t \geq 0 \tag{3-24}$$

3.3 Case Study

The case study used for this section is based on an example from (Zhao & Tseng, 2003) which is about constructing a public parking garage by a county in the Washington, D.C. area. The maximum number of floors the parking garage can have is assumed to be 6 (*i.e. UB=6*). The risk free interest rate and inflation rate are 8% and 5% respectively (*R*=8% and *I* = 5%). In this example, each level represents a phase. The life time of this garage is assumed to be 15 years (T= 14). After the 15 years, the parking garage generates no more revenue nor cost for the owner. This could be a result of a Build Operate Transfer (BOT) contract in which the garage is transferred to a public authority after 15 years of operation by a private entity.

3.3.1 Cost parameters

(Zhao & Tseng, 2003) gathered the cost data from a feasibility study done by the county. Their construction costs are presented in Table 3.

Table 3 Construction cost summary for parking public parking garage without synergies between levels [1]

Costs type	Value
Site Preparation	$300,000

[1] Source: Zhao, T., and Tseng, C.L. (2003). Valuing flexibility in infrastructure expansion. *Journal of infrastructure systems, 9(3), 89-97.*

Fixed cost for foundation	$1 million
Variable cost for foundation	$100,000/level
Superstructure and miscellaneous	$800,000/level
Construction cost for expansion	$850,000/level

As it can be seen in Table 3, the synergy between costs of different levels was not reported. In other words, the benefit of constructing multiple levels together is not reflected in the costs. We modify the costs by introducing some parameters A_i, B_i, C_i, to account for the cost interdependencies for the three types of variable costs of Table 3. The summary of the modified costs are presented in Table 4.

Table 4 Modified construction costs accounting for cost interdependencies among levels

Costs type	**Value**
Site Preparation	$300,000
Cost for building a foundation for only one level	$\$1\ M + A_1 \times \$100{,}000$
Cost for building a foundation for two levels	$\$1\ M + 2A_2 \times \$100{,}000$
Cost for building a foundation for three levels	$\$1\ M + 3A_3 \times \$100{,}000$
Cost for building a foundation for four level	$\$1\ M + 4A_4 \times \$100{,}000$
Cost for building a foundation for five levels	$\$1\ M + 5A_5 \times \$100{,}000$
Cost for building a foundation for six levels	$1 M + $600,000
Superstructure and miscellaneous for one level	$B_1 \times \$800{,}000$
Superstructure and miscellaneous for two levels parallel	$B_2 \times \$800{,}000/level$
Superstructure and miscellaneous for three levels parallel	$B_3 \times \$800{,}000/level$
Superstructure and miscellaneous for four levels parallel	$B_4 \times \$800{,}000/level$
Superstructure and miscellaneous for five levels parallel	$B_5 \times \$800{,}000/level$
Superstructure and miscellaneous for six levels parallel	$800,000/level
Construction cost for expansion for one level	$C_1 \times \$850{,}000$
Construction cost of expansion for two levels parallel	$C_2 \times \$850{,}000/level$
Construction cost of expansion for three levels parallel	$C_3 \times \$850{,}000/level$
Construction cost of expansion for four levels parallel	$C_4 \times \$850{,}000/level$

Construction cost of expansion for five levels parallel	$C_5 \times \$850{,}000/level$
Construction cost of expansion for six levels parallel	\$850,000/level

Using the modified values from Table 4, we can calculate the cost parameters of our model as illustrated in Table 5.

Table 5 Model input Cost Parameters

Costs Parameter	Value
$Cost_1$: $investment\ construction\ cost\ of\ the\ foundation\ for\ 1\ level$	$\$1\ M + \$300{,}000 + A_1 \times \$100{,}000$
$Cost_2$: $investment\ construction\ cost\ of\ the\ foundation\ for\ 2\ levels$	$\$1\ M + \$300{,}000 + 2A_2 \times \$100{,}000$
$Cost_3$: $investment\ construction\ cost\ of\ the\ foundation\ for\ 3\ levels$	$\$1\ M + \$300{,}000 + 3A_3 \times \$100{,}000$
$Cost_4$: $investment\ construction\ cost\ of\ the\ foundation\ for\ 4\ levels$	$\$1\ M + \$300{,}000 + 4A_4 \times \$100{,}000$
$Cost_5$: $investment\ construction\ cost\ of\ the\ foundation\ for\ 5\ levels$	$\$1\ M + \$300{,}000 + 5A_5 \times \$100{,}000$
$Cost_6$: $investment\ construction\ cost\ of\ the\ foundation\ for\ 6\ levels$	\$1.9 M
CC_{ii}: $overall\ investment\ or\ construction\ cost\ of\ level\ i\ by\ itself$	$B_1 \times \$800{,}000 + C_1 \times \$850{,}000/level$
CC_{ij}: $overall\ investment\ or\ construction\ cost\ of\ two\ levels\ i\ and\ j\ in\ parallel$	$B_2 \times \$800{,}000 + C_2 \times \$850{,}000/level$
CC_{ij}: $overall\ investment\ or\ construction\ cost\ of\ three\ levels\ i\ through\ j\ in\ parallel$	$B_3 \times \$800{,}000 + C_3 \times \$850{,}000/level$
CC_{ij}: $overall\ investment\ or\ construction\ cost\ of\ four\ levels\ i\ through\ j\ in\ parallel$	$B_4 \times \$800{,}000 + C_4 \times \$850{,}000/level$
CC_{ij}: $overall\ investment\ or\ construction\ cost\ of\ five\ levels\ i\ through\ j\ in\ parallel$	$B_5 \times \$800{,}000 + C_5 \times \$850{,}000/level$

CC_{16}: $overall\ investment\ or\ construction\ cost\ of\ all\ levels\ 1\ through\ 6$ $1.65 $M/level$

The recently introduced parameters, A_i, B_i, C_i, are based on economy of scale. For simplicity, assume for each number of levels that are going to be done in parallel, i, $A_i = B_i = C_i$. Table 6 stores an example of values for the economy of scale input parameters. Sensitivity analysis on these inputs can be done to illustrate how these values can affect optimal decisions.

Table 6 Input values for economy of scale adjustment parameters

$\boldsymbol{i}$	**Description**	$\boldsymbol{A_i = B_i = C_i}$
1	Level is constructed alone	1.5
2	Two levels constructed together	1.4
3	Three levels constructed together	1.3
4	Four levels constructed together	1.2
5	Five levels constructed together	1.1

The maintenance and operation costs are assumed to be constant for each phase (level). The operation cost is $100 per parking space. Since each level has 100 parking spaces, the operating cost per phase implemented is $10,000.

3.3.2 Revenue and profit parameters

The annual revenue of each occupied parking space is $3,600 from parking fees. Obviously, if a parking space is empty it will not generate any revenue. The actual revenue during each time, t, from the parking garage is calculated using equation (38).

$$\text{Revenue}_t = \$3{,}600 \times \min\{demand_t, Capacity_t\} = \$3{,}600 \times \{demand_t, 100 \times po_t\} \quad (38)$$

Where po_t is the variable which represented the number of phases (parking levels) that have been completed at time t.

The Revenue is therefore related to the demand at each time. The predicted average annual demand of (Zhao & Tseng, 2003) is presented in Table 7. They predicted this value based on the average and standard deviation gained from historical data of 10 years prior to the time of their study.

Table 7 Average demand at each time[2]

Year	Time t	Average demand (units)
1	0	250
2	1	263
3	2	276
4	3	290
5	4	305
6	5	320
7	6	336
8	7	353
9	8	371
10	9	390
11	10	410
12	11	431
13	12	452
14	13	475
15	14	499

3.3.3 Time related parameters

The site preparation and foundation construction time are added together to form the infrastructure preparation time parameter, CT_i. The time period of the study is 15 years.

[2] Source: Zhao, T., and Tseng, C.L. (2003). Valuing flexibility in infrastructure expansion. *Journal of infrastructure systems, 9(3), 89-97.*

Starting from year 0, T is equal to 14. The construction times that are based on the number of phases being done in parallel, CT_{ij}, and CT_i, are presented in Table 8 and Table 9.

Table 8 Construction time for foundation and site preparation time based on potential number of levels

ub	$\boldsymbol{CT_{ub}\ (months)}$
1	6
2	6.5
3	7
4	7.5
5	8
6	9.5

Table 9 Construction time for each level based on number of levels being done in parallel

# of levels being constructed in parallel	$\boldsymbol{CT_{ij}\ \ months/level}$
1	3
2	2.7
3	2.5
4	2.2
5	2
6	1.9

The expansion decisions are assumed to be made at the beginning of years.

3.4 Results and Sensitivity Analysis

The MIP formulation was solved using Xpress Optimization suite. In general, the running time for a problem of the size of the case study which has 401 variables and 1016 constraints was on average 3 seconds using an Intel Core i5-2400 CPU @3.10 GHz computer with 4.00 GB of installed memory (RAM).

The optimal decision for the case study is to build three levels (capacity of 300) at year 1 which is the first year that we can construct any level. Preparing the construction site and constructing the foundation will consume a portion of year 0, therefore year 1 is the soonest that we can construct any level. The objective function value that represents the NPV of the firm's future budget is equal to $8,178,280. Given that the initial budget was $8 M, the overall benefit (profit) of doing the project is $178,280 in today's money. Figure 6 illustrates the number of parking levels during each time that are the optimal solution.

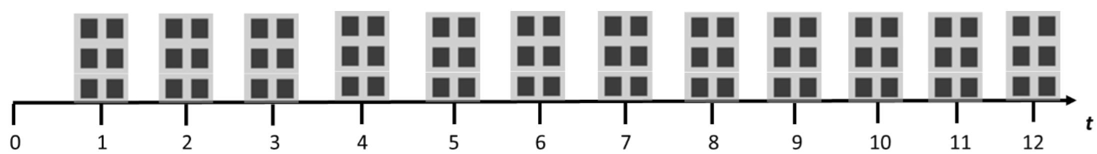

Figure 6 Number of parking levels during each time for the base case

3.4.1 Sensitivity Analysis on initial budget when risk free interest rate is 8%

In the base case example, the available budget was set to be $8 M. We vary this amount and solve the problem for each case. The results are depicted in Figure 7. The results indicate that the only profitable option when the risk free interest rate is 8% is to invest and build 3 phases together all at the soonest possible time. Any other investment option is not economically feasible and therefore it is optimal to not invest at all in the parking garage project. The option of investing becomes feasible when we have at least approximately $7,678,000 available at time zero. If the budget is less, we do not have enough funds to build the foundation and build three levels at year one which is the only economically feasible option and therefore as explained earlier the optimal decision is to not invest.

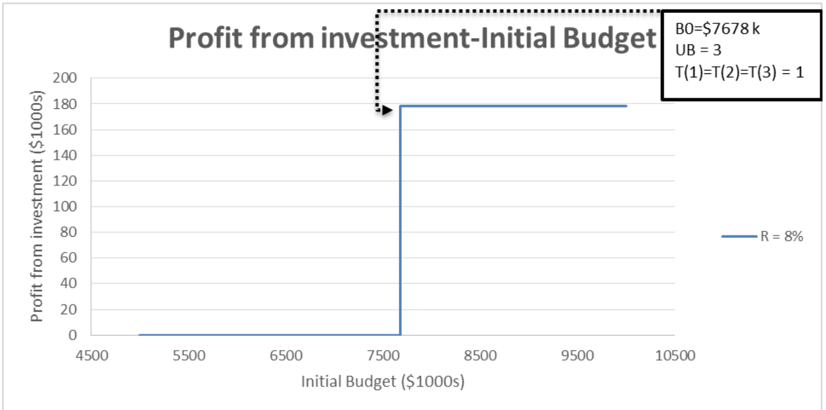

Figure 7 Profit from investment based on initial budget (Risk free interest rate = 8%)

3.4.2 Sensitivity Analysis on initial budget when risk free interest rate is 5%

Figure 8 illustrates the optimal decisions and their respective profits depending on the available budget at time zero.

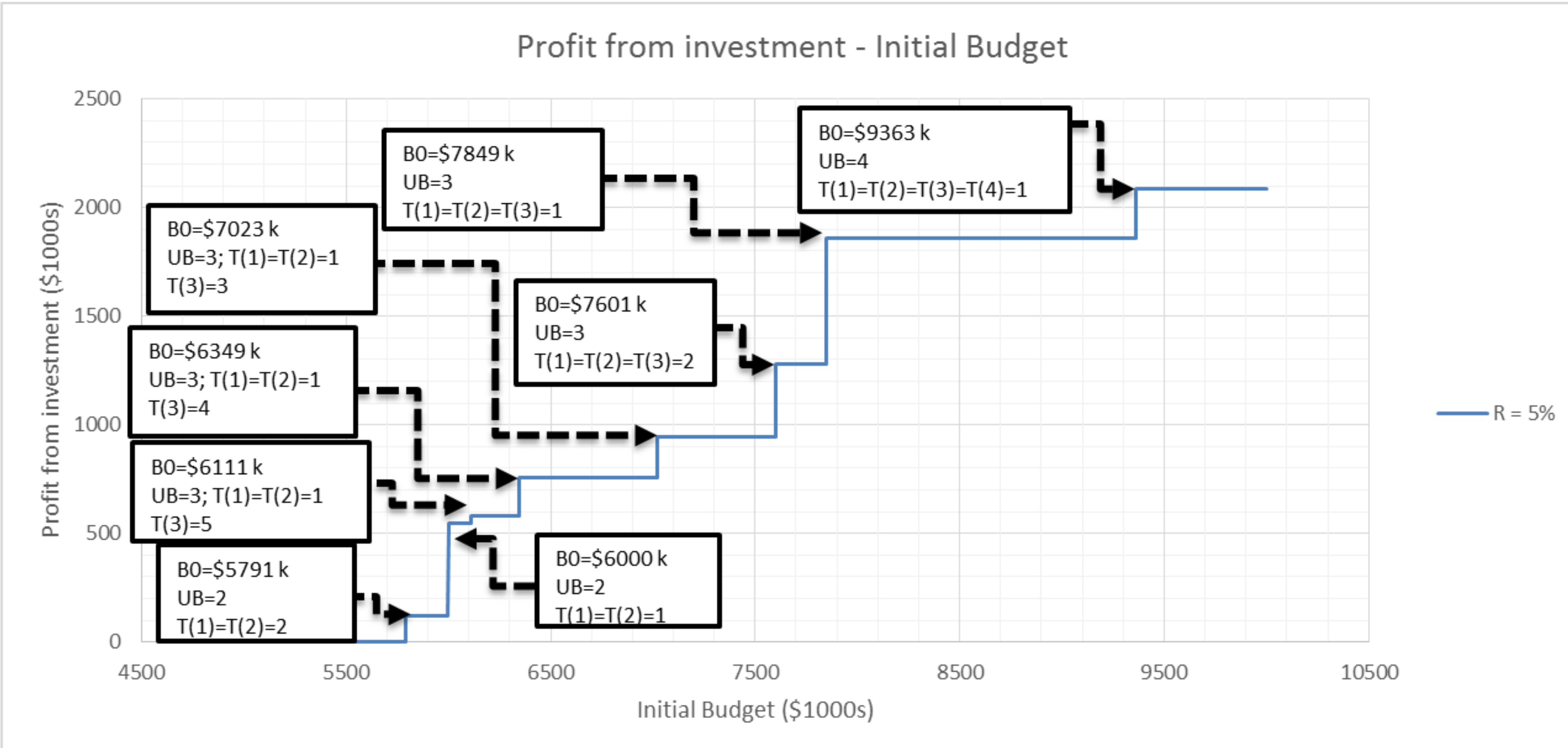

Figure 8 Profit from investment based on initial budget (Risk free interest rate = 5%)

Based on Figure 8, if the initial budget is less than or equal to $5,791,000, the optimal decision is not to invest at all in the parking garage project. However, if the available budget exceeds that amount and is still less than $6M, the project will be favorable because we have enough capital available to invest in two phases (levels) simultaneously and benefit from the economy of scale in construction. The optimal decision in this case is to invest in phases 1 and 2 together at time (year) 2. We cannot invest in phases 1 and 2 at year one due to lack of available budgets.

If the capital exceeds $6M but is still less than $6,111,000, the optimal decision is to build two phases. However, this time since we have enough budget, we can build the phases at the earliest time (year 1). As a result the NPV of the profit from the investment will increase by approximately $424,000.

If the initial budget is equal to or greater than $6,111,000 and less than $6,349,000, we are capable of building three levels. The optimal decision in this case is to invest and build levels 1 and 2 at year 1 and build level 3 at year 5. Figure 9 depicts the optimal parking garage levels during each time for this case. The increase in the initial budget with respect to the previous case allows us to invest in more phases and therefor increase the profit from investing in the parking garage project.

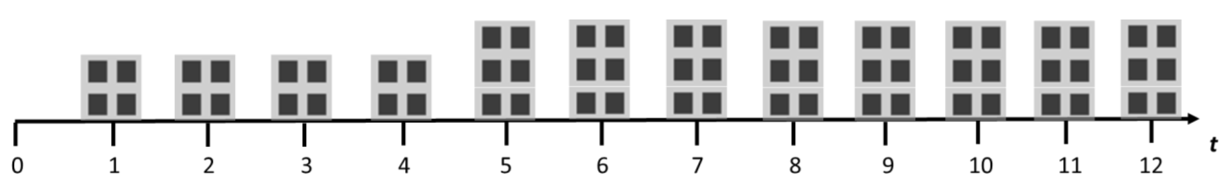

Figure 9 Number of parking levels during each time for the case of risk free interest rate of 5% and initial budget between $6.111 M and $6.349 M

As the initial budget increases, the profit from investing in the parking garage increases as well. This general trend continues until we exceed the initial budget of $9,363,000. After this amount, the profit from investing in the parking garage project will not increase anymore because we already are gaining the biggest benefit of economy of scale by building all 4 levels at the soonest time possible (year 1). Adding another level is not beneficial due to the combination of maintenance costs, construction costs and the demands at different times.

Table 10 summarizes the results of the model under different initial available budget assumptions.

Table 10 Optimal solution results for risk free interest rate of 5% by varying initial budget

B0 (Initial budget) $1000s	***Z (NPV of money) $1000s***	***NPV profit from investment ($1000s)***	***UB***	***Year investing in phase 1***	***Year investing in phase 2***	***Year investing in phase 3***	***Year investing in phase 4***
3500	3500	0	0				
5780	5780	0	0				
5790	5790	0	0				
5791	5912.42	121.42	2	2	2		
5999	6120.42	121.42	2	2	2		
6000	6544.96	544.96	2	1	1		
6110	6654.96	544.96	2	1	1		
6111	6693.45	582.45	3	1	1	5	
6348	6930.45	582.45	3	1	1	5	
6349	7106.56	757.56	3	1	1	4	
7022	7779.56	757.56	3	1	1	4	
7023	7964.92	941.92	3	1	1	3	
7600	8541.92	941.92	3	1	1	3	
7601	8881.01	1280.01	3	2	2	2	
7848	9128.01	1280.01	3	2	2	2	
7849	9708.4	1859.4	3	1	1	1	
9362	11221.4	1859.4	3	1	1	1	
9363	11450	2087	4	1	1	1	1
10000	12087	2087	4	1	1	1	1

Many managerial insights can be drawn by looking at Figure 8 and Table 10. For example, we can see that there is a big increase in the NPV from investing in the parking garage project when the initial available budget is marginally increased. For instance, if a firm has an initial available budget of $5.9 M, they can increase the NPV of their profit by $424,000 if they borrow an amount as small as $100,000.

3.4.3 Sensitivity Analysis on initial budget when risk free interest rate is 2%

Figure 10 and Table 11 illustrate how the NPV of investment in the parking garage varies by varying the initial budget available when the risk-free interest rate is 2%. The general trend is similar to what was observed when the risk-free interest rate was 5%. If the initial budget is more than $9,586,000 the optimal decision is to construct four levels at year 1. Increasing the budget further will not cause any changes.

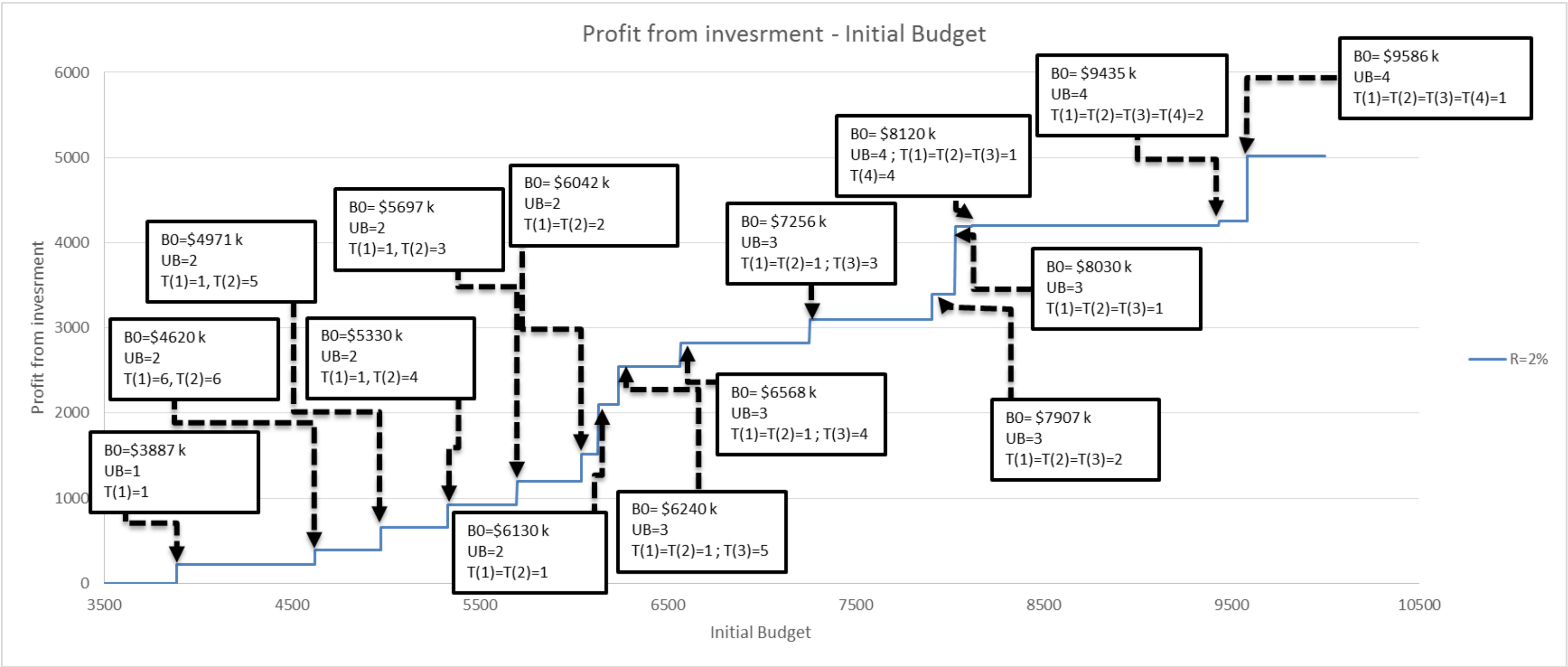

Figure 10 Profit from investment based on initial budget (Risk free interest rate = 2%)

However, the optimal decision in lower available budgets is different for the case with a higher risk free interest rate (5%). In the current case, just building one level is still a profitable investment. This is mainly because, we get to collect the profit of having a one level parking lot for the time period under study (15 years). Since the risk-free interest rate is lower, the value of the future money is closer to the value of money today.

Using the results of the model, similar managerial insights and suggestions can be made in this case as well. These suggestions could be with regard to the size of a loan and the acceptable interest of a loan.

Table 11 Optimal solution results for risk free interest rate of 2% by varying initial budget

B0 (Initial budget) $1000s	***Z (NPV of money) $1000s***	***NPV profit from investment ($1000s)***	***UB***	***Year investing in phase 1***	***Year investing in phase 2***	***Year investing in phase 3***	***Year investing in phase 4***
3500	3500	0	0				
3886	3886	0	0				
3887	4113.48	226.48	1	1			
4619	4845.48	226.48	1	1			
4620	5012.72	392.72	2	1	6		
4970	5362.72	392.72	2	1	6		
4971	5627.88	656.88	2	1	5		
5329	5985.88	656.88	2	1	5		
5330	6256.88	926.88	2	1	4		
5696	6622.88	926.88	2	1	4		
5697	6899.82	1202.82	2	1	3		
6041	7243.82	1202.82	2	1	3		
6042	7555.86	1513.86	2	2	2		
6129	7642.86	1513.86	2	2	2		
6130	8226.5	2096.5	2	1	1		
6239	8335.5	2096.5	2	1	1		
6240	8786.89	2546.89	3	1	1	5	
6567	9113.89	2546.89	3	1	1	5	
6568	9384.89	2816.89	3	1	1	4	
7255	10071.9	2816.9	3	1	1	4	

7256	10348.8	3092.8	3	1	1	3	
7906	10998.8	3092.8	3	1	1	3	
7907	11299.6	3392.6	3	2	2	2	
8029	11421.6	3392.6	3	2	2	2	
8030	12223.1	4193.1	3	1	1	1	
8119	12312.1	4193.1	3	1	1	1	
8120	12321.5	4201.5	4	1	1	1	4
9434	13635.5	4201.5	4	1	1	1	4
9435	13690.4	4255.4	4	2	2	2	2
9585	13840.4	4255.4	4	2	2	2	2
9586	14603	5017	4	1	1	1	1
10000	15017	5017	4	1	1	1	1

3.4.4 Summary of sensitivity analysis on different interest rates and different available budgets

Figure 11 captures the results for all three aforementioned cases. It is worth noting that for the case with a 10% interest rate, no investment happens at all. From the results, we can conclude that for lower risk free interest rates, the value of the project increases. This is mainly due to the fact that many of the earnings are gained in the future and are in terms of future money. Another conclusion is that as the risk-free rate decreases, many options become economically viable and also investing in fewer phases that requires less initial capital becomes an option.

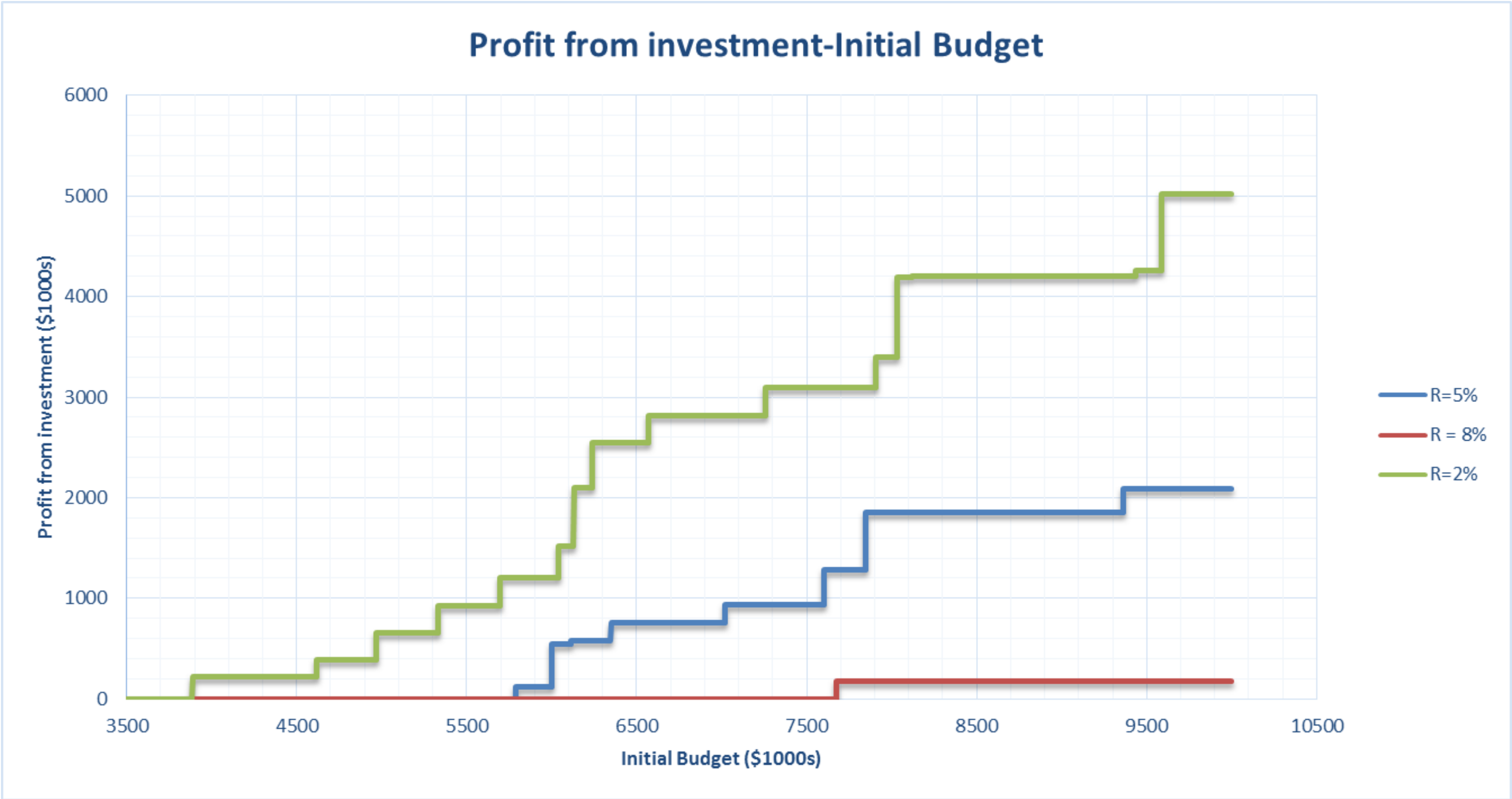

Figure 11 Profit from investment as a function of both risk-free interest rate and initial available budget

Chapter 4: Stochastic Multi-phase Project Optimization

This chapter focuses on finding the optimal timing and phasing of multi-phase projects when some parameters of the project are subject to uncertainty. Hedging unfavorable outcomes of uncertain parameters is one of the main purposes of following a multi-phase paradigm.

4.1 Source of Uncertainty: Demand

In the deterministic model presented in the previous chapter, one of the main drivers of profits/revenues of a project was the demand for the different phases of the infrastructure. Assuming that the unit price of a service/good does not change drastically, the overall profit is proportional to the demand.

Infrastructure projects such as transportation infrastructures are built to serve their consumers for long periods. Prediction and estimates of the close future are oftentimes wrong or full of uncertainties. These uncertainties increase as the prediction duration increases. Therefore, many of the infrastructure planning problems are associated with huge uncertainties. These uncertainties are such that even professionals do not agree on a unique predictive model. Figure 12 illustrates the demand forecast results of analysis of four different consultants all for the same toll road. The vertical scale is intentionally left out due to confidentiality. However, the difference between the highest and lowest forecast at different times (in percentages) were reported. These differences can be found in Table 12.

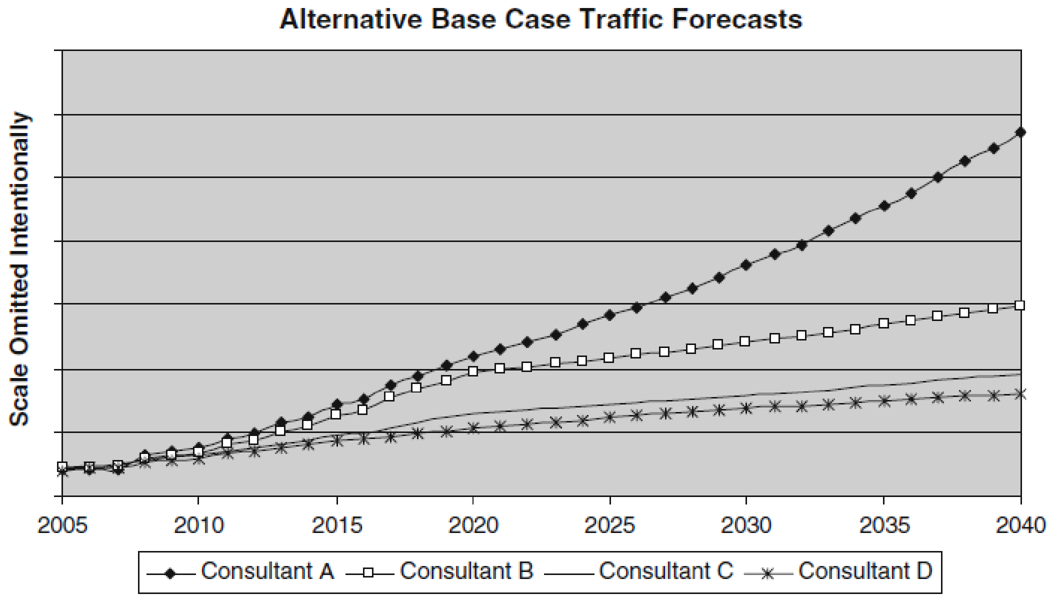

Figure 12 Same toll road, different traffic forecasts (Source: Bain (2009)[3])

Table 12 Difference between the highest and lowest base-case forecast as a function of time for the same road (Source: Bain (2009)[4])

Forecast period (from project opening) (years)	Difference between the highest and lowest base-case forecast (%)
5	26%
10	66%
15	106%
20	130%
25	164%
30	204%
35	255%

Table 12 also shows how the predictions of professional firms vary from each other as the prediction duration increases. To deal with these uncertainties, there are many naïve approaches such as using expected values of outcomes or using the prediction of one consultant. Obviously, even if the best consultant is picked, the predictions cannot

[3] Source: (Bain, 2009)
[4] Source: (Bain, 2009)

deemed to be accurate. There always exists the likelihood of observing an output which is very far from the point predictions. This illustrates the need for modelling the uncertainties. One of the most common methods for modelling uncertainties in infrastructure analysis is real option valuation.

4.2 Introduction to Real Options Analysis

Real Options are derived from options analysis. Options are financial derivatives. They are defined as the right, but not the obligation to, purchase or sell a product. This right has an expiration date. A premium should be paid to acquire the option, which is based on the value of the product, or the stock, and the time remaining until the option expiration date. In finance, the right to purchase a stock is termed a *call option*, and the right to sell a stock is called a *put option*. In terms of allowable exercise dates, options are usually either *American* or *European*. European options can be only exercised on the expiration date. American options can be exercised at any date prior to the expiration date, including the expiration date. American options are usually worth more as they have fewer restrictions.

There are also other options called *exotic options*. An example of an exotic option is a *Bermudan option*. A Bermudan option can be seen as a special case for the American option, and it can be exercised early on any of so many pre-specified dates. For example, it can be exercised on a particular day of each month.

The financial options theory focuses on the options in which the underlying asset is stock. Whenever the underlying asset is tangible, like an infrastructure project, the nomenclature of real options is used.

Real options can be defined as the right to take certain business initiatives. These rights add flexibility to decision-making; therefore, real options can be seen as flexibilities. Real option valuation is used in corporate finance as a tool to evaluate investment decisions in uncertain environments where adding flexibility is valuable. This approach is used as a replacement or an addition to the traditional DCF analysis, which has several shortcomings. Some major shortcomings of traditional DCF analysis are:

1) Small or no flexibility with a high degree of commitment (i.e. allowing for no flexibility),

2) Neglecting volatility and using expected values, and

3) Difficulty in finding and updating risk adjusted discount rates.

Real-option valuation overcomes these shortcomings by inserting flexibility using options and utilizing decision nodes. This way, investments are pursued if uncertain conditions are favored. Real-option valuation technics generally use risk-neutral probabilities and risk-free discount rates for financial evaluations because an interim decision can change the risk-adjusted discount rate evaluated at the beginning of the project.

In terms of applications, real options can be on the project scope, like the option to expand and the option to contract. They could be about the timing of a project. Some of the most popular timing options are the option to defer and the option to abandon. The option to abandon allows management to bail out of a project and possibly get a salvage value. These options are similar to put options. The option to defer, on the other hand, allows management to start an investment/project at any time at which it is the most valuable. This option is analogous to a call option because it gives the right to invest in the project at pre-specified times in the future. An example of an option to delay for a road widening investment, which has value in volatile conditions, is illustrated in the following. Although this example is an extreme case in which uncertainty fades away as a result of waiting, it manages to convey the purpose of the example in a simple and tangible fashion.

Assume that a lane expansion investment is under consideration. The current road has a capacity which can service an Average Annual Daily Traffic (AADT) of 6000 Vehicles/day. The current demand for this road has a ballpark figure of 7000 AADT. If a lane is added and the road is expanded, the capacity will increase to 9000 AADT. The cost of such an expansion is $500k. Future demand is uncertain. After one year, the demand might go up to 8000 AADT with a probability of q=0.5, or it might decrease to 6000 AADT with a probability of 1-q=0.5. After the change in demand at the end of the first year, demand will remain the same during all the following years. Also assume there is no investment production delay, which means that as soon as the investment is done, the road is available for use. The additional revenue gained from serving an additional demand of 1000 AADT, and 2000 AADT are $100k and $200k per year, respectively. The interest rate for this example is 10 percent. The additional revenue paths in case the investment takes place are illustrated in Figure 13.

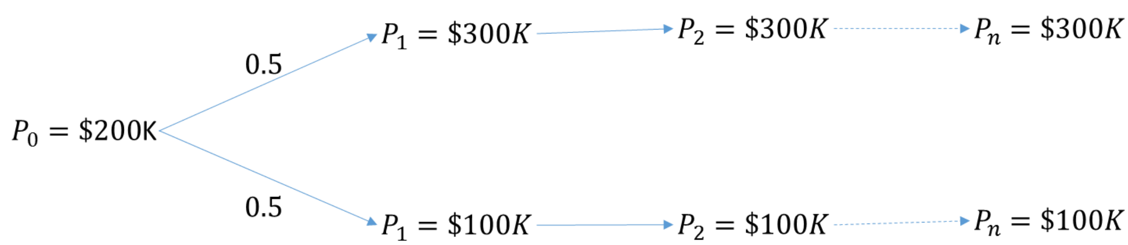

Figure 13 Price paths for deferment option

Performing Discounted Cash Flow (DCF) analysis for this example at time 0 will lead to a positive Net Present Value ($NPV = -I + P_0 + \sum_{t=1}^{\infty} \frac{100k}{(1+r)^t} = \$600k$).

Based on traditional DCF analysis, the positive NPV implies that the investment is worthy and should take place. This positive NPV is a result of the upward path in which the additional revenue is $200k annually. If the possibility and the right exist to allow deferring the investment for one year, until the definite demand and hence additional revenue can be observed and the decision would be only to invest if the demand is the better case, the NPV of such an option to wait one year is calculated as follows:

$$NPV = -\frac{500k}{(1+r)} \times q + \left(\sum_{t=1}^{\infty} \frac{200k}{(1+r)^t}\right) \times q \approx \$773k.$$

The NPV shows that waiting until uncertainties become certain is more valuable than investing right away. The increase in the NPV, $173k, is the value of the option to defer. Adding the option to defer adds more flexibility which has value. This example also emphasizes that performing only traditional DCF analysis is not enough, and it points out one of the downsides of solely performing DCF analysis and using expectations and point estimates. Many other examples about investment under uncertainties and real options are available in (Dixit & Pindyck, 1994). The previous example also is an illustration of binomial lattice analysis which is a popular method for valuing real options. In general, real options are typically valued using three methods:

1- Black-Scholes formula

2- Binomial/Trinomial lattice methods

3- Monte Carlo simulation

4.2.1 Black-Scholes option pricing model

The Black-Scholes formula was first derived in 1973 using delta hedging to value financial European options (Black & Scholes, 1973). This formula was derived with the assumptions of having no dividend yields for the stock, having no arbitrage, and having stock prices that are log-normally distributed; hence, the returns on the stocks are normally distributed. The formula derived by the authors is applicable to European options.

4.2.2 Binomial lattice framework for modeling uncertainties

The Binomial lattice method is based on the diffusion of possible outcomes of an uncertain matter. Based on it, after each time-step, the uncertain element can move two ways: upward or downward.

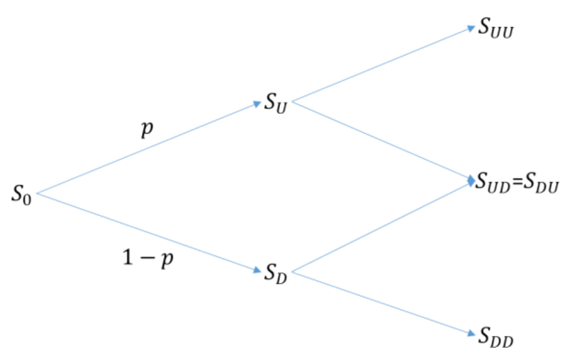

Figure 14 Illustration of a binomial lattice framework

Figure 14 shows the diffusion for two periods. During the first period, the uncertain parameter value either moves up to S_U with a risk-neutral probability p, or moves down to S_D with probability (1-p). At the next time step, the uncertainty could be S_{UU}, S_{UD}, S_{DU}, or S_{DD}. If S_{DU}=S_{UD}, the binomial tree is called a *recombining tree*. The value of the uncertain parameter at each step t can be calculated using equation (4-1).

$$S_t = S_0 \times U^n \times D^{(t-n)}, \qquad (4\text{-}1)$$

where U is the size of an upward movement and D is the size of a downward movement for each step, n is the number of upward movements until step t, and S_0 is the value of the uncertain parameter at step 0. In the general binomial model it is assumed that p, D and U are constant during all periods and depend on the uncertainty of the uncertain article. After the uncertainty fluctuation is modeled using a binomial lattice, a dynamic approach is undertaken to find the best decision. In this approach, at each time step the best option is explored contingent on the outcomes of the decision taken at its previous time stage. The binomial lattice approach can be applied to any type of option.

4.2.3 Monte-Carlo simulation for valuing real options

Monte Carlo simulation is popular for valuing all types of options. The phased investment problem can be modelled as an option to expand. In terms of type of auction based on possible execution times, it is an exotic option because the decision regarding the expansion is only made at discrete times (i.e. at the beginning of each year). Since the option is exotic, Black-Scholes cannot be used to find its value. Using Monte Carlo simulation could be computationally burdensome, especially if a high degree of reliability is desired. The simulation outcomes could be based on the probabilities from a binomial/trinomial lattice framework. The next section, focuses on briefly describing some of the research done related to real options and infrastructures, in particular transportation infrastructures.

4.3 Additional Literature Review on Real Option Analysis in Infrastructure evaluation

Real options have gained attention from the researchers in infrastructures' valuation. Much of this research surrounds Public Private Partnership (PPP) projects in which the public transfers the right and risks of maintaining and operating an infrastructure through a contract. Thus, the risk of the uncertainties, such as the uncertainty in demand, is transferred to the private party. In return, the public provides risk-sharing mechanisms or gives incentives to the private party.

Using real options, (Cui, et al., 2004) introduce an alternative to the conventional warranty clauses stated inside contracts. This alternative transfers the responsibility of maintaining highways whenever certain thresholds are met, and it typically gives contractors more flexibility in selecting the methods and materials for construction. They introduce a warranty option that provides the right to purchase the warranty if certain conditions occur at the end of the construction. They also introduce a method to compare bids based on the different possible warranties firms may provide in their bid proposals. They show that this optional warranty has more value compared to the conventional warranty.

In another study, (Cui, et al., 2008) a binomial lattice model is used to value a ceiling option available in a pavement maintenance warranty clause.

(Brandai, et al., 2012) reviewed the PPP agreement made in Sao Paulo, Brazil, and they reviewed the incentives the government provided for the concessionaire so they would invest in the project. Some of the incentives were:

a) Financial subsidies,

b) Partial exchange-rate guarantees, and

c) Minimum demand guarantees (MDG).

They calculated the risk for the concessionaire as the probability of reaching a negative NPV and showed how subsidies and MDG decreases the risks. Based on the total expected costs of these risk-sharing mechanisms, they mention how a good combination of MDG and subsidies can help reduce the risk for the concessionaire. They finally conclude that the risk of the concessionaire decreases as the portion of MDG to subsidy increases.

(Park, et al., 2013) use real-option valuation for valuing underground water and sewer systems. Unlike toll lanes, their sources of uncertainty are not from uncertain demand. Instead, they are from uncertain Operation and Maintenance costs.

(Cruz & Marques, Submitted 2013) discuss different types of uncertainties associated with PPP projects. They put emphasis on cost overruns, demand forecasting and capital costs. In their paper, they evaluate the options related to demand uncertainty, which has to do with capacity optimal usage. They apply their framework on a healthcare case study. The case study is a PPP arrangement for constructing and operating a Hospital. The concessionaire gets paid based on the number of patients served. They consider two values for the two types of treatments: (1) inpatient and (2) ambulatory. The demands of these types of treatments are uncertain, which leads to two sources of uncertainties. Monte Carlo sampling technique is used to calculate the expected option value. They conclude that adding flexibility inside the contracts, as expected, increases the value of the project whenever uncertainty is present.

(Brandao, et al., 2005) elaborate and discuss the benefits of using binomial decision trees with risk-neutral probabilities over using binomial lattices with risk-adjusted probabilities. Their argument is that risk-neutral probabilities eliminate the need for creating a replicating portfolio at each time step. Instead of the common practice of having uncertainties in the value of the project, they assume that the cash flows, which are used to value a project, have uncertainties. They do their modeling in three steps. In Step One, using risk-adjusted probabilities and using the expected cash flows, the value of the project is calculated. Step Two focuses on finding the standard deviation of the returns of the project by running a Monte Carlo simulation. In the Third Step, a binomial lattice is constructed using the standard deviation from Step Two and the project's initial value from Step One.

(Kruger, 2012) analyzes the expansion of a two-lane road in Sweden using a binomial lattice. He assumes the only source of uncertainty is in traffic demand.

4.4 Modeling Approach

The mathematical model presented in this chapter shares most of the assumptions stated in the deterministic model. The main difference is in the method used for dealing with uncertainties. We modify the deterministic model and provide a stochastic optimization model. A scenario based multi-phase optimization approach is undertaken.

4.4.1 Scenario generation for uncertainties

To model uncertainties and some of the possible outcomes of the uncertain demand we build scenarios. The scenarios are built based on the trinomial lattice trees used in real option valuation. The trinomial lattices are for estimating the demand for each one year period in the future. We then calculate the quantity of demand served that is equal to the minimum of supply provided at each time and the demand at that time based on the number of phases built during each phase. Based on the demand served we can then calculate the overall profit of each scenario. However, we also need to calculate the probability of each scenario. This probability is calculated using the risk-neutral probabilities of an up-ward move, neutral move, and down-ward move used for generating the trinomial lattice. Figure 15 exhibits the possible outcomes when we have 2 time steps after the current time step (T=3). In this case we have $3^{3-1} = 9$ possible scenarios.

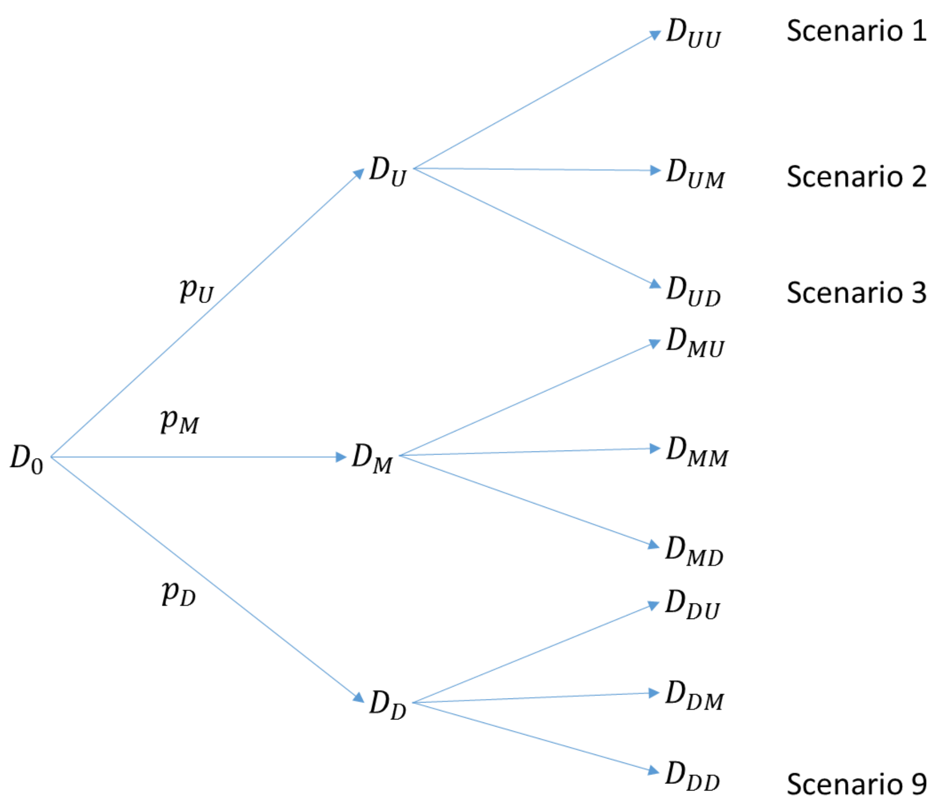

Figure 15 Scenario generation for demand (trinomial lattice)

4.4.2 Multi-phase stochastic optimization

To solve the problem in the presence of uncertainty we model the problem as a multi-stage stochastic optimization problem. This modelling is scenario based. Each scenario is based on the outcomes of the trinomial lattice at each of the stages as explained in the previous subsection.

The general format for a multi-stage scenario based stochastic optimization objective function is shown in (4-2). In multi-stage optimization problems, the most important decision variables are those which are related to the first stage. These decisions are commonly referred to the "here and now" decisions. The other decisions that are made in future times are known as "recourse decisions" and they depend on future scenarios. In (4-2), for example, variables x are here and now decisions and variables y that depend on scenarios, ω, are future recourse decisions.

$$\max z = f(x) + Exp[g(x, y_\omega)]$$

$$h(x, y_\omega) \leq A$$

$$q(x, y_\omega) = B$$

(4-2)

The objective function value is usually the expected value of all scenarios.

4.5 Stochastic Single-Project Phased-investment Problem (SSPP) MIP

The multi-phase scenario based stochastic optimization problem is described in this section.

4.5.1 SSPP parameters and variables

The main here and now decision variable is the infrastructure being prepared, ub, and the main recourse decision variable is the implementation of phases, $x_{i,j}^{t,\omega}$. All the variables and parameters used in the SSPP model are summarized in Table 13.

Table 13 variables and parameters used in SSPP

Variables	
ub	Number of phases selected for implementation
Variables regarding infrastructure	
it_{ub}	Construction duration for infrastructure based on actual number of phases to be implemented
$cost_{ub}$	Construction cost for the infrastructure required for implementing ub phases.
Variables regarding phases	
$x_{ij}^{t,\omega}$	Binary variable that equals 1 if phases i through j start their implementation at time t in scenario ω
Other main variables	
b_t^{ω}	Available budget at the beginning of each time t in scenario ω
n_t^{ω}	Number of phases that have already been implemented or are being implemented at time t in scenario ω
Variables used for linearization	
$cn_{i,t}^{\omega}$	Binary representation of number of phases that have already been implemented or are being implemented at time t for linearization for scenario ω
nc_i	Binary variable for linearization of infrastructure cost
Parameters	
T	Planning period
UB	Maximum number of phases of project
$B0$	Initial available budget
R	Risk free interest rate
I	Inflation rate

M_T	Big-M value used for timing
Times	Set of time periods = 0 .. T
Phases	Set of phases = 1 .. UB
$P_{i,t}^{\omega}$	Profit gained from first i phases at time t in scenario ω
IT_{ub}	Preparation (construction) time for infrastructure of ub phases
$ICOST_{ub}$	Preparation (construction) cost for infrastructure of ub phases
$PT_{i,j}$	Duration required for implementation of each phase $i-j$ when phases $i-j$ are implemented together
$PCOST_{i,j}$	Construction cost of each phases between i,j if phases $i-j$ are being done together
$OCPA_{i,t}$	Operation cost at time t when Number of phases that have already been implemented or are being implemented at time t is i
P_{ω}	Probability of scenario ω

4.5.2 MIP formulation for SSPP

4.5.2.1 Objective Function

The objective function is maximizing the Expected Net Present Value (ENPV) of the portfolio that is being built based on our investments. The objective function is presented in (4-3). It contains the budget at the beginning of the last time period, T, and the profits gained during the last time period and thereon, and the costs incurred during the last time period and thereon. All the costs are subject to inflation.

$$\max z = \sum_{\omega} P_{\omega} \times \frac{B_T^{\omega} + \sum_i P_{i,T}^{\omega} \times cn_{i,T}^{\omega} - \left(\sum_i \sum_j (j-i+1) \times x_{i,j}^{t,\omega} \times PCOST_{i,j} + \sum_i OCPA_{i,T} \times cn_{i,T}^{\omega}\right) \times (1+I)^T}{(1+R)^T} =$$

$$E\left[\frac{B_T^{\omega} + \sum_i P_{i,T}^{\omega} \times cn_{i,T}^{\omega} - \left(\sum_i \sum_j (j-i+1) \times x_{i,j}^{t,\omega} \times PCOST_{i,j} + \sum_i OCPA_{i,T} \times cn_{i,T}^{\omega}\right) \times (1+I)^T}{(1+R)^T}\right] \quad (4\text{-}3)$$

The constraints of the problem are categorized into infrastructural constraints, phase constraints, scheduling constraints, and budgetary constraints.

4.5.2.2 Infrastructural constraints (SSPP)

The infrastructural constraints are listed below:

$$n_t^{\omega} \leq ub \qquad \forall t, \omega \tag{4-4}$$

$$cost_{ub} = \sum_i ICOST_i \times nc_i \tag{4-5}$$

$$it_{ub} = \sum_i IT_i \times nc_i \tag{4-6}$$

$$ub = \sum_i i \times nc_i \tag{4-7}$$

$$\sum_i nc_i \leq 1 \tag{4-8}$$

Constraints (4-4) limit the total number of phases that are implemented at each time for each scenario to the maximum invested infrastructure. Constraints (4-5) and (4-6) are for calculating the infrastructure cost and install duration. For them being linear, we need to express the number of phases selected for implementation, ub, using binary variables. This is done using constraints (4-7) and (4-8).

4.5.2.3 Phase related constraints (SSPP)

Constraints (4-9) – (4-14) are the phase related constraints:

$$n_0^{\omega} = \sum_i \sum_j (j - i + 1) \times x_{i,j}^{0,\omega} \qquad \forall \omega \tag{4-9}$$

$$n_t^{\omega} = n_{t-1}^{\omega} + \sum_i \sum_j (j - i + 1) \times x_{i,j}^{t,\omega} \qquad \forall t \geq 1, \omega \tag{4-10}$$

$$n_t^{\omega} = \sum_i i \times cn_{i,t}^{\omega} \qquad \forall t, \omega \tag{4-11}$$

$$\sum_i cn_{i,t}^{\omega} \leq 1 \qquad \forall t, \omega \tag{4-12}$$

$$\sum_{j \geq i} x_{i,j}^{t,\omega} \leq \sum_{l \leq i-1} \sum_{t' \leq t} x_{l,i-1}^{t',\omega} \qquad \forall i \in \{2, \dots, UB\}, t, \omega \tag{4-13}$$

$$\sum_t \sum_{i \leq l} \sum_{j \geq l} x_{i,j}^{t,\omega} \leq 1 \qquad \forall l \in \{1, \dots, UB\}, \omega \tag{4-14}$$

Constraints (4-9) and (4-10) are for calculating the number of phases implemented/being implemented at different times for each scenario. Constraints (4-11) and (4-12) are for representing the number of phases that are implemented/are being implemented using binary variables so that we would have linear constraints when calculating the different costs and times for phases for each scenario. Constraints (4-13) prevent implementation of succeeding phases prior to the implementation of phases that are preceding them for each scenario. Constraints (4-14) prevent the assignment of a phase to two different groups of phases for all scenarios.

4.5.2.4 Scheduling related constraints (SSPP)

Constraints (4-15) – (4-17) are the schedule related constraints:

$$it_{ub} \leq \sum_j \sum_t t \times x_{1,j}^{t,\omega} + M_T \times \left(1 - \sum_j \sum_t x_{1,j}^{t,\omega}\right) \qquad \forall \omega \tag{4-15}$$

$$t \times x_{i,j}^{t,\omega} + (j - i + 1) \times PT_{i,j} \times x_{i,j}^{t,\omega} \leq \sum_{l \geq j+1} \sum_{t' \geq t} t' \times x_{j+1,l}^{t',\omega} + M_T \times \Big(1 - \sum_{l \geq j+1} \sum_{t' \geq t} x_{j+1,l}^{t',\omega}\Big) \qquad \forall i, j \in \{2, \dots, UB\} \mid j \geq i,\ t, \omega \tag{4-16}$$

$$\sum_i \sum_{j \geq i} x_{i,j}^{t,\omega} \leq 1 \qquad \forall t, \omega \tag{4-17}$$

Constraints (4-15) assure that for each scenario, the first group of phases are implemented after the infrastructure is completed and successfully implemented. Constraints (4-16) ensure that each phase is implemented after the completion of implementation of its preceding phases for each scenario. Constraints (4-17) prevent multiple groups of phases to start their implementation together for all scenarios.

4.5.2.5 Budget related constraints (SSPP)

Constraints (4-18) – (4-21) are the budget related constraints:

$$b_1^{\omega} = B0 - cost_{ub} - \left(\sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{0,\omega} + \sum_i OCPA_{i,0} \times cn_{i,0}^{\omega}\right) \times (1 + R) + \sum_i P_{i,0}^{\omega} \times cn_{i,0}^{\omega} \qquad \forall \omega \tag{4-18}$$

$$b_t^{\omega} = \left(b_{t-1}^{\omega} - \left(\sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{t-1,\omega} + \sum_i OCPA_{i,t-1} \times cn_{i,t-1}^{\omega}\right) \times (1 + I)^{t-1}\right) \times (1 + R) + \sum_i P_{i,t-1}^{\omega} \times cn_{i,t-1}^{\omega} \qquad \forall t \in \{2, \dots, T\}, \omega \tag{4-19}$$

$$cost_{ub} + \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{0,\omega} + \sum_i OCPA_{i0} \times cn_{i,0}^{\omega} \leq B0 \forall \omega \tag{4-20}$$

$$\left(\sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{t-1,\omega} + \sum_i OCPA_{i,t-1} \times cn_{i,t-1}^{\omega}\right) \times (1 + I)^{t-1} \leq b_t^{\omega} \qquad \forall t \geq 1, \omega \tag{4-21}$$

Constraints (4-18) and (4-19) are the updates on the available budget at the beginning of each time period. For each scenario, the available budget at the beginning of each time period is equal to remaining budget from the previous time period in the current period's value (incorporating time value of money) plus the profits earned at the end of the last period as a result of implemented phases in the previous period. Constraints (4-20) and (4-21) are the budget limitations during different times for each scenario.

4.5.2.6 Variable domain constraints (SSPP)

Finally, constraints (4-22) and (4-23) are the variable domain constraints:

$$ub, it_{ub}, cost_{ub}, b_t^{\omega}, n_t^{\omega} \geq 0 \tag{4-22}$$

$$x_{i,j}^{t,\omega}, cn_{i,t}^{\omega}, nc_i \in \{0,1\} \tag{4-23}$$

4.5.2.7 Non-anticipatively constraints (SSPP)

These constraints are for restricting the future decisions that are taken at future steps to have the same value regardless of the scenario.

$$x_{i,j}^{t,\omega_1} = x_{i,j}^{t,\omega_2} \qquad \forall \omega_1, \omega_2, t, (i,j) \tag{4-24}$$

4.5.2.8 The complete MIP model for SSPP

The complete multi-phase mixed integer programming model is summarized in the following:

$$\max z = \sum_{\omega} P_{\omega} \times \frac{B_T^{\omega} + \sum_i P_{i,T}^{\omega} \times cn_{i,T}^{\omega} - \left(\sum_i \sum_j (j-i+1) \times x_{i,j}^{t,\omega} \times PCOST_{i,j} + \sum_i OCPA_{i,T} \times cn_{i,T}^{\omega}\right) \times (1+I)^T}{(1+R)^T} =$$

$$E\left[\frac{B_T^{\omega} + \sum_i P_{i,T}^{\omega} \times cn_{i,T}^{\omega} - \left(\sum_i \sum_j (j-i+1) \times x_{i,j}^{t,\omega} \times PCOST_{i,j} + \sum_i OCPA_{i,T} \times cn_{i,T}^{\omega}\right) \times (1+I)^T}{(1+R)^T}\right] \tag{4-3}$$

Subject to:

$$n_t^{\omega} \leq ub \qquad \forall t, \omega \tag{4-4}$$

$$cost_{ub} = \sum_i ICOST_i \times nc_i \tag{4-5}$$

$$it_{ub} = \sum_i IT_i \times nc_i \tag{4-6}$$

$$ub = \sum_i i \times nc_i \tag{4-7}$$

$$\sum_i nc_i \leq 1 \tag{4-8}$$

$$n_0^\omega = \sum_i \sum_j (j - i + 1) \times x_{i,j}^{0,\omega} \qquad \forall\omega \tag{4-9}$$

$$n_t^\omega = n_{t-1}^\omega + \sum_i \sum_j (j - i + 1) \times x_{i,j}^{t,\omega} \qquad \forall t \geq 1, \omega \tag{4-10}$$

$$n_t^\omega = \sum_i i \times cn_{i,t}^\omega \qquad \forall t, \omega \tag{4-11}$$

$$\sum_i cn_{i,t}^\omega \leq 1 \qquad \forall t, \omega \tag{4-12}$$

$$\sum_{j \geq i} x_{i,j}^{t,\omega} \leq \sum_{l \leq i-1} \sum_{t' \leq t} x_{l,i-1}^{t',\omega} \qquad \forall i \in \{2, \dots, UB\}, t, \omega \tag{4-13}$$

$$\sum_t \sum_{i \leq l} \sum_{j \geq l} x_{i,j}^{t,\omega} \leq 1 \qquad \forall l \in \{1, \dots, UB\}, \omega \tag{4-14}$$

$$it_{ub} \leq \sum_j \sum_t t \times x_{1,j}^{t,\omega} + M_T \times \left(1 - \sum_j \sum_t x_{1,j}^{t,\omega}\right) \qquad \forall\omega \tag{4-15}$$

$$t \times x_{i,j}^{t,\omega} + (j - i + 1) \times PT_{i,j} \times x_{i,j}^{t,\omega} \leq \sum_{l \geq j+1} \sum_{t' \geq t} t' \times x_{j+1,l}^{t',\omega} + M_T \times \Big(1 - \sum_{l \geq j+1} \sum_{t' \geq t} x_{j+1,l}^{t',\omega}\Big) \qquad \forall i, j \in \{2, \dots, UB\} \mid j \geq i,\ t, \omega \tag{4-16}$$

$$\sum_i \sum_{j \geq i} x_{i,j}^{t,\omega} \leq 1 \qquad \forall t, \omega \tag{4-17}$$

$$b_1^\omega = B0 - cost_{ub} - \left(\sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{0,\omega} + \sum_i OCPA_{i,0} \times cn_{i,0}^\omega\right) \times (1 + R) + \sum_i P_{i,0}^\omega \times cn_{i,0}^\omega \tag{4-18}$$

$$b_t^{\omega} = \left(b_{t-1}^{\omega} - \left(\sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{t-1,\omega} + \sum_i OCPA_{i,t-1} \times cn_{i,t-1}^{\omega}\right) \times (1 + I)^{t-1}\right) \times (1 + R) + \sum_i P_{i,t-1}^{\omega} \times cn_{i,t-1}^{\omega} \qquad \forall t \in \{2, \dots, T\}, \omega \tag{4-19}$$

$$cost_{ub} + \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{0,\omega} + \sum_i OCPA_{i,0} \times cn_{i,0}^{\omega} \leq B0 \tag{4-20}$$

$$\left(\sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{t-1,\omega} + \sum_i OCPA_{i,t-1} \times cn_{i,t-1}^{\omega}\right) \times (1 + I)^{t-1} \leq b_t^{\omega} \qquad \forall t \geq 1, \omega \tag{4-21}$$

$$ub, it_{ub}, cost_{ub}, b_t^{\omega}, n_t^{\omega} \geq 0 \tag{4-22}$$

$$x_{i,j}^{t,\omega}, cn_{i,t}^{\omega}, nc_i \in \{0,1\} \tag{4-23}$$

$$x_{i,j}^{t,\omega_1} = x_{i,j}^{t,\omega_2} \qquad \forall \omega_1, \omega_2, t, (i,j) \tag{4-24}$$

4.6 Solution Methods for Single-Project Phased-investment Problem (SSPP)

We are interested in the outcomes of the here and now decision variables standing at the current time. Throughout the remainder of this section, the variable that we are interested in its value is the variable that determines the size of the infrastructure, ub. ub is the variable that states how many of the phases would be carried out during the planning period. Based on the value of ub we prepare the infrastructure for carrying out up to ub phases during the planning period. Note that although we have prepared the infrastructure for ub phases, it might turn out that for some scenarios, some recourse actions are abandoning the project at some stage or carrying out less phases.

We can solve the problem and find ub using commercial solvers such as XPRESS for problems which do not have many scenarios. The commercial solvers usefulness suddenly drops as the size of the problem increases.

For the parking example, in the presence of 15 time periods and the trinomial tree platform for the diffusion of uncertainty, we have a total of $3^{14} \approx 4.78e6$ scenarios. Since many of the variables and constraints are scenario related, the complete multi-stage mathematical model will be very large and solving such a large MIP using merely commercial solvers is unrealistic. Two approaches are considered for such cases in this research. The first approach is to decompose the problem by solving the problem for each scenario independently. The second approach is using simulation to compare the bounds on the expected objective function values for different values of the here and now decision. Both approaches are explained in more detail in the following subsections.

4.6.1 Solving the problem for all scenarios

A popular method for solving multi-stage scenario based stochastic optimization problems is to solve the deterministic problem for each single scenario. The most common solution to the first stage "here and now" decision variable could then be picked as the solution to proceed upon.

This approach has its benefits. One of the most important benefits it has is its independence on probabilities of the scenarios. All scenarios have equal value in this method. This is valuable especially in cases that we cannot calculate or predict the probabilities accurately. It is also profitable when we have many scenarios and as a result the probability of most of those scenarios are practically zero (less than machine epsilon). The running time of each scenario is also very small in comparison to the running time of the overall stochastic problem since the problem for each scenario is a deterministic problem with much fewer constraints and variables.

The downside of this approach is the curse of dimensionality. If we have many scenarios, even though solving each one of them might be very fast, the overall time required to solve all of the scenarios may be very large. This causes this method to look impractical. To fix this impracticality, many reduce the number of scenarios. In doing this, most of the scenario generation and reduction studies try to minimize what is called a distance function. This distance function is the difference between the original probability distribution of the original scenario tree and the probability distribution of the reduced scenario tree. For example, the distance between two scenarios ω_i and ω_j can be calculated using the norm of their differences ($\left|\omega_i - \omega_j\right|^n$). The overall distance for example could be calculated as (4-25).

$$\sum_{k\in RemainingScenarios} P_k \min_{j\notin RemaininScenarios} \left|\omega_k - \omega_j\right|^n \tag{4-25}$$

The reduced scenario tree has a smaller probability base (fewer scenarios). When the size of the base of the reduced scenario tree is given (the number of scenarios to be preserved are given) or equivalently the number of scenarios that should be deleted are given, the problem is relatively solvable. However when the sizes are not given as an input, the problem becomes difficult and different heuristics have been proposed to solve the problem. The most famous among the heuristics are the Forward and Backward methods. In the Forward method, we select scenarios to preserve iteratively. In the backward method, we iteratively select scenarios to delete. The interested reader in the theory of these methods is referred to (Heitsch & Römisch, 2003) and (Dupačová, et al., 2003). The forward and backward algorithms have been very popular in stochastic optimization studies (Xing, et al., 2006), (Razali & Hashim, 2010), (Growe-Kuska, et al., 2003), (Sharma, et al., 2013), (Siahkali & Vakilian, 2010), (Pedrasa, et al., 2011), (Feng & Ryan, 2013), (Park, et al., 2016).

Another approach for scenario reduction is clustering. (Beraldi & Bruni, 2014) use a cluster based approach. They cluster the scenarios from the decision tree. Then select one representative scenario per cluster to remain in the reduced tree.

Most, if not all of these methods, mainly rely on the probability distributions and they eliminate scenarios with some loss. This is while solving the problem for all scenarios (without any loss in scenarios) still has some benefits especially when robust optimization is important. In addition, all of them might eliminate some scenarios for which the problems' solution (ub value) might be different from the solution of all the remaining reduced number of scenarios. To overcome these shortcoming we present a method for solving all scenarios by "elimination via guessing". This method, that we call it the Solve-Search-Delete (SSD algorithm), is applicable to the cases in which the value of the "here and now" decision is monotonically related to the outcome of the scenario.

For example, in the phased investment example, we have to show that when all input parameters are fixed, for a better scenario, the problem's here and now variable, ub, is always larger or equal to the that from the worse scenario. In simpler words, when the outcomes are better, we never build a smaller infrastructure than when the outcomes are worst.

4.6.1 Preliminaries for the SSD algorithm

Before we proceed with explaining the SSD algorithm, we will discuss the main reason this algorithm works.

Theorem: We could "find" the optimal here and now decision variables' value, ub value, of a scenario ω, ub^{ω}, without solving the scenario if the optimal ub values for two scenarios that are better, p, and worse, n, than ω are equal and the monotonicity condition is satisfied: If $ub^{p} = ub^{n} \Rightarrow ub^{\omega} = ub^{p}$.

Before we prove the theorem, it is important to clearly define what we call a "better" and "worse" scenario. We say scenario a is better than scenario b if the value of the uncertain parameter during all stages is more favorable or equally favorable in a in comparison to b. This means that if the uncertain parameter is demand which is positive, the demand of scenario a should always be larger or equal to the demand of scenario b. Figure 16 illustrates the case where the green scenario is better than all scenarios. The red and dark blue scenarios are not comparable since the value of the uncertain parameter crosses during a stage in these two scenarios. It is easy to show that the optimal objective function value for the better scenario is better since all other parameters except for the uncertain parameter are the same.

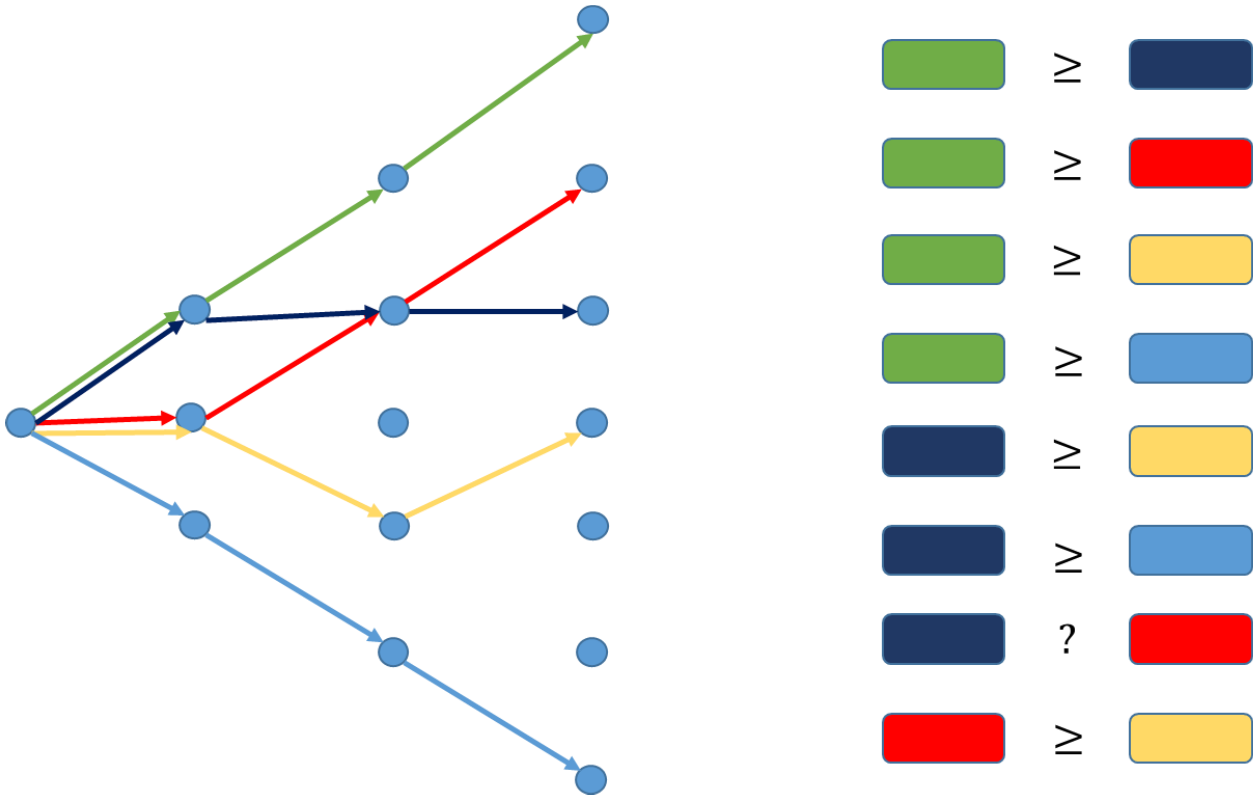

Figure 16 Comparison among different scenarios

Note that this declaration of better and worse scenarios is very robust. One might say that the dark blue scenario has a better objective function value than the red because it's value is higher in two stages (in comparison to one). The net area between these two scenarios is a driving factor in this claim. However, we cannot clearly say that because, due to limitations in supply, during the first two stages, we might not be able to serve a demand beyond the demand of the red scenario. If the supply increases during the last stage however, we can satisfy more overall demand for the red scenario.

Now that we have declared what is a better and a worse scenario, let's prove the theorem. We prove the theorem by proving parts a and b listed below.

a) The optimal ub of a worse scenario is always a feasible solution to the better scenario: This is very trivial since we can always set the value of the decision variables of the better scenario equal to the ones from the worse scenario and get a feasible solution. The ub here acts as a lower bound.

b) The optimal ub solution of a better scenario is an upper bound to the ub: This is a direct fact from the monotonicity requirement.

Once we have a lower bound and an upper-bound that are equal, the solution is going to be equal to any of these bounds.

The idea behind the SSD algorithm is to progressively solve scenarios and compare the solved scenarios. Once two solved scenarios have the same here and now decision variable, ub, and are comparable (one scenario is better than the other scenario), using the theorem, delete the scenarios in-between the comparable scenarios and set their ub to equal the ub of the comparable scenarios. As we proceed, we are expanding the list of already solved scenarios and therefore, more compatible scenarios would be found. This process would continue until there are no more remaining scenarios that we should solve as they are either deleted by comparison or solved. After this step the optimal ub for all scenarios is known.

4.6.2 The Solve-Search-Delete (SSD) algorithm

The SSD algorithm can be broken into smaller pieces. The first piece is initialization. The other major components are the search and delete components. Each of these parts are explained in more detail in the following subsections.

4.6.2.1 Initialization of SSD algorithm

During this step that is only executed once at the beginning of the algorithm, we build the required search arrays and populate them with some non-arbitrary values. The initializations step, executes all the steps except for the search step once. The pseudo-code for this part is:

- ❑ $Remaining\_scenarios$ ← All Scenarios
- ❑ Create an empty array for each possible value of ub: $ub\_search[|ub|]$
- ❑ Solve the MIP for the best scenario: 1
    - ❑ Add the scenario to its corresponding ub_search value based on its ub value:
        - ❑ $ub\_search[ub^1]$ $.append(1)$.
    - ❑ Remove scenario 1 from the list of available projects:
        - ❑ $Remaining\_scenarios$ $.pop(1)$
- ❑ Solve the MIP for the worst scenario: N
    - ❑ Add the scenario to its corresponding ub_search value based on its ub value:
        - ❑ $ub\_search[ub^N]$ $.append(N)$.
    - ❑ Remove scenario N from the list of available projects:
- ❑ $Remaining\_scenarios$ $.pop(N)$

The initialization starts by putting all the scenarios in the list of remaining scenarios. Then we create an array that would be used for searching for compatible scenario pairs. This array has $|ub|$ cells, where $|ub|$ is the number of possible outcomes for the here and now decision variable, ub. In the case of the parking example, this value would be $UB + 1$. Each cell of this array will store the IDs for the solved scenarios that their "here and now decision variable", ub, values are equal to ub_i.

Next, we populate the built array. During initialization, this is done wisely. Instead of selecting random scenarios, we pick the best (ID=1) and worst scenario (ID=N). We solve the deterministic MIP problem for each scenario ω using Xpress. Based on the outcome of its here and now decision variable, ub^ω, we assign it to its appropriate cell in ub_search. Then since these scenarios are solved, we remove them from the list of remaining scenarios. After the initializations step we have $N-2$ remaining scenarios.

4.6.2.2 Solve and Search steps of the SSD algorithm

Now that we have at least a minimum basis for searching (2 scenarios added during initialization) and also the we know the bounds of the values for ub from the initialization step, we can illustrate the solve and search steps. In the solve step, we randomly select a scenario ω from the list of remaining scenarios. We solve the deterministic model using the parameters of this scenario with Xpress. Say, the optimal here and now decision value for this scenario is ub^{ω}. We search within the list $ub\_search[ub^{\omega}]$. We see if there is any scenario in that list which it is either better or worse than scenario ω. Recall that the ranking of scenarios were explained earlier. If we find such a compatible scenario, we proceed to the delete step of the algorithm. If not, we simply add the solved scenario to the search list and delete it from the list of remaining projects.

Figure 17 illustrates how the search step of the SSD algorithm works. Assume that until this point in time within the algorithm, $ub\_search[ub_1]$ has been populated with scenario 1 (the one with the dashed line) and $ub\_search[ub_2]$ has been populated with scenarios α and β (in this order). Scenarios α and β are those which are solid and darker but not bolded. Also, assume that we randomly have selected the bolded scenario ω from the list of remaining projects and have solved the deterministic model for this scenario and its optimal here and now decision value is ub_2. To find comparable pairs, we search within $ub\_search[ub_2]$. We start with comparing scenario ω with scenario α. These two scenarios are non-comparable because they cross each other in a stage.

Then we proceed to comparing scenario ω with scenario β. These two scenarios are comparable and scenario ω is the better scenario. Here we proceed to the deleting step of the algorithm.

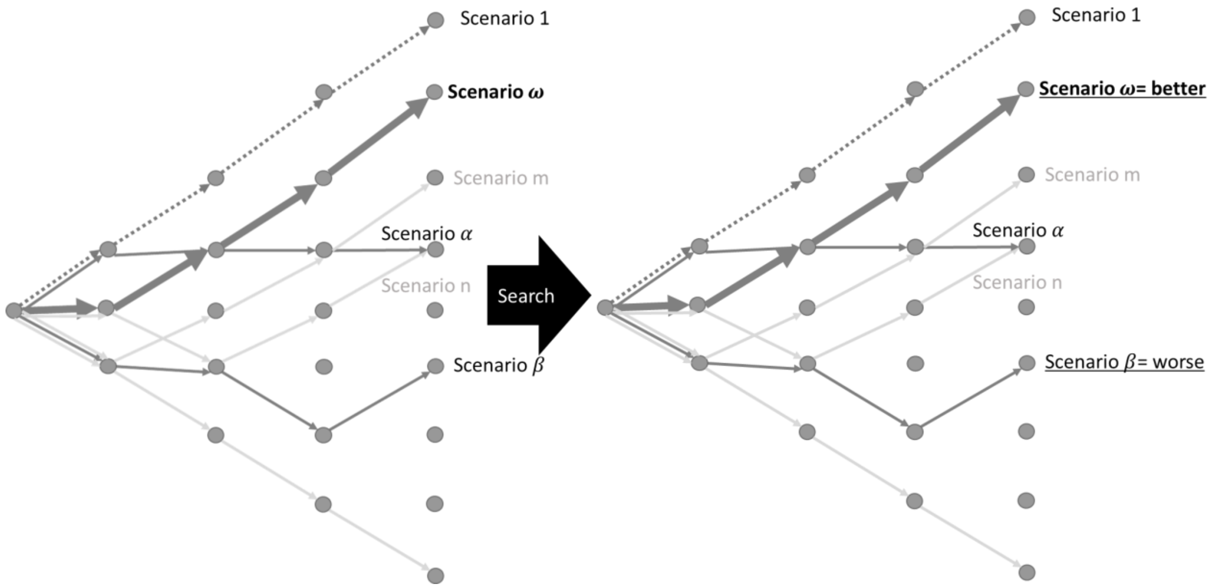

Figure 17 Search of "guessing" Algorithm Illustrative Example

4.6.2.3 Delete step of the SSD algorithm

In this step, if we have found a comparison in the previous step, we delete all the scenarios that are sandwiched by the better and worse scenario found in the prior step. Figure 18 Illustrates this. During this step, scenarios n and m are sandwiched by scenario ω and β and therefore are deleted from the list of remaining scenarios. Their $ub^n = ub^m$ is equal to $ub_2 = ub^\omega$.

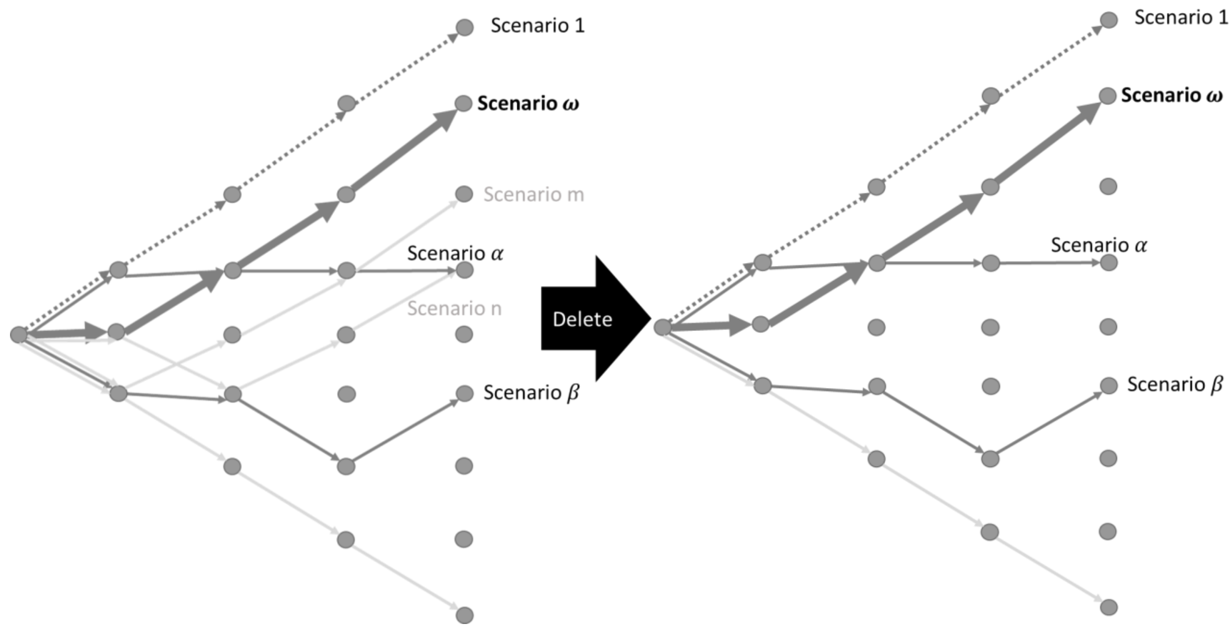

Figure 18 Deleting step in "guessing" algorithm illustrative example

After this step we proceed by adding scenario ω to the search list and also deleting it from the list of remaining scenarios. This process of solving-searching and deleting is continued until no more scenarios remain. The overall algorithm minus the initialization step is summarized below:

- ❑ While size($remaining\_scenarios$) ≥ 1:
  - ❑ Select a scenario, ω randomly from $remaining\_scenarios$
  - ❑ Solve deterministic MIP for scenario ω and find ub^{ω}
  - ❑ Search for compatible scenario c in $ub\_search[ub^{\omega}]$,
  - ❑ If compatible pair found (c exists):
    - ❑ Delete all remaining scenarios in between ω and c.
  - ❑ Add ω to ub_search: $ub\_search[ub^{\omega}].append(\omega)$
  - ❑ Remove ω from list of $remaining\_scenarios$: $remaining\_scenarios.pop(\omega)$

Now that the SSD algorithm is explained, we proceed by describing the second solution method for solving SSPP.

4.6.3 Optimization via Simulation for evaluating different here and now decisions of the SSPP

Solving the SSP problem to its entirety is a very cumbersome and complex as illustrated before. However, by relaxing the non-anticipatively constraints (4-24), we will be able to find good approximation solutions for the here and now decision variable possible values. In order to solve the stochastic problem without neglecting the probabilities of the scenarios, we utilize a random selection (and simulation) scheme. In this scheme, using the probabilities of moving upward, downward, and not changing in each step of the trinomial lattice, we perform a random walk. This random walk is an iteration of one run for the simulation. The purpose of this simulation is to find the best possible outcome for the here and now decision variable, ub. If the possible outcomes of this decision variable is limited, we can look at each different outcome as a policy and use simulation to pick the best policy. The performance measure calculated for comparing the different policies is the bound on the expected objective function value that is very appropriate as it is the closest simple measure we have to the objective function of the SSPP. Since the outcome of the decision variable is fixed for each of these policies, in each iteration of each simulation run, we only have to solve the relaxed deterministic model for each simulation outcome with a fixed here and now decision variable that is very fast and easy using commercial solvers.

4.6.3.1 Simulation algorithm for SSPP

The simulation algorithm for one run is presented below:

- ❑ For different possible (groups of) values of ub; i= 1..UB:
    - ❑ Set $ub = i$

- ❑ Counter = 1
- ❑ Total_obj_func[i] = 0
- ❑ While Counter $\leq$ Max_iterations:
  - ❑ Based on $P_u,\ P_m,\ and\ P_d$ simulate a random walk for the uncertain parameter, ω.
  - ❑ Solve “remaining relaxed MIP” using Xpress:
    - ❑ Update Total_obj_func: Total_obj_func[i] $+= Z^{\omega}$
  - ❑ Counter $+= 1$
- ❑ Avg_obj_func[i] = Total_obj_func[i] / Max_iterations

Note that during the simulation, we are approximating the expectation of the maximization of the relaxed SSPP problem. The relaxed SSPP problem, as mentioned earlier, is the SSPP problem without the nonanticipativity constraints. By relaxing these constraints, and fixing the here and now decision variable in each simulation run, each simulation iteration becomes independent from the other simulation iterations (scenario become independent). As a result, the expectation of the maximum will be equal to the maximum of the expectation. Since the relaxed problem is giving us an upper bound of the non-relaxed “original” SSPP, the performance measure of the simulation is giving us the expected upper bound value for each possible value of the here and now decision variable (4-26).

$$P_M = E[\max z_r] = \max E[z_r] \tag{4-26}$$

In order to achieve certain confidence for the performance measure (upper bound on the expected objective function value) calculated during each simulation run, we need to have enough iterations in each run of the simulation. We can find the number of iterations needed, m, based on the equation for the confidence intervals. For cases that

$m \geq 30$, we can use the following equation for estimating the performance measure for system a in the first run, Y_{a_1}:

$$\overline{Y_{a_1}} \pm z_{\alpha/2} \times \frac{\sigma}{\sqrt{m}} \tag{4-27}$$

where $\alpha = 1 - c.c$ and $c.c.$ is the confidence coefficient, $\overline{Y_{a_1}}$ is the point estimate for this performance measure, and the term $z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$ is known as the margin of error, E. By fixing the value of this parameter, we can find the number of iterations needed using equation (4-28).

$$E = z_{\alpha/2} \times \frac{\sigma}{\sqrt{m}} \quad \rightarrow \quad m = \frac{z_{\frac{\alpha}{2}}^2 \times \sigma^2}{E^2} \tag{4-28}$$

Once we have the estimates for the performance measures during each run of the simulation per policy, we can use those values for comparing the different policies. Generally, when we want to compare two systems a and b using simulation, we look at the average performance measure of the simulation runs. Say Y_{a_i} is the average performance measure for system a at run i of the simulation. Then the overall average of the system after A runs is $\overline{Y_a} = \frac{1}{A}\sum_{i \in \{1,2,\ldots,A\}} Y_{a_i}$. If the overall average performance measure of system b after B runs is $\overline{Y_b}$, we can compare the two systems by looking at where the difference of the performance measures fall with respect to 0. If, $a - b < 0$ this means that $a < b$ and if $a - b > 0$ this means that $a > b$. However if $a - b$ does not fall clearly on one of the sides of zero then we cannot draw any conclusions and we need more iterations so that $a - b$ would finally fall on one of the sides of zero. Typically, we can estimate $a - b$ (the difference between the performance measures of the two systems) with a confidence interval of $100(1-\beta)\%$ using the outputs of the simulation runs and equation (4-29).

$$\overline{Y_a} - \overline{Y_b} \pm t_{1-\frac{\beta}{2},\ \nu} \times S.e.(\overline{Y_a} - \overline{Y_b}) \tag{4-29}$$

where $S.e.(\overline{Y_a} - \overline{Y_b})$ is the standard error of $\overline{Y_a} - \overline{Y_b}$ and ν is the degree of freedom used for finding the critical value of the t-test. When the variances of the runs from Y_a and Y_b are not equal, we can use the following equations to find the standard error and the degree of freedom:

$$S.e.(\overline{Y_a} - \overline{Y_b}) = \sqrt{\frac{S_a^2}{n_a} + \frac{S_b^2}{n_b}} \tag{4-30}$$

$$\nu = round\left(\frac{\left(\frac{S_a^2}{n_a} + \frac{S_b^2}{n_b}\right)^2}{\left[\frac{\left(\frac{S_a^2}{n_a}\right)^2}{n_a - 1}\right] + \left[\frac{\left(\frac{S_b^2}{n_b}\right)^2}{n_b - 1}\right]} \right) \tag{4-31}$$

The overall procedure of comparing two policies is following two sampling events. Assume that event A is that the performance measure calculated per each simulation run is actually incorrect and the actual value for that performance measure falls out of the margin of error plus the point estimate. Event B is that the comparison has an error. We can used Bonferroni's method to find a bound on the overall confidence interval. The probability of at least one of the steps of finding a point estimate for the performance measures per each run and comparing two policies based on their avg. performance is erroneous is $P(A \cup B)$. $P(A \cup B)$ however is bound by $P(A) + P(B)$. So the overall confidence of both steps together would be at least $1 - (P(A) + P(B))$.

4.6.3.2 Common Random Number (CRN) Simulation algorithm for SSPP

We could potentially decrease the number of simulation runs needed using the concept of common random numbers (CRN). CRN is a popular variance reduction technique that is mainly used when we are interested in comparing performance measures of two or more different policies. By using the same random numbers per each simulation iteration of the simulation runs, we are adding dependency between the simulation runs. The covariance will cause the variance to decrease based on equation (4-32).

$$Var(\bar{a} - \bar{b}) = Var(a) + Var(b) - 2Cov(a, b) \tag{4-32}$$

The CRN simulation algorithm is summarized below:

- ❑ Counter = 1
- ❑ Total_ob_func = zeros(UB)
- ❑ While Counter ≤ Max_iterations:
    - ❑ Based on $P_u,\ P_m,\ and\ P_d$ simulate a random walk for the uncertain parameter, ω.
    - ❑ For different possible (groups of) values of ub; i= 4..5:
        - ❑ Set $ub = i$
        - ❑ Solve "remaining MIP" using Xpress:
            - ❑ Update Total_obj_func: Total_obj_func[i] $+=$ Z^{ω}
    - ❑ Counter $+=$ 1
- ❑ Avg_obj_func[i] = Total_obj_func[i] / Max_iterations

4.7 Example and results for Single-Project Phased-investment Problem (SSPP) using solution methods

In this section, we again solve the mentioned parking garage example from (Zhao & Tseng, 2003) . The possible outcomes of the uncertainty (demand) are built based upon a trinomial lattice model. The same way that it was built in (Zhao & Tseng, 2003). During each stage of the uncertain parameter can either increase in value with probability $P_U = 0.288$, not change with probability $P_M = 0.626$, or decrease in value with probability $P_D = 0.086$. The demand at time 0 is 250. If the demand increases at stage 1, it will increase to $e^{\ln(250)+D}$. If it decreases, it will decrease to $e^{\ln(250)-D}$. If it does not change it will remain as $e^{\ln(250)} = 250$. By knowing the number of upward movements (UP), and downward movements (DN) before each stage, we can find the demand during that stage using equation (4-31).

$$Demand = e^{\ln(250)+(UP-DN)\times D} \tag{4-31}$$

where D was calculated to be 0.2087 based on historical data (Zhao & Tseng, 2003). The remaining of the parameters are the same as the ones mentioned in (3.3) unless explicitly mentioned otherwise.

We consider two set of values for parameters ($B0$, I, R), that are the initial investment budget, inflation rate, and risk-free interest rate, respectively. These two set of values are selected such that for one of them, the maximum ub value would be less than the maximum allowed, UB, and in the other one it would be equal. The set of parameter values are: (a) $B0 = 6{,}000$, $I = 5\%$, and $R = 2\%$ and (b) $B0 = 10{,}000$, $I = 2\%$, and $R = 2\%$.

4.7.1 Solving the parking SSPP problem for all scenarios using SSD algorithm

In this example, we can show that the requirement of "better scenarios give higher ub values" for the SSD algorithm is satisfied by looking at the sensitivity of the solution of ub to the initial available budget, $B0$. Note that we are using a proxy that better scenarios are somehow equivalent to bigger initial budgets. This proxy is valid due to the way a better scenario is defined throughout this work. The sensitivity analysis on $B0$ shown in Figure 19 verifies that this condition is satisfied for $R = 2\%, \mathrm{I} = 5\%$. This condition is also true for the other set of parameters ($I = R = 2\%$).

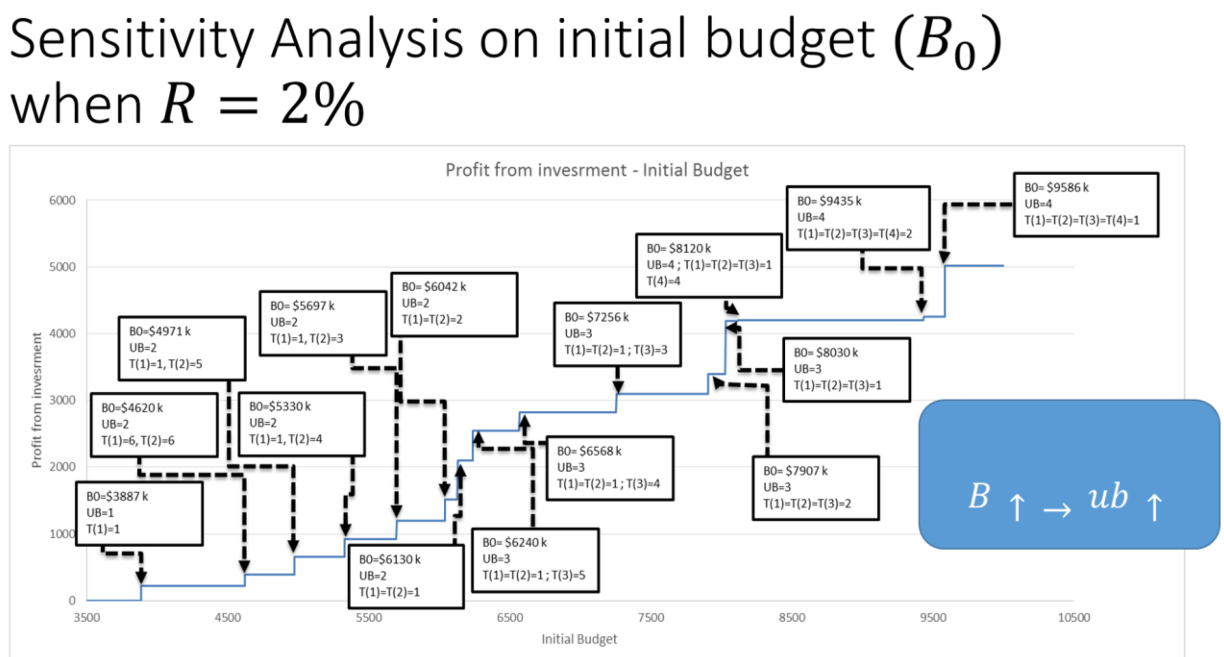

Figure 19 Verifying the correctness of the condition required for the SSD algorithm for the Parking garage example and R=2%

Note that due to the trinomial lattice framework used for modeling the uncertainty in demand (profit), the number of scenarios exponentially increases by increasing the time steps (Stages).

Solving the problem for each scenario independently as a deterministic problem is very time consuming. While the average running time for each scenario is 0.3 seconds, due to having approximately 4.78e6 scenarios, the total time for solving all scenarios would be roughly 400 hours. By keeping this in mind, we use the SSD algorithm to find the optimal here and not decision variable's value for all the scenarios.

For the case that the inputs are, $B0 = 6{,}000$, $I = 5\%$, and $R = 2\%$, the number of scenarios remaining based on time executed is depicted in Figure 22. As it is illustrated, the ub value for all scenarios is solved or found in 2 hours and 6 minutes and 19,334. Simply running the algorithm for 30 minutes reduces the number of remaining scenarios to 340,418, according to Figure 23.

Even in the best scenario, the maximum value for ub was 2 as it is shown in Figure 20. The most common value for ub is 2. As illustrated in Figure 21, most of the ub values are deleted and very few of them are solved. In fact, based on Table 14, on average, for each scenario solved, at least 100 scenarios were deleted. For each scenario with $ub = 2$ solved, on average 1431 scenarios with $ub = 2$ were deleted!

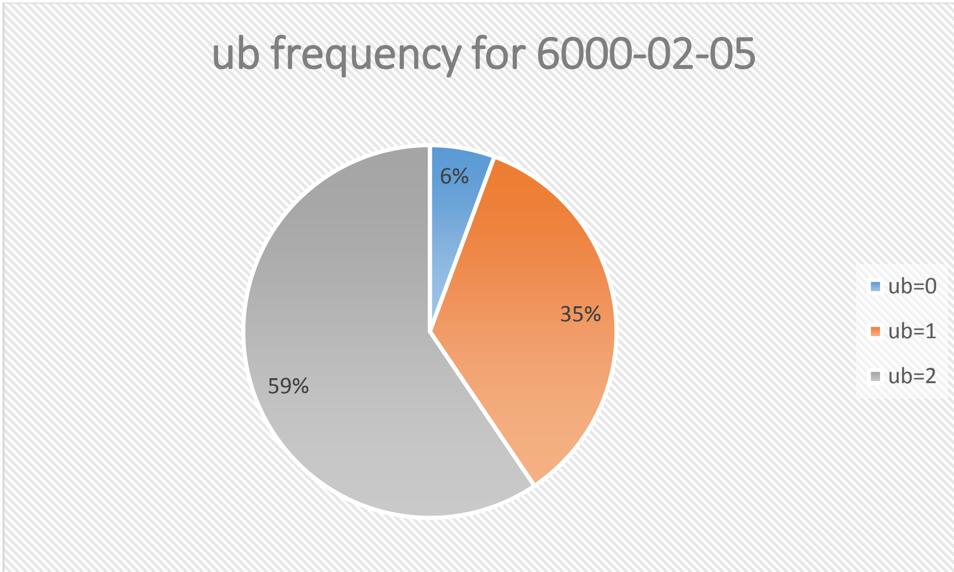

Figure 20 Pie-chart of frequency of ub values for 6000-02-05

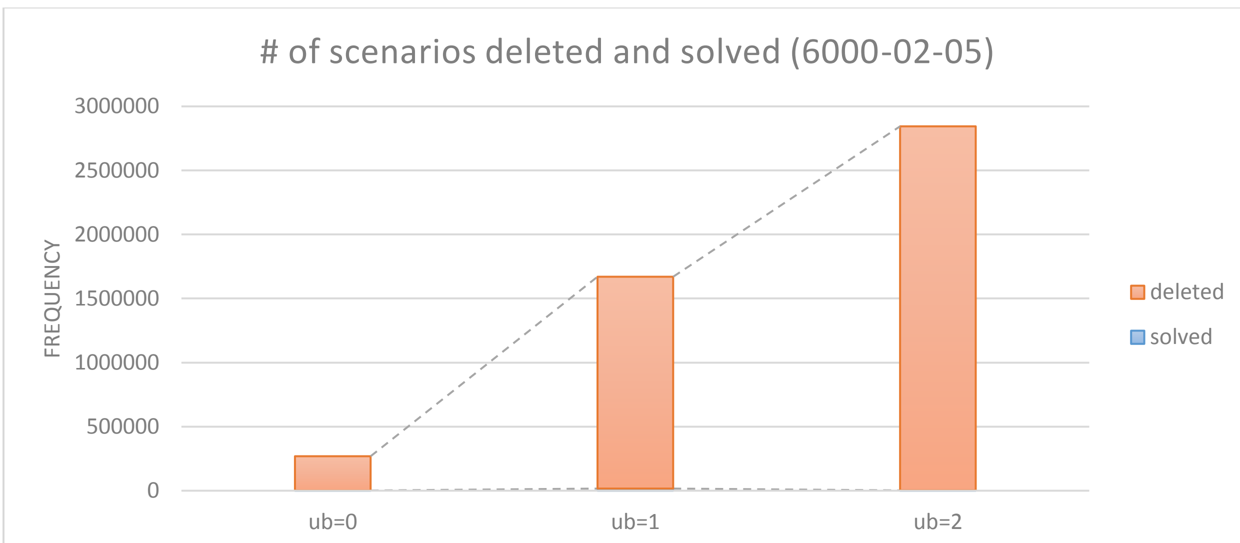

Figure 21 Number of scenarios deleted and solved (6000-02-05)

Table 14 ratio of scenarios deleted over solved for different ub values (6000-02-05)

ub	deleted/solved
ub=0	229.49
ub=1	102.18
ub=2	1431.84

Supposedly ub could be 6 under different environmental input parameters. To illustrate the performance of the model in those cases, we run the SSD algorithm for the other set of input parameters. This set of parameters are $B0 = 10000$, $I = 2\%$, $R = 2\%$. Under this new setting, the maximum value for ub is 6 in the good scenarios.

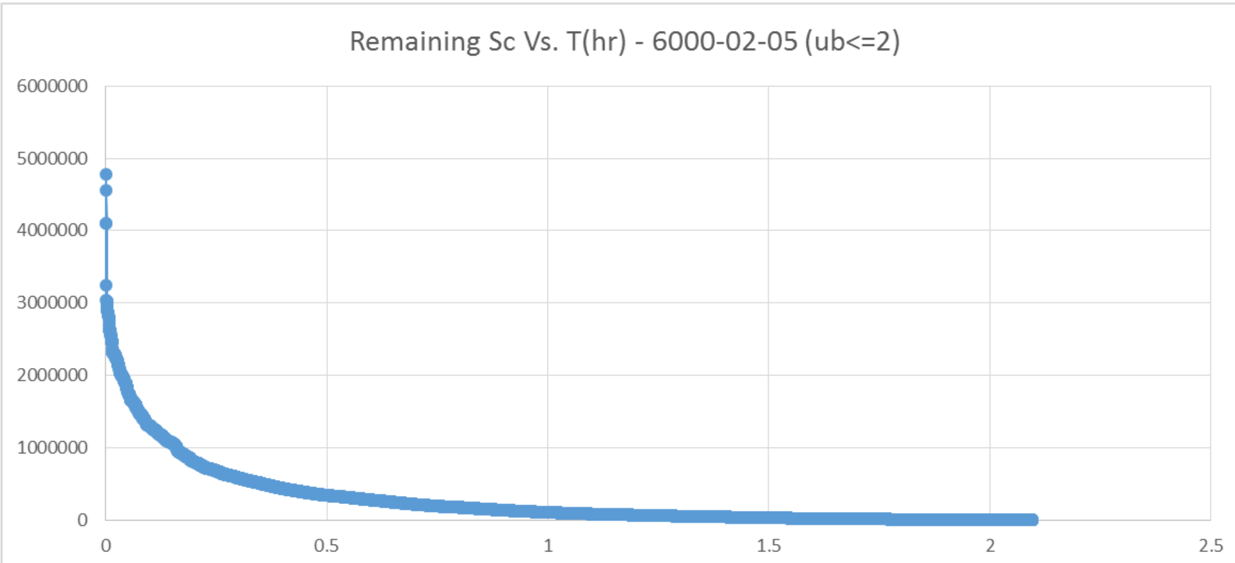

Figure 22 Remaining number of scenarios vs time (hr) for B0=6000, I=5%, R=2%

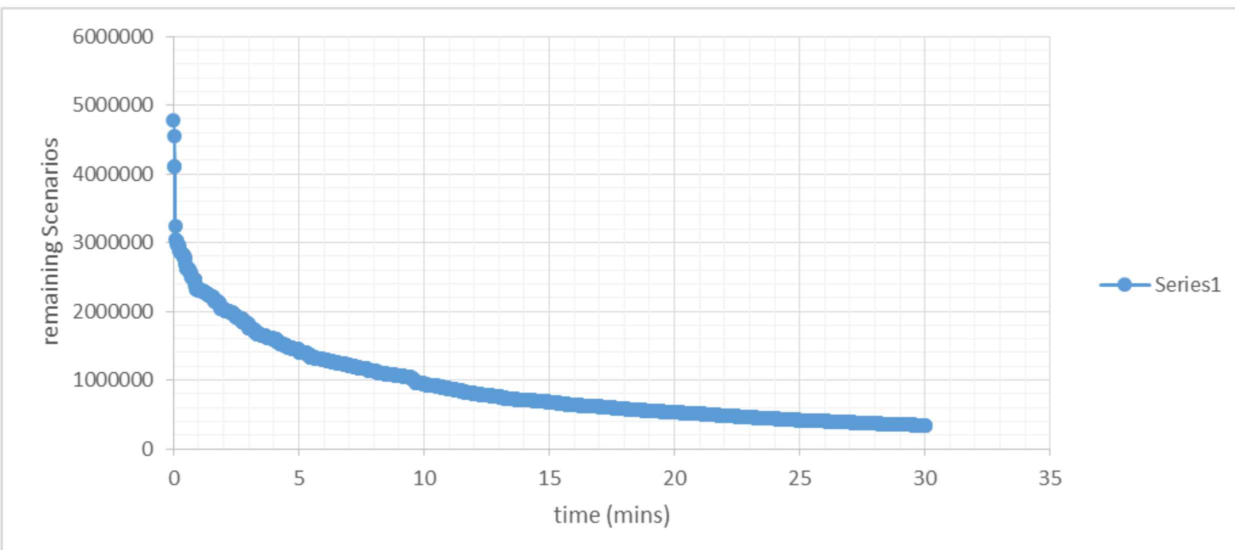

Figure 23 Remaining number of scenarios based on (guessing algorithm ran for 30 minutes) for B0=6000, I=5%, R=2%

Figure 24 and Figure 25 illustrate how the number of scenarios decrease as time proceeds for the case that the input parameters are $B0 = 10000, I = 2\%, R = 2\%$. For these inputs, the value for all scenarios are found after approximately 32.83 hours and

281,283 iterations. This is still a lot (approx. 91.8%) less than the 400 hours needed to solve all the 4.78e6 scenarios. The number of remaining scenarios after 30 minutes of execution of the algorithm is: 3,318,770. This decrease in the rate of solving/finding all scenarios in comparison to the $B0 = 6{,}000, I = 5\%, and\ R = 2\%$ case is due to more possible outcomes for the ub variable in this case and also due to fewer number of scenarios deleted per iteration. The average running time per iteration is higher in the second case (0.42 secs vs 0.39 secs). This increase in running time per iteration is due to the additional searches needed to find compatible scenarios.

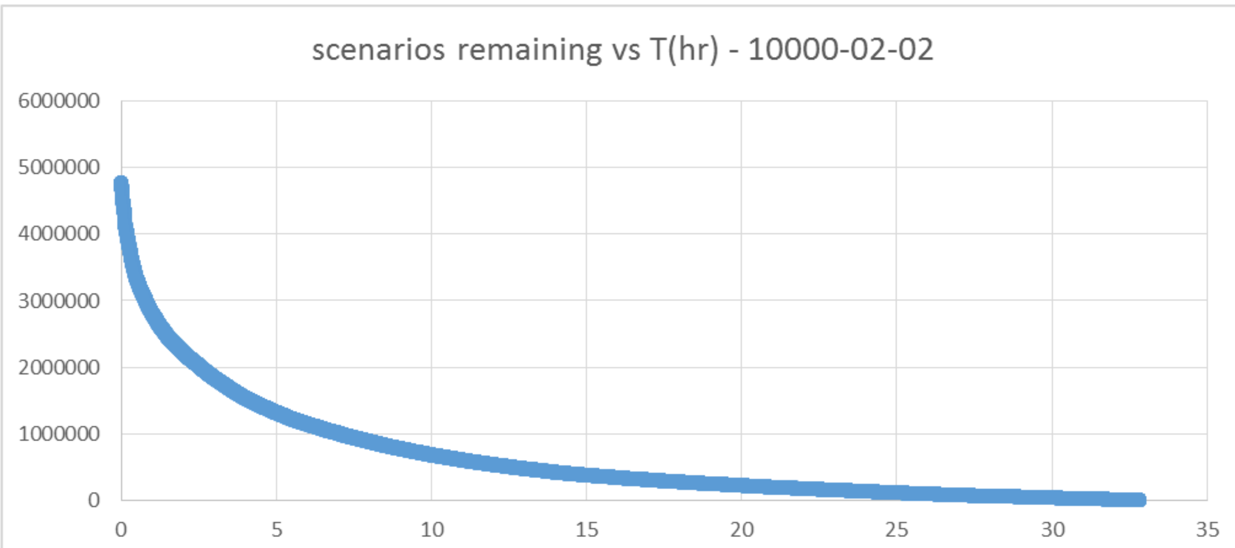

Figure 24 Remaining number of scenarios vs time (hrs) – B0=10000, R=2%, I=2%

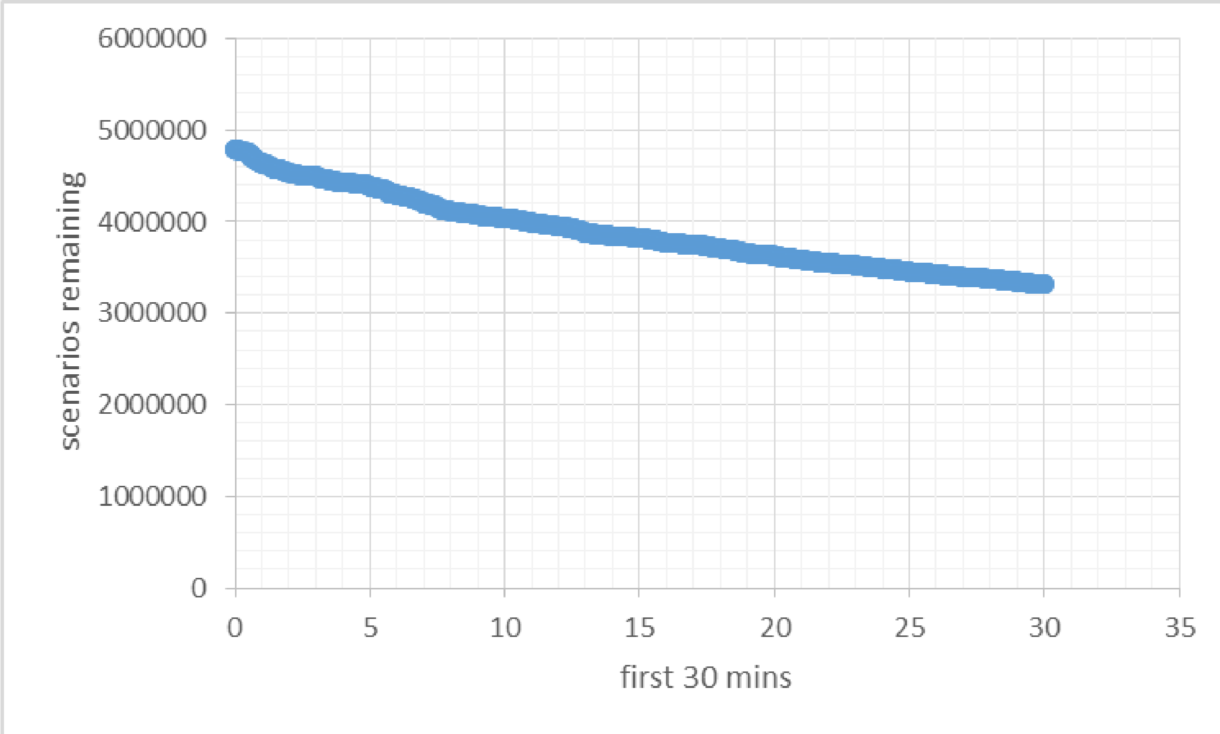

Figure 25 Remaining number of scenarios vs time (first 30 mins)- B0=10000, R=2%, I=2%

The percentage of scenarios with different ub values for B0=10,000, R=2%, and I=2% can be viewed in Figure 26. Based on this, the most common ub is $ub = 2$ with $ub = 3$ and $ub = 4$ trailing it. $ub = 4$ is the optimal here and the now decision value for about 21% of all the scenarios. One could therefore decide to build the infrastructure for up to 4 phases if the cost of that is not a lot so that in the future if any of those 21% of the scenarios happen, they would gain the most. However, if the stakeholders are only interested in the most common size of infrastructure, they should invest in the infrastructure required for 2 phases.

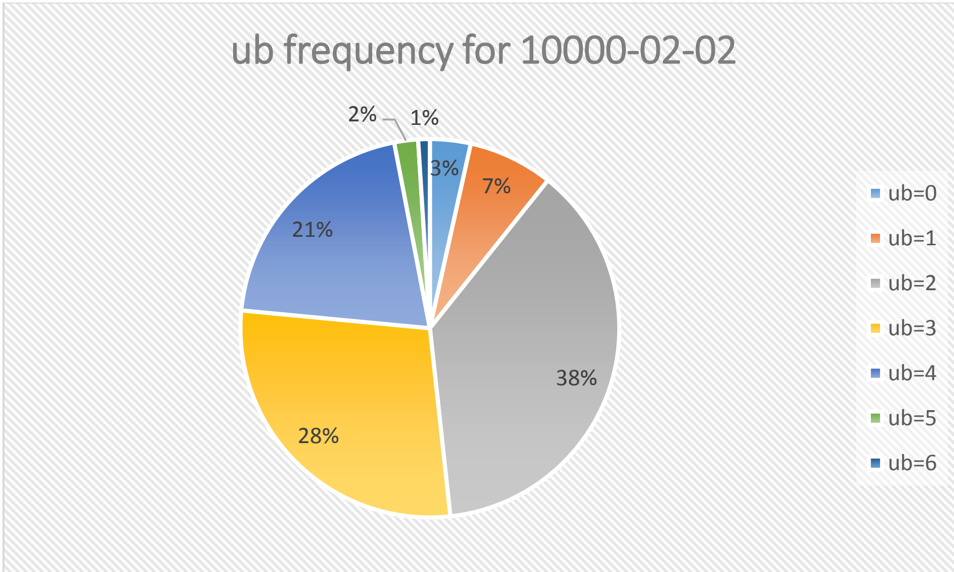

Figure 26 Pie-chart of frequency of ub values for 10000-02-02

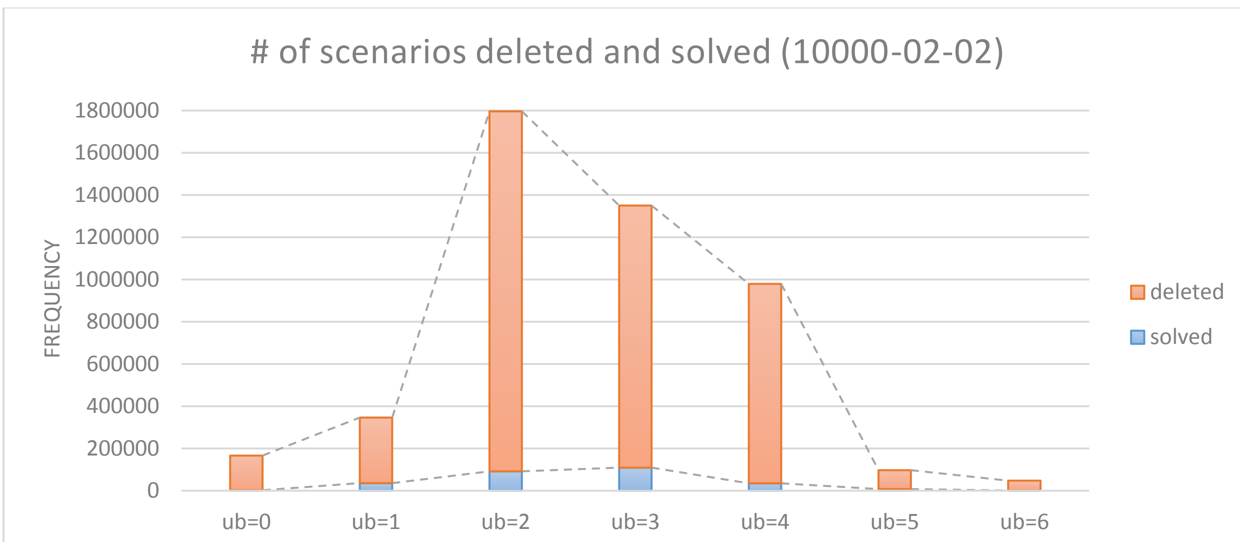

Figure 27 Number of scenarios deleted and solved (10000-02-02)

Table 15 ratio of scenarios deleted over solved for different ub values (10000-02-02)

ub	deleted/solved
ub=0	81.50
ub=1	8.71
ub=2	18.62
ub=3	11.38
ub=4	27.09
ub=5	11.87
ub=6	78.04

Figure 27 and Table 15 illustrate the ratio of number scenarios deleted over number of scenarios solved for each possible value of the here and now decision variable. As it can be seen, again, most of the largest ratios are for the maximum and minimum value of the here and now decision variables.

4.7.2 Solving the parking SSPP problem using simulation

For the parking planning stochastic problem, since we have 3^{14} scenarios, we cannot solve the stochastic optimization problem efficiently. However since the possible outcomes for the here and now decision are limited to 7 cases, we can find a good approximate solution to the here and now variable using the proposed approach. Before proceeding with the algorithm, we can easily find the final expected objective function value for the case in which $ub = 0$. The optimal objective function value for this case is equal to $B0$ as we will not be doing any investments and therefore the NPV will not change. We are then left with 6 remaining policies. Before proceeding with running the simulation for all possible policies, we perform a sensitivity analysis on the number of iterations to visually identify the competing policies. Table 16 summarizes the estimate of the bound on the objective function value for each policy under certain number of iterations for simulation. The parameters used for the problem are: $B0 = 10{,}000$ and risk free interest rate and inflation rate of 2%. The highlighted cells mark the highest bound on the objective function value. As it can be seen, it seems that policies 4 and 5 are competing and the best policy is always among these two.

Table 16 Preliminary analysis for identifying competing policies

# simulations per ub	ub = 1	ub = 2	ub = 3	ub = 4	ub = 5	ub = 6

500	10213.1	12034.23	13477.8	14030.38	14003.16	13924.37
1000	10213.1	12035.17	13495.39	13987.5	14010.25	13959.66
5000	10213.04	12039.68	13448.54	14017.93	13949.38	13911.69
10000	10213.06	12042.3	13466.07	13968.63	13925.26	13909.88

For this case, initially, to estimate the variance and average and to find the number of iterations needed per simulation run, we simulate each policy using an initial guess of 500 iterations. Table 17 summarizes the outcomes for the case in which we start with initial budget of $B0 = 10{,}000$ and risk free interest rate and inflation rate of 2% for policies 4 and 5. Note that the difference in the bound on the objective function value of policies 4 and 5 and the same number of iterations between Table 16 and Table 17 is due to different random seeds.

Table 17 Outcome of 500 iterations for B0=10,000 - I=2%, R=2%

Policy	**Policy 4: ub=4**	**Policy 5: ub=5**
Avg. O.F	14060.72	13759.65
Variance	2.31E+06	2.05E+06

If we accept a margin of error of 20, with a confidence coefficient of 0.935, the number of iterations needed is:

$$\max\left\{\frac{1.514^2 \times 2313306.68}{20^2}, \frac{1.514^2 \times 2046324.99}{20^2}\right\}$$

$$= \max\{13256.38, 11726.44\} = 13{,}257$$

Based on the above calculation, perform the simulation runs with ≈13,500 iterations per run. We want to use simulation to see which one of the policies 4 or 5 are better.

So we are interested in the difference between the bounds on the objective function values that are the outcomes of each simulation run.

We start off with 2 simulation runs per policy and perform a t-test for the difference between the two policies. The difference is not meaningful and we keep on adding new runs to decrease the standard error and therefore make the differences meaningful. We stop at 12 simulation runs because the difference is meaningful and also the lower bound of the difference is approximately equal to 2 times the accepted error used for finding the number of iterations needed per run, 20. The summary of the results and calculations are provided in Table 18. From Table 18, we can say that the difference in bounds on the objective function value between policy 4 and 5 is within the following range:

$$(\bar{F}_4 - \bar{F}_5) - t_{0.05,\ \nu} \times S.E.\left(\bar{F}_4 - \bar{F}_5\right) \leq F_4 - F_5 \leq (\bar{F}_4 - \bar{F}_5) + t_{0.05,\ \nu} \times S.E.(\bar{F}_4 - \bar{F}_5)$$

$$39.87302 \leq F_4 - F_5 \leq 58.89203$$

Table 18 Summary of simulation runs and t-test for policies 4 and 5

Simulation Run	Avg. obj. func. Val.		Variance	
#	**Policy 4**	**Policy 5**	**Policy 4**	**Policy 5**
1	13990.18972	13954.7378	2121269	2176838
2	13990.28721	13945.57042	2096654	2241521
3	13999.90093	13955.14194	2095595	2201416
4	14015.0007	13923.48089	2103278	2205163
5	14001.22141	13953.46931	2093618	2196729
6	13996.39225	13945.92813	2122183	2185678
7	13974.10661	13963.34725	2080020	2191451
8	13991.85706	13946.82828	2156072	2217405
9	14003.82788	13935.93889	2106939	2223724

10	14021.29753	13957.63289	2099822	2226776
11	13993.86869	13954.15718	2087012	2227678
12	13990.17502	13939.30168	2112367	2207908
	Analysis		Equation	
Avg of all runs	13997.34375	13947.96122	$\bar{F} = \frac{\sum_{i \in 1..12} F_i}{12}$	
Variance of all runs	140.3822104	110.5356856	σ^2	
Difference in Avg. of p4-p5	49.38252819		$\bar{F}_4 - \bar{F}_5$	
Standard Error	4.572726175		$S.E. = \sqrt{\frac{\sigma_4^2}{12} + \frac{\sigma_5^2}{12}}$	
Degree of freedom	21.69306586		$\nu = \left(\frac{\left(\frac{\sigma_4^2}{12} + \frac{\sigma_5^2}{12}\right)^2}{\left[\frac{\left(\frac{\sigma_4^2}{12}\right)^2}{11}\right] + \left[\frac{\left(\frac{\sigma_5^2}{12}\right)^2}{11}\right]} \right)$	
2 t-test critical value for 95% confidence	2.079613845		$T^{-1}(0.05, \nu)$	

So we can approximately say that policy 4 is better than policy 5 with an overall confidence coefficient no worse than:

$$1 - \big(P(A) + P(B)\big) = 1 - 0.065 - 0.050 = 0.885$$

The running time per each run depends on the number of iterations.

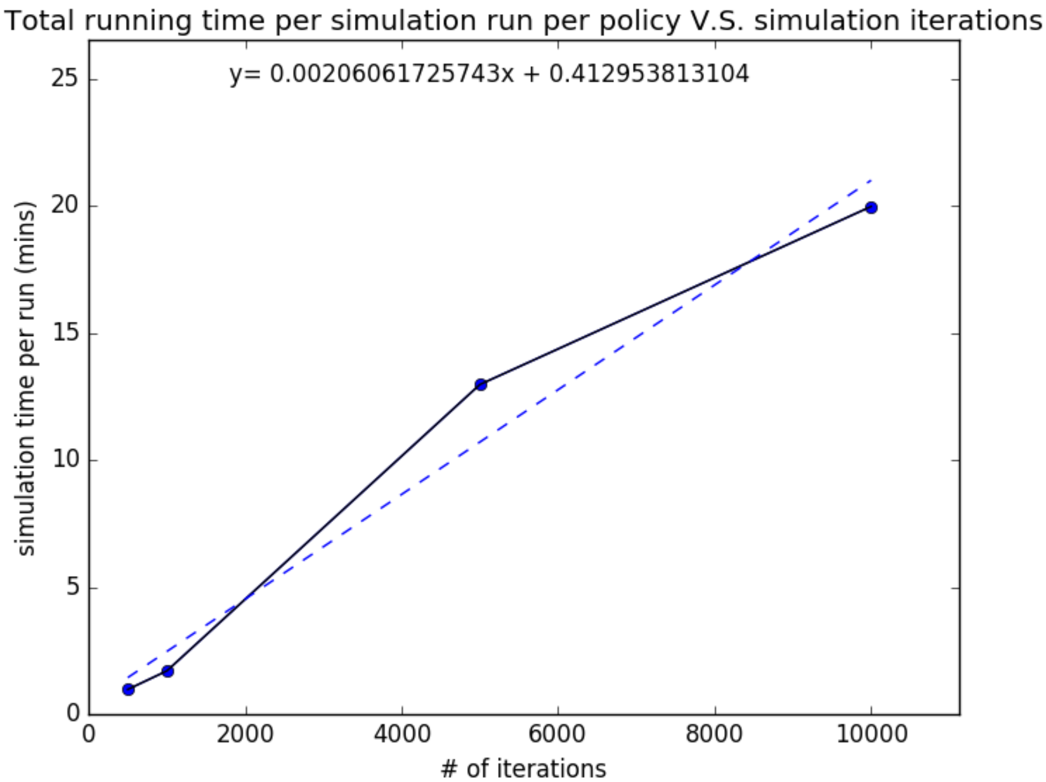

Figure 28 Running time per run v.s. number of iterations

Using the equation from fitting a 1d line to the data as show in Figure 28, we can calculate the expected time for each run for each policy with 13,500 iterations to be about 30 minutes:

$$t = 0.00206 \times (13500) + 0.41295 \approx 28.22\, mins$$

Based on this, tabulating Table 18 that has 2 policies and 12 runs, requires approximately 11 hours and 17 minutes. Via extrapolation, we can see for running 12 simulation runs for all 6 policies, a total of 33 hours and 52 minutes would have been required. This is a lot of time for an overall confidence of no more than 90%.

However, as mentioned earlier, CRN can potentially yield to a better comparison with higher confidence. The results for one run of the simulation via CRN is summarized in Table 19.

Table 19 Results for CRN simulation (B0=10000, I=R=2%)

run #	Avg. Obj. func. Value		differences	running time (mins)
	ub =4	ub = 5		
1	13997.56301	13953.52996	44.03305	57.3207
2	13994.05845	13948.48336	45.57509	58.3452
3	14015.35577	13970.48748	44.86829	58.568
Analysis				
Avg of difference			44.825	
STD of difference			0.772	
Degree of freedom			2	
2 t-test critical value for 99% confidence			9.925	

Based on results from Table 19, we can see that the difference between policies 4 and 5, falls within the range expressed below with 99% confidence:

$$44.825 - 9.925 \times \frac{0.772}{\sqrt{3}} \leq \bar{a} - \bar{b} \leq 44.825 + 9.925 \times \frac{0.772}{\sqrt{3}}$$

$$40.401 \leq \bar{a} - \bar{b} \leq 49.249$$

Note that the overall confidence here is at least 1-(0.01 + 0.065) that equals 92.5 %. This confidence is achieved in less than 3 hours and with a lot fewer simulation runs. The results are however, consistent as policy 4 is the best for both simulation cases (with and without CRN).

Chapter 5: Deterministic Multi-project Phased-investment Optimization

In this chapter of the dissertation, we take the deterministic single-project phased-investment problem of chapter 3 and extend it to a multi-project setting. In the case of having a portfolio of projects, one of the most important driving factors influencing the investment decision making is the availability of funds. In the presence of more funds, the decision makers can more easily disaggregate the pool of projects into separate individual projects that could be treated individually. However, in the more realistic case, under limited funds, the amount of investment in a project greatly affects the availability of funds for future projects. Once some funds are invested in a project, the available budget is decreased by the amount of investment. However, the budget will begin to increase upon arrival of revenues resulted from the investments.

Some of the different attributes can make a project become favorable are listed below:

- Certainty of the outcomes: as the certainty increases, we can be sure about the availability of the funds in future and can make rigorous decisions.
- Profitability: the amount (or expected amount) that a project can generate profit greatly affects its favorability.
- Revenue collecting period: the period (or time periods) that we receive the revenue from the investment is also very important. Generally, the sooner we start receiving revenue, the better the project is. The available funds increase by depositing the revenue received. This opens the opportunity for better investments at and after the time of receiving the revenue(s). In addition,

faster returning projects, accelerate the time at which the investment breaks even and starts to be profitable.

5.1 Introduction and Literature review for Multi-project optimization

The overall goal for the Deterministic Multi-project Phased-investment Problem (DMPP) is to assist managers in making a simultaneous project selection and scheduling decision. We categorize the studies into different classes. The first class is the general project selection and portfolio optimization class. Another class of research is the research that considers interdependencies among projects in the project selection problem. Another class is the class of project scheduling problems. The intersection of selection and scheduling is another area studied. Finally, the most similar class to this study is the project selection and scheduling with interdependencies among projects. The remaining of this section, looks at the problems in each of these classes in more detail.

5.1.1 Project selection and portfolio optimization

In the presence of more than one candidate project, the main problem that concerns management is the selection of a few projects from a pool of existing projects. This problem is known as the project selection problem. While the project selection problem with one main constraint is small modification to the knapsack problem, many variants of the project selection problem are widely studied in the literature. Usually these variants are either combining the project selection problem with other problems such as scheduling or markup estimation (Shafahi & Haghani, 2014). Or, they are modifying the objective function or modeling the problem as a multi-objective optimization problem (Shafahi, 2012). In the latter case, the problem is usually rebranded as "portfolio optimization". For a relative recent survey on portfolio optimization please read (Mansini, et al., 2014).

One of the approaches taken in portfolio optimization and in the presence of multiple objectives is building a single objective based on the priorities and weights of each of the objectives. In this case, by taking a weighted average of each objective, we are left with a single objective and hence can thereon treat the problem as project selection problem with a new hybrid objective. Many of the studies in project portfolio selection focus on the multi-criteria optimization aspect and utilize qualitative methods such as AHP, weighting methods, and ranking methods. For the purposes of this study, we direct our attention to cases with one objective keeping in mind that we can transform a multi-objective problem (a portfolio problem) into a single objective problem using the mentioned weighted average method. There are many methods for modeling the project selection problem. Two of which are goal programming (Badri, et al., 2001), and mathematical modeling (integer programming).

5.1.2 Project selection with interdependencies among projects

In the realm of project selection, a majority of the problems assume that the projects have minimum or zero interactions on each-other. Even-though this is the case for some projects, most projects somehow affect other projects that at least fall within the same category. The interdependencies and interactions among projects mainly fall into three different categories, namely: benefit, cost, and outcome. The benefit category refers to an increase in profit of a given project as a result of doing another project which is related (dependent) to that project. The Outcome category refers to the increase in the probability of success for a given project if an earlier project which is in the same category is completed. Finally, the cost category refers to the decrease in costs and all other resources which a given project is consuming if an earlier project of that kind is completed (Shafahi & Haghani, 2013).

These interdependencies among projects have been addressed in previous researches (Killen & Kjaer, 2012) (Liesiö, et al., 2008) (Bhattacharyya, et al., 2011) (Dickinson, et al., 2001).

The project selection and decision making problems which consider interdependencies have been dealt with using different solution techniques. Some have used goal programming (Santhanam & Kyparisis., 1996) (Lee & Kim, 2000) (Lee & Kim, 2001). Others have approached the problem with linear programming, branch and bound, or using heuristic approaches (Iniestra & Gutierrez, 2009) (Schmidt, 1993) (Carazo, et al., 2010) . Constraint Programming is also another approach used for solving problems of this type (Liu & Wang, 2011).

5.1.3 Project scheduling

As mentioned before, some problems related to project selection are hybrid problems that mix project selection with other famous problems. One of the most studied problems is project scheduling. Scheduling is the profession of finding "appropriate" times for execution of projects or activities. For a survey on deterministic project scheduling, we refer the interested reader to (Kolisch & Padman, 2001). The integration of project scheduling and project selection is very important since once we have selected a project we should also determine its schedule. (Chassiakos & Sakellar, 2005)

Perform time-cost analysis of construction projects, formulate the problem using integer programming, and present an approximation solution method that solves the model by decreasing the number of integer variables and constraints. Their study focuses on a single project. (Icmeli, et al., 1993) conduct a survey of different problems that are related to scheduling.

If the scheduling is affected by the constraints enforced by the limitations of funds and other resources, the scheduling problem is known as a "Resource-constrained Project Scheduling Problem" (RCPSP). RCPSP is an NP-hard problem and hence many researches have focused on developing heuristics and solution algorithms. (Chen, et al., 2010) use ant-colony to solve the RCPSP. (Chan, et al., 1996) model a construction scheduling project as a RCSP and use genetic algorithms to solve it. GA is a very popular meta-heuristic for solving these classes of problems. (Gonçalves, et al., 2008) also use GA to solve the RCPSP problem when we have more than one project. (Kim & Ellis Jr, 2008) use GA with elitism to solve the scheduling problem for single projects. They assume that problem with more than 60 activities are considered large. (Brucker, et al., 1998) and (Dorndorf, et al., 2000) solve the RCPSP problem using a branch and bound procedure. (Zhang, et al., 2006) use Particle Swarm Optimization (PSO) for solving RCPSP with the objective of minimizing duration and compare the performance of PSO with GA. (Hartmann & Kolisch, 2000) and (Kolisch & Hartmann, 2006) compare some of the different heuristics that are being used for solving the resource constraint scheduling problem. The resource constraint scheduling problem has many variations, some of which can be found in (Hartmann & Briskorn, 2010).

Another famous scheduling problem studied, especially in computer science is job shop scheduling. The focus of this problem is scheduling the jobs and assigning them to the machines that can execute them. Most of the Job Shop Scheduling problems assume that all jobs have to be executed. In other words, there is no selection of the jobs.

5.1.4 Intersection of selection and scheduling

In contrast to the amount of research available about project selection, and project scheduling, very few research focus on the intersection of these two problems (Carazo, et al., 2010) (Sun & Ma, 2005). Many of the studies are dedicated to developing different solution algorithms since each problem by itself is complex and thus the intersection would be even more complex. Some example studies are: (Coffin & TAYLOR III, 1996) that use filtered beam search heuristic to solve the problem.

5.1.5 Project selection and scheduling with interdependencies among projects

In the subject of the mixture of project scheduling and selection, yet very few studies exist that in addition to modeling these two problems simultaneously, consider some interdependencies among projects as well.

In (Tao & Schonfeld, 2006) and (Tao & Schonfeld, 2007), the authors consider the problem of scheduling and selection of interdependent transportation projects. They capture interdependencies beyond more than just pairwise dependence between projects. They develop an island model for solving the problem. Island models are variants of the traditional GA models that generally achieve better results in comparison to traditional GAs. In another study by (Shayanfar, et al., 2016), the authors try to prioritize the projects and compare three different metaheuristics, namely GA, SA, and TS and conclude that for their application of scheduling and selection with interdependencies, GA yields the most consistent solutions.

In (Zuluaga, et al., 2007) a MIP formulation for the selection and scheduling problem is presented that includes three types of interdependencies among projects: resource, technical, and benefit. The authors also include scheduling relationships. In their example they have a project with negative NPV that after running their model it is not included! This could have been prevented by preprocessing!

In (Ballou & Tayi, 1996) a framework for facilitating software maintenance projects and their staffing is provided. Initially, the selection process is modeled as an IP and afterwards, for the selected projects, staffs are assigned based on a transportation algorithm.

(Tofighian & Naderi, 2015) use ant-colony to solve the integrated selection and scheduling problem. They consider two objectives: maximizing benefit and minimizing the maximum level of required resources. The only type of interdependency they model is mutual exclusiveness. Their study lacks re-investment strategies.

The study by (Jafarzadeh, et al., 2015) has re-investment strategies such that the profit yielded from completing projects can be invested for implementing other projects. The planning horizon in their study is flexible and one objective of their study is to find the best time horizon. Although they consider re-investments they do not model interdependencies among projects and assume that each project is independent. They model the problem as an MIP and find commercial solvers sufficient enough for solving their proposed problem. Another study that allows for re-investment is (Belenky, 2012). In one of the generalized cases in the study, scheduling interdependence and priorities are considered.

(Carazo, et al., 2010) allow transfer of unused funds between the current and next time period within their modeling. They consider existing synergies among projects when they are done at the same time. Their MIP model is non-linear and hence they solve the model using a two-step method. In the first step, Tabu search is done and in the second step scatter search is done.

There are also some studies which model the integrated project selection and scheduling problem as an uncertain problem. (Huang & Zhao, 2014) present a study in which uncertainty is added to the integrated project selection and scheduling problem. Apart from the scheduling restrictions and dependency of the projects, they assume projects are independent. To solve the problem, GA is used. This research introduces the flexibility in project start times. Each project can start anytime within its certain time-frame. (Sefair & Medaglia, 2005) also incorporate uncertainty into their modeling and have minimal interdependencies among projects. They simultaneously maximize the NPV of their portfolio and minimize the variance to minimize risk. The MIP model is solved using commercial solvers. Fuzzy logic has also been used as a means for modeling the uncertainties (Coffin & Taylor, 1996).

To the best of the knowledge of the author, no study exists that models the integrated project selection and scheduling for a pool of projects that themselves can be broken down into sequential phases. This study aims to fill this gap by expanding the single project deterministic model into a model that can handle the optimization of a pool of phased projects. The traditional models for selection and scheduling are special cases of our model in which all of the projects are single phase projects.

Some of the other contributions of the DMPP introduced in this dissertation is listed below:

- Developing a mathematical model that assists managers in making selection and scheduling decisions for cases that some or all of the projects are made up from smaller sequential phases. In addition to considering the following facts:
    - Time and cost interrelationships among phases.
    - Transfer of available funds between time periods.
    - Ability to abandon a project prior to the end of its duration.
    - Ability to start investment into a project within a time-frame.
- Presenting a preprocessing step that can reduce the number of variables and constraints without any compromises in terms of the objective function value.
- Developing a heuristic to solve the problem.

5.2 MIP model for Deterministic Multi-project Phased-investment Problem (DMPP)

We present a deterministic MIP model that can assist managers in identifying which projects to invest in and for the selected projects, which phases to invest in and when. In the presented MIP model, we allow the stakeholders to abandon the project when the profits do not cover the costs.

The input parameters for this problem are categorized into environmental parameters and project parameters. The environmental parameters are those which are common for all projects. Each project by itself has many attributes associated with it. Some examples are: duration of project, the time frame in which we can start the project within, the costs for its infrastructure and phases, and etc. The parameters are aggregated in Table 20.

Table 20 Parameters for deterministic multiple phased investment project optimization

Environmental Parameters	
$B0$	Initial budget available at t=0
R	Risk free interest rate
I	Annual inflation rate
T	Planning time period
$times$	Set of times = {0,1,…,T}
M_T	Big-M used in constraints = $(2 \times T + 1)$
Project Related Parameters	
UB_k	The upperbound on the number of phases associated with project k
$AvailableT_k$	The first time in which project k is available for investment (The soonest that we can start the preparation of the infrastructure)
$LatestT_k$	The latest time that we can start the preparation of the infrastructure required for project k
$Duration_k$	The maximum duration that project k can be ongoing
$P_{k,i,t}$	Profit gained from first i phases of project k at time t (@ AvailableT)
$IT_{k,ub}$	Preparation (construction) time for infrastructure of project k for ub phases
$ICOST_{k,ub}$	Preparation (construction) cost for infrastructure of project k for ub phases (@ AvailableT)
$PT_{k,i,j}$	Duration required for implementation of each phase $i-j$ when phases $i-j$ are implemented together for project k
$PCOST_{k,i,j}$	Construction cost of each phases between i,j for project k if phases $i-j$ are being done together based on the money of the first time the project is available (@ AvailableT)
$OCPA_{k,i,t}$	Operation cost at time t for project k when number of phases that have

	already been implemented or are being implemented at time t is i (@ AvailableT)
$projects$	Set of projects
$phases$	Set of phases = {1,…, max(UB_k)}

The variables used are summarized in Table 21.

Table 21 Variables for deterministic Multi project Phased investment model

Variables	
ub_k	Number of phases of project k selected for implementation
Variables regarding infrastructure	
ist_k	The actual time in which we start investing in the infrastructure needed for project k
it_k	Construction duration for infrastructure of project k based on actual number of phases to be implemented
$costub_k$	Construction cost for the infrastructure required for implementing ub phases of project k
$tcostub_{k,t}$	A variable that equals $costub_k$ at time ist_k
Variables regarding phases	
$x_{k,i,j}^t$	Binary variable that equals 1 if phases i through j of project k start their implementation at time t
Variables regarding projects	
$finished_{k,t}$	Binary variable that equals 1 if project k is ended at time t (we have exceeded its duration)
$suspend_{k,t}$	Binary variable that equals 1 if we abandon project k at time t
$n_{k,t}$	Number of phases that have already been implemented or are being implemented at time t for project k
Other main variables	
b_t	Available budget at the beginning of each time t
Variables used for linearization	

$cn_{k,i,t}$	Binary representation of number of phases that have already been implemented or are being implemented at time t for linearization for project k
$nc_{k,i}$	Binary variable for linearization of infrastructure cost for project k
$indict_{k,t}$	Binary variable that equals 1 if project k is starts at time t (used for binary representation of ist_k)

The objective function is maximizing the NPV of the available budget at the beginning of time T plus the future costs and revenues afterwards from all projects where the profits and costs are subject to inflation and interest rates. It is expressed in (5-1)

$$\max z = \frac{b_T + \sum_k \sum_i P_{k,i,T} \times cn_{k,i,T} \times (1+R)^{T-AvailableT_k}}{(1+R)^T} -$$

$$\frac{\sum_k \left(\sum_i \sum_j (j-i+1) \times x_{k,i,j}^t \times PCOST_{k,i,j} + \sum_i OCPA_{k,i,T} \times cn_{k,i,T} \right) \times (1+I)^{T-AvailableT_k}}{(1+R)^T} \tag{5-1}$$

For each project, the constraints of the problem are categorized into infrastructural constraints, phase constraints, and scheduling constraints. The budgetary constraints are common. The infrastructural constraints for the projects are listed below:

$$n_{k,t} \leq ub_k \qquad \forall k, t \tag{5-2}$$

$$costub_k = \sum_i ICOST_{k,i} \times nc_{k,i} \qquad \forall k \tag{5-3}$$

$$it_{k,ub} = \sum_i IT_{k,i} \times nc_{k,i} \qquad \forall k \tag{5-4}$$

$$ub_k = \sum_i i \times nc_{k,i} \qquad \forall k \tag{5-5}$$

$$\sum_i nc_{k,i} \leq 1 \qquad \forall k \tag{5-6}$$

$$ub_k \le UB_k \qquad \forall k \tag{5-7}$$

$$ist_k = \sum_t t \times indict_{k,t} \qquad \forall k \tag{5-8}$$

$$\sum_t indict_{k,t} = 1 \qquad \forall k \tag{5-9}$$

$$tcostub_{k,t} \ge costub_k - M_T \times \left(COSTUB_{k,UB_k} + 1\right) \times \left(1 - indict_{k,t}\right) \quad \forall k, t \ge AvailableT_k \tag{5-10}$$

Constraints (5-2) limit the total number of phases that are implemented for each project at each time to the maximum invested infrastructure. Constraints (5-3) and (5-4) are for calculating the infrastructure cost and install duration for each project. For them being linear, we need to express the number of phases selected for implementation, ub_k, using binary variables. This is done using constraints (5-5) and (5-6). Constraints (5-7) limit the size of the infrastructure. Constrains (5-8) and (5-9) are also for constructing the binary representation of ist_k. Constraints (5-10) assure that $tcostub_{k,t} = costub_k$ when $t = ist_k$.

Constraints (5-11) – (5-16) are the phase related constraints.

$$n_{k,0} = \sum_i \sum_j (j - i + 1) \times x^0_{k,i,j} \qquad \forall k \tag{5-11}$$

$$n_{k,t} = n_{k,t-1} + \sum_i \sum_{j \ge 1} (j - i + 1) \times x^t_{k,i,j} - suspend_{k,t} \qquad \forall k, t \ge AvailableT_k \tag{5-12}$$

$$n_{k,t} = \sum_i i \times cn_{k,i,t} \qquad \forall k, t \ge 1 \mid t \ge AvailableT_k \tag{5-13}$$

$$\sum_i cn_{k,i,t} \le 1 \qquad \forall k, t \ge AvailableT_k \tag{5-14}$$

$$\sum_{j\geq i} X_{k,i,j}^{t} \leq \sum_{l\leq i-1}\sum_{t'\leq t} x_{k,l,i-1}^{t'} \qquad \forall k, i \in \{2,\dots,UB_k\}, t \geq AvailableT_k \qquad (5\text{-}15)$$

$$\sum_{t}\sum_{i\leq l}\sum_{j\geq l} x_{k,i,j}^{t} \leq 1 \qquad \forall k, l \in \{1,\dots,UB_k\} \qquad (5\text{-}16)$$

Constraints (5-11) and (5-12) are for calculating the number of phases implemented/being implemented at different times for each project. Constraints (5-13) and (5-14) are for representing the number of phases that are implemented/are being implemented using binary variables so that we would have linear constraints when calculating the different costs and times for phases for each project. Constraints (5-15) prevent implementation of succeeding phases prior to the implementation of phases that are preceding them for each project. Constraints (5-16) prevent the assignment of a phase to two different groups of phases for all projects.

The scheduling constraints (5-17) - (5-23) are presented below:

$$it_k + ist_k \leq \sum_j \sum_t t \times x_{k,1,j}^{t} + M_T \times \left(1 - \sum_j \sum_t x_{k,1,j}^{t}\right) \qquad \forall k \qquad (5\text{-}17)$$

$$t \times x_{k,i,j}^{t} + (j-i+1) \times PT_{i,j} \times x_{k,i,j}^{t} \leq \sum_{l\geq j+1}\sum_{t'\geq t} t' \times x_{k,j+1,l}^{t'} + M_T \times \left(1 - \sum_{l\geq j+1}\sum_{t'\geq t} x_{k,j+1,l}^{t'}\right) \qquad \forall k, i, j \in \{2,\dots,UB_k\} \mid j \geq i,\ t \qquad (5\text{-}18)$$

$$\sum_i \sum_{j\geq i} x_{k,i,j}^{t} \leq 1 \qquad \forall k, t \geq AvailableT_k \qquad (5\text{-}19)$$

$$-M_T \times \left(1 - finished_{k,t}\right) \geq ist_k + Duration_k - t \qquad \forall k, t \geq 1 | t \geq AvailableT_k \qquad (5\text{-}20)$$

$$T \times (UB_k + 1) \times \left(1 - finished_{k,t}\right) \geq \sum_{t'\geq t} n_{k,t'} \qquad \forall k, t \geq AvailableT_k \qquad (5\text{-}21)$$

$$-M_T \times \left(1 - \sum_j \sum_{t' \geq 1} x_{k,1,j}^{t'}\right) - M_T \times finished_{k,t} + t + 1 \leq ist_k + Duration_k \qquad \forall k, t \geq 1 | t \geq AvailableT_k \tag{5-22}$$

$$AvailableT_k \leq ist_k \leq LatestT_k \quad \forall k \tag{5-23}$$

Constraints (5-17) assure that for each project, the first group of phases are implemented after the infrastructure is completed and successfully implemented. Constraints (5-18) ensure that each phase is implemented after the completion of implementation of its preceding phases for each project. Constraints (5-19) prevent multiple groups of phases to start their implementation together for all projects. Constraints (5-20) and (5-22) force a project to be finished whenever the last time in which the project could be ongoing is passed. Constraints (5-21) force the number of phases ongoing for a project to be zero if the project is finished. Constraints (5-23) ensure that each project starts within its allowable range.

The budgetary constraints (5-24) – (5-27) are expressed in the following:

$$b_1 = B0 - \left(\sum_k tcostub_{k,0} + \sum_k \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{k,i,j} \times x_{k,i,j}^0 + \sum_k \sum_i OCPA_{k,i,0} \times cn_{k,i,0}\right) \times (1 + R) + \sum_k \sum_i P_{k,i,0} \times cn_{k,i,0} \tag{5-24}$$

$$b_t = \left(b_{t-1} - \sum_k \left(tcostub_{k,t-1} + \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{k,i,j} \times x_{k,i,j}^{t-1} + \sum_i OCPA_{k,i,t-1} \times cn_{k,i,t-1}\right) \times (1 + I)^{t - AvailableT_k - 1}\right) \times (1 + R) + \sum_k \sum_i P_{k,i,t-1} \times cn_{k,i,t-1} \times (1 + R)^{t - AvailableT_k - 1} \qquad \forall t \in \{2, \dots, T\} \tag{5-25}$$

$$\sum_k tcostub_{k,0} + \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{k,i,j} \times x_{k,i,j}^0 + \sum_i OCPA_{k,i,0} \times cn_{k,i,0} \leq B0 \tag{5-26}$$

$\sum_k tcostub_{k,t} + \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{k,i,j} \times x_{k,i,j}^t + \sum_i OCPA_{k,i,t} \times cn_{k,i,t} \times$

$$(1 + I)^{t - AvailableT_k} \leq b_t \qquad \forall t \geq 1 \tag{5-27}$$

Constraints (5-24) and (5-25) are the updates on the available budget at the beginning of each time period. The available budget at the beginning of each time period is equal to remaining budget from the previous time period in the current period's value (incorporating time value of money) plus the profits earned at the end of the last period as a result of implemented phases from all projects in the previous period. Constraints (5-26) and (5-27) are the budget limitations during different times. Finally, the variable domain constraints are shown in (5-28) and (5-29).

$$ub_k, ist_k, it_k, costub_k, tcostub_k\ , b_t, n_{k,t} \geq 0 \tag{5-28}$$

$$x_{k,i,j}^t, cn_{k,i,t}, nc_{k,i}, indict_{k,t}, finished_{k,t}, suspend_{k,t} \in \{0,1\} \tag{5-29}$$

5.3 Solution methods for the Deterministic Multi-project Phased-investment Problem (DMPP)

To solve this problem fast, we present a heuristic algorithm. Prior to the heuristic algorithm, we propose a pre-processing step that can potentially decrease the problem size without any loss in the value of the optimal solution.

5.3.1 Pre-processing step for DMPP

In order to reduce the running time, we follow the simple pre-processing algorithm presented below. This algorithm helps eliminate "non-optimal" phases for each project. The main intuition behind the algorithm is that if a project, p, has N potential phases but even in the presence of an enormous amount of funds, no more than N' phases of it will be executed, we can decrease the potential number of phases from N to N' ($N \rightarrow N'$) without much or any loss in the optimal objective function value of the entire portfolio of projects. Note that we might be able to come up with pathological examples in which we might lose a lot in terms of the objective function value. However, these cases rarely, if ever, exist in reality. A pathological example is as follows. Assume we have two single phase projects A and B. Project A starts sooner than project B. However, it does not end prior to the start time of project B. Also, for simplicity assume that there is no flexibility in the start time and end time of either project. A schematic of the cash flow time-chart of these two projects can be seen in Figure 29.

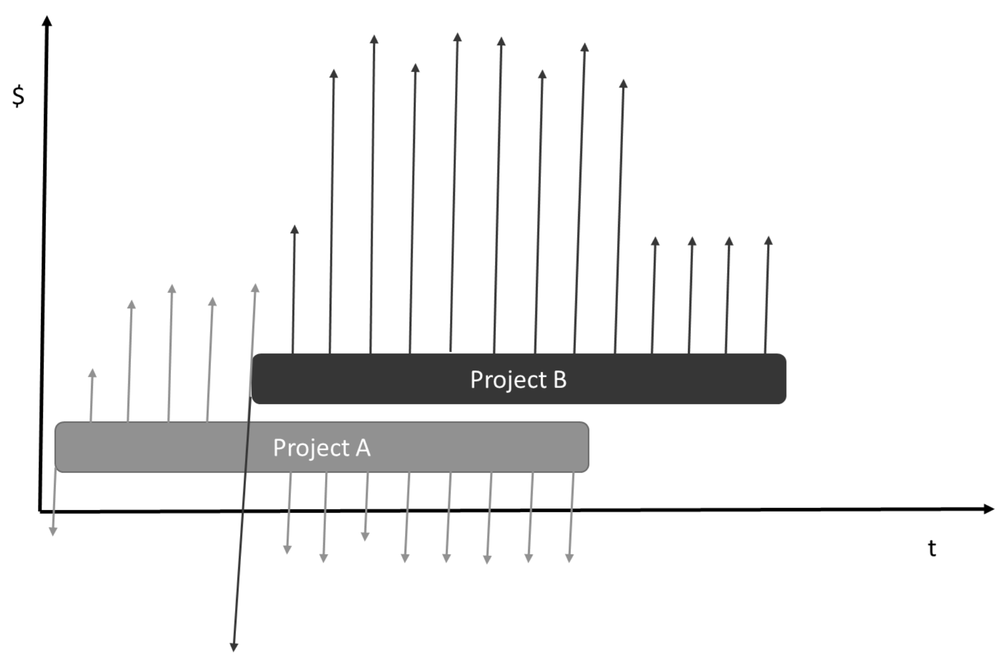

Figure 29 Pathological example for the downside of the preprocessing algorithm

In this example, assume that project A by itself is non-profitable under the existing interest rate and if executed, it will yield a loss of NPV_A. Therefore it should not be selected. However, project B is very profitable and should be selected. The selection of project B will yield an profit of NPV_B. Now assume that the initial budget required for project B exceeds the available budget at the start time of project B. Hence, project B cannot be executed. This is while, due to the cash flow of project A, if project A is executed, by the time we get to the start time of project B, we have gained enough funds that the execution of project B becomes feasible. In this example, we accept some loss in order to make more profit in the long run.

Note that this pathological example would not have existed if we were allowed to abandon project A prior to its end time. In this case, after the project started to become unfavorable, we would have abandoned the project. Since we allow for project abandonment, we do not need to worry about such pathological examples.

5.3.1.1 Pre-processing intuition

The intuition of this algorithm is to effectively reduce the UB_k parameter for projects, which is the maximum potential number of phases for project k. This parameter is given as an input for each different project. However, by decreasing the value of it, we can considerably reduce the number of variables and constraints needed. Note that many of the variables such as nc, x, nt depend on the maximum number of phases of the project.

5.2.1.2 Pre-processing algorithm

The initialization step of the algorithm is increasing the starting budget, $B0$, to a large amount. All the other input parameters remain intact. After this step, we iterate over all the projects, running the mathematical model (DMPP) for each project individually.

For this, we assume that the pool of projects of DMPP only has one project. Once the problem is solved, we compare the optimal value for ub_k, ub_k^* with the input parameter UB_k. Two cases might happen:

- $0 < ub_k^* \leq UB_k$: In this case, the project is overall economically feasible and we keep this project in the pool of available projects. Although, we update its UB value to the preprocessed value of UB, $UB_k = ub_k^*$.
- $ub_k^* = 0$: In this case, the project was never economically viable under the input parameters. We delete the project from the list of available projects.

The algorithm is summarized in the flowchart illustrated in Figure 30.

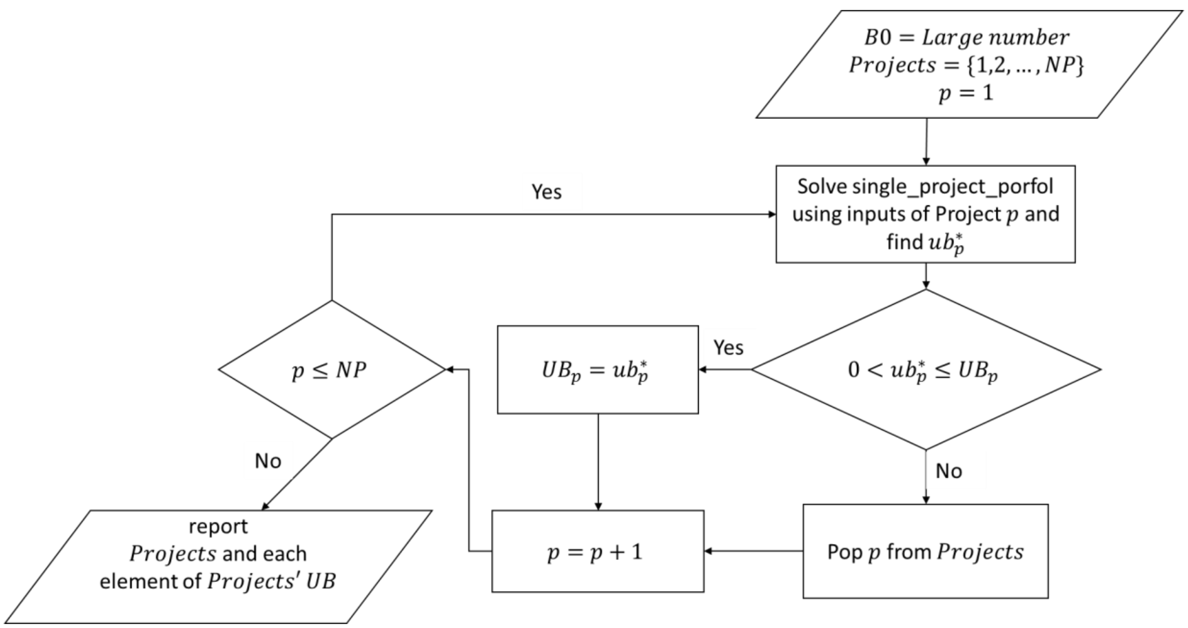

Figure 30 Flowchart for pre-processing

5.3.2 Solution heuristic for DMPP

We could assume that the optimal solution of the selection problem and scheduling problem of DMPP is built by adding projects one by one and limiting their UB value in an “optimal” way. We are not aware of this “optimal” sequence and “optimal” way of limiting the UB values for the projects. Our heuristic however tries to imitate this selection and mutation behavior by adding some randomness into the process. So, the general idea behind the heuristic is: 1) select a project from the pool of projects; 2) limit its UB value; 3) try to add the project to the list of projects that are already undertaken in previous steps. These general steps are repeated until all projects have been considered and no more projects remain in the pool of remaining projects. Each of the steps are explained in more detail in the next sub-sections.

5.3.2.1 Step1 of heuristic- Candidating: Selecting a project from the pool of remaining projects for consideration

In this step, we are interested in picking a project from the pool of available projects. For this step, we should have a measure of favorability for each project. This measure directly influences the order of selecting projects since we set the probability of selecting a project proportional to this measure. Some examples of the measures of favorability for the projects are:

- $Profit_{k,ub}$: NPV of profit for each project given its optimal ub value from preprocessing. This is easily calculated during the preprocessing step. This value is equal to the objective function value minus the initial available budget that was an input.

- $\frac{1}{duration_k}$: This is based on the fact that as the duration of projects become shorter, the investment period shrinks and therefore we can invest the funds faster.
- $\frac{1}{TotalCost_k}$: This measure favors the projects that require less overall capital.
- $\frac{Profit_{k,ub}}{TotalCost_k}$: This measure combines the first and third measures.
- $\frac{Profit_{k,ub}}{LastAvTime_k}$: This measure favors projects with higher profits from the pre-processing step and the ones which have to be selected sooner.

In the examples solved, we compare the performance of the heuristic when the favorability measure is $\frac{Profit_{k,ub}}{LastAvTime_k}$, H2, with the heuristic when the favorability measure is $Profit_{k,ub}$, H1.

Once this measure for all projects, M_k, is evaluated, we calculate the probability of selecting a project using equation (5-30) which is a simple normalization of the selection probabilities.

$$Select\_Prob_k = \frac{M_k}{\sum_{q \in Projects} M_q} \tag{5-30}$$

After this, a random number $r \sim U(0,1)$, is drawn form an uniform distribution between 0 and 1. Using r, we select a project, ps, for the second step of the heuristic.

5.3.2.2 Step2 of heuristic- Limitation: Limiting the candidate project's phases using mutation

In this step, for the selected project that is under consideration at step 1, ps, we decrease its UB value, UB_{ps} to $\overline{UB}_{ps} \sim [1, UB_{ps}]$ with mutation probability, $Limit\_Prob$.

5.3.2.3 Step3 of heuristic- Inclusion/Not inclusion: Deciding whether to include the project from step1 and step2 in the solution or not

In this step, we solve a modified version of DMPP, MOD_DMPP, for ps and $\overline{UB_{ps}}$. The parameters of MOD_DMPP, is updated every time a project is added to the pool of selected projects. More detail on MOD_DMPP can be found in the next section. MOD_DMPP attempts to add a project that is given as an input and find the optimal scheduling for it based on availability of budgets at different times. The scheduling of the project is done optimally by solving the MOD_DMPP model using Xpress or any other commercial solvers.

We decide to accept (pass) the candidate project from step 1 and put it in the list of projects that would be executed or reject it based on the outcome of MOD_DMPP. A project ps is accepted if once the MOD_DMPP is solved we have $ub_{ps}^{*} = \overline{UB_{ps}}$. This means that project ps has fulfilled its potential and all of its phases would be implemented. Accepting a project is from hereon referred to *case a* throughout the rest of this section. If $ub^{*} \neq \overline{UB_{ps}}$, two other cases are possible:

Case b) $0 < ub_{ps}^{*} < \overline{UB_{ps}}$: In this situation, we do not accept the project as is. We add a mask of the project to the list of available projects with a change in its measure of favorability. Note that the change in the measure of favorability of this project would affect the probability of selecting all remaining projects.

Case c) $ub^*_{ps} = 0$: in this situation, we are not able to implement any portion of project ps at any time. So, we simply delete the project.

5.3.2.4 Step4 of heuristic- Removing: Deleting the selected project from step1 and returning back to step1

After step 3, we delete ps from the list of remaining projects. In this step, after one iteration, either the number of remaining projects have decremented by one or remained unchanged. It remains unchanged if in step 3, *case b* happens. In the other situations, the size of remaining projects is lessened by one. While the size of remaining projects is greater than 0, we loop and go to step 1.

Once no more projects remain, one major iteration is finished. We add up the individual benefits (differences in the objective function) made when a project was accepted and added. The value gained is the final overall profit for this major iteration. We can run another major iteration by changing the starting seed and starting off again with all projects.

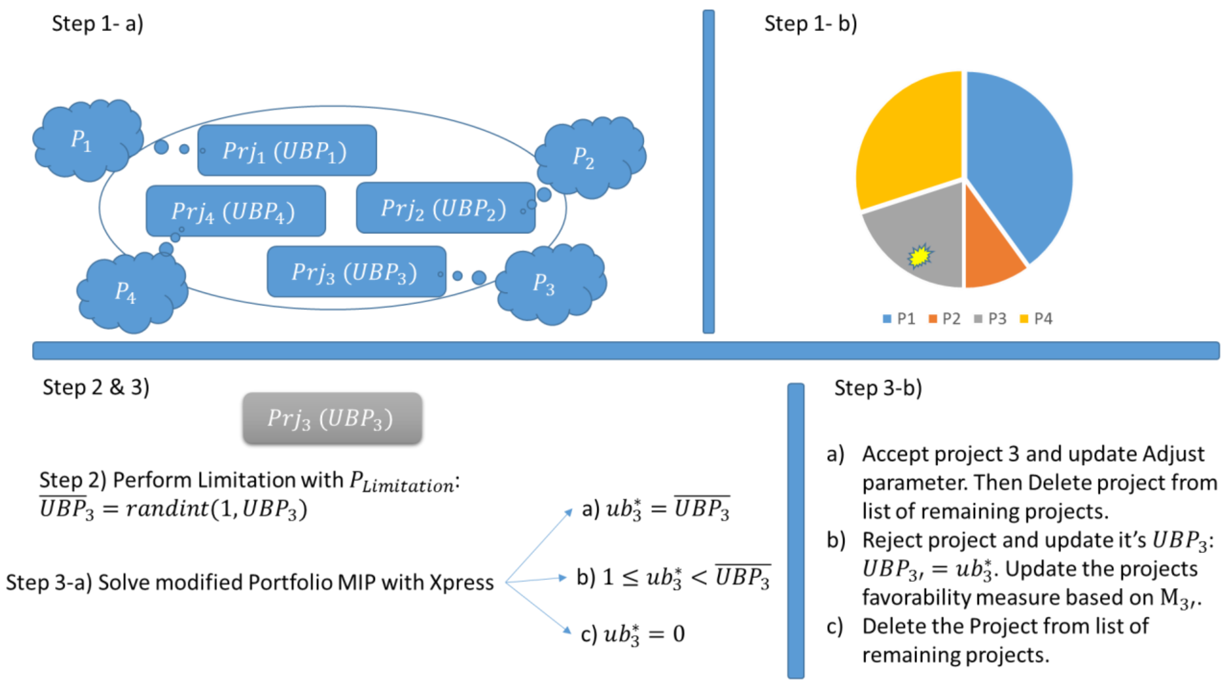

Figure 31 Steps 1 and 3 for heuristic for the case which project 3 gets is candidate selected from a list of 4 remaining projects

Figure 31 gives an illustration of steps 1-3. In this example, we start off step 1 with 4 projects. And during step 1, we select project 3 to be considered for being accepted. Figure 32 depicts 3 complete iterations of one major iteration of the heuristic. Note that the remaining pool of projects after an iteration is the pool of projects the next iteration starts with. In this example, in iteration 1, project 3 is selected for consideration. It is accepted (*case a*) and removed from the pool of available projects. In iteration 2, project 2 is chosen for consideration. In it not accepted as-is nor rejected (*case b*) and therefore a mask of the project, project 2', with modified UBP is added to the pool. In iteration 3, project 4 is picked for consideration. It is rejected (*case c*) and removed from the pool of available projects.

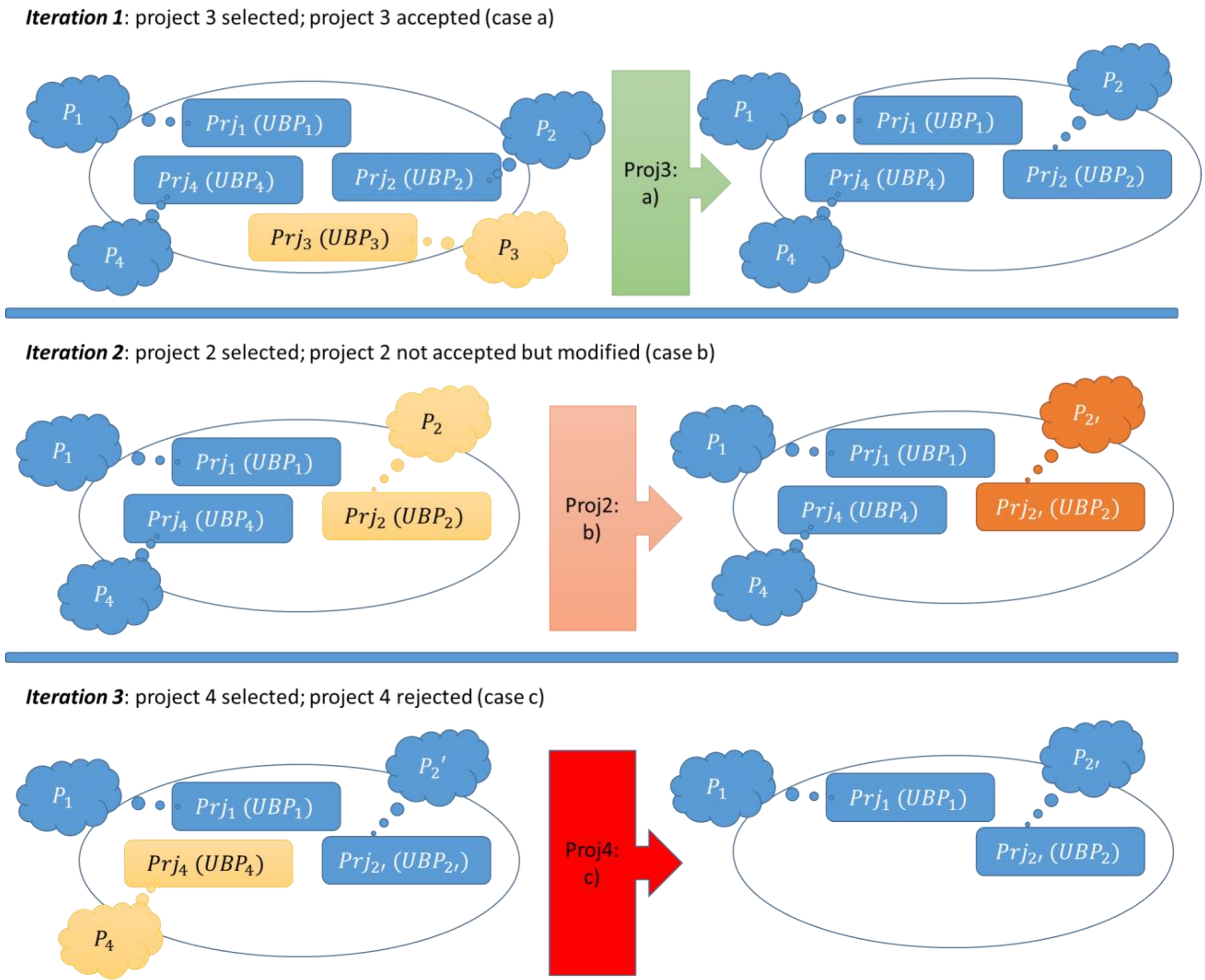

Figure 32 3 example iterations for heuristic

5.3.2.5 Summary pseudo code of the heuristic

The gist of one major iteration of the heuristic algorithm is summarized in the following pseudo-code:

- ❑ Calculate “favorability measure” for each project. i.e. M_p for each preprocessed project when it is done by itself.
- ❑ Set adjustment parameters for MOD_DMPP equal to 0.
- ❑ $remaining\_projects \leftarrow preprocessed\_projects$
- ❑ While $size(\ remaining\_projects\ ) \geq 1$

- Calculate the probability of selecting projects: $P_p = \frac{M_p}{\sum_{p \in remaining\_projects} M_p}$
- Select a project ps randomly using roulette wheel and it's probability of selection.
- With P_{Limit} Limit the UB for ps: $\overline{UB_{ps}} = randint(1, UB_{ps})$. If not limited: $\overline{UB_{ps}} = UB_{ps}$
- Solve the MOD_DMPP for ps with $\overline{UB_{ps}}$ to optimality using Xpress.
- If $ub^* = \overline{UB_{ps}}$:accept project and update adjustment parameters for MOD_DMPP.
- Else if $ub^* = 0$: the project will be deleted, do not update the adjustment parameters.
- Else: Add a mask of the ps, ps' to $remaining\_projects$ with $UB_{ps'} = ub^*$ and calculate its "favorability measure", $M_{ps'}$.
- Remove ps from list of remaining projects: $remaining\_projects.pop(ps)$

5.3.2.6 Modified version of DMPP, MOD_DMPP, used in the heuristic for DMPP

This modified version is for the execution of the heuristic. MOD_DMPP takes in one projects' input and some adjustment input parameters from its previous iterations. These input parameters modify the available budget at the beginning of each period. They also keep update of the objective function value. So, in total, we have $T + 1$ budget adjustment input parameters that updated after each iteration of the heuristic if the project being considered during that iteration is added to the list of selected projects.

The constraints for MOD_DMPP are mainly similar to the constraints of DMPP for one project. The only constraints that change are constraints (5-26)-(5-28) of DMPP. The modified constraints add an adjustment parameter to the RHS of the budget limitation constraints.

$$tcostub_{ps,0} + \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{ps,i,j} \times x_{ps,i,j}^0 + \sum_i OCPA_{ps,i,0} \times cn_{ps,i,0} \leq B0 + \boldsymbol{Adjustment_0} \quad (5\text{-}26')$$

$$tcostub_{ps,t} + \sum_i \sum_{j \geq i} (j - i + 1) \times PCOST_{ps,i,j} \times x_{ps,i,j}^t + \sum_i OCPA_{ps,i,t} \times cn_{ps,i,t} \times (1 + I)^{t - AvailableT_{ps}} \leq b_t + \boldsymbol{Adjustment_t} \qquad \forall t \geq 1 \quad (5\text{-}27')$$

We also relax the domain of b_t in (5-28) of DMPP and allow it to also take on negative values. This is allowed since, we are solving MOD_DMPP for each individual project. As a result, even-though the costs for a single project might exceed the initial budget, we are covering for that shortage in costs using the Adjustment parameter:

$$b_t \ is_{free} \qquad \forall t \quad (5\text{-}28'')$$

The budget adjustments, Adjustment, are updated once a project passes step 3 of the heuristic and is included in the selected projects list. The detailed updates are done using (5-31) and (5-32):

$$Adjustment(0) = Adjustment(0) - tcostub_0^* + \sum_i \sum_j (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{t\,*} + \sum_i OCPA_{i,0} \times cn_{i,0}^* \quad (5\text{-}31)$$

$$
\begin{aligned}
Adjustment(t) \\
&= Adjustment(t) + b_t^* \\
&- \left(tcostub_t^* \right. \\
&\left. + \sum_i \sum_j (j - i + 1) \times PCOST_{i,j} \times x_{i,j}^{t\ *} + \sum_i OCPA_{i,t} \times cn_{i,t}^* \right) \\
&\times (1 + I)^{t - AvailableT} - B_0 \times (1 + R)^t
\end{aligned}
\tag{5-32}
$$

In (5-31) and (5-32), all variables with an asterisk on top of them are the optimal solution of the problem of solving MOD_DMPP for the project at step 3 of the heuristic. If the project is not accepted in step 3, the Adjustments do not get updated.

5.4 Solved Example for the Deterministic Multi-project Phased-investment Problem (DMPP)

In order to illustrate the performance of the Preprocessing Step and the Heuristic, we generate two pools of projects. One that is relatively small and can be solved to optimality relatively fast using commercial solvers. The other is a large size problem. The constant parameters for the generated small size and large size problems are summarized in Table 22:

Table 22 Parameters for generated case examples

Parameter Name	Small Example	Large Example
Number of Projects	5	50
Time period (T)	10	60

Inflation rate, *I*	5%	5%
Risk free interest rate, *R*	2%	2%
Max UB for all projects	8	8

5.4.1 Potential savings from the pre-processing step

The savings from preprocessing for both the small and large cases can be seen from the summarization of the number of variables and constraints before and after preprocessing. The numbers reported in Table 23 are based on solving DMPP with Xpress.

Table 23 Number of variables and constraints before and after pre-processing step

Description	**Without (W/O) pre-processing**	**With (W) pre-processing**
	Large Example	
Number of vars after presolving by Xpress	31,107	10,933
Number of Constraints after presolving by Xpress	23,562	7,829
Remaining Number of Projects	50	33
	Small Example	
Number of vars after presolving by Xpress	321	273
Number of Constraints after presolving by Xpress	261	236
Remaining Number of Projects	5	4

As it can be seen, in the large case, the total number of variables and constraints after pre-processing is about 66% less than what it is before preprocessing. For the small case, the saving is around 13%. Nevertheless, the preprocessing step will never adversely affect the optimal solution and it is relatively fast as it is only solving the DMPP problem for just one project at a time.

5.4.2 The effect of favorability measure in the performance of the heuristic

In order to verify the performance of the heuristic, we ran the heuristic for the small case example using the simplest favorability measure for each project (*Profit*). Note that the Profit for each project by itself could be easily stored while performing the pre-processing step. After the pre-processing step, only the projects with Profit >0 remain. The optimal objective function value for the small case with $B0 = 5000$ was 12510.2. Using the Profit favorability measure for each project we were rarely able to reach this value with mutation probability of 0. This was mainly due to the setup of the small case. One of the projects that could have been started at the final time periods required a lot of funds. These funds would have become available if the prior projects were done. However since the profit of the latest project was relatively high, its favorability was high and it rarely was selected as the last project. As a result the project was rejected in most iterations due to unavailability of funds during that iteration.

Inspired by what we had learned, we considered an alternative favorability measure for the projects. This favorability was the ratio of Profit/LastAvTime for each project. LastAvTime is the LatestT parameter defined in Table 20. The two different favorability measures compared are summarized in Table 24.

Table 24 Two heuristics compared

Heuristic Name	**Heuristic 1 (H1)**	**Heuristic 2 (H2)**
Favorability Measure	Profit	Profit/LastAvTime

In order to capture the performance of each heuristic in different situations of availability of funds, we considered different levels of initial budget available. The different available budget tiers for each one of the small and large projects are summarized in Table 25.

Table 25 Different tiers of available funds for each case

Case	B0	Description
Small case example		
S-1	10,000	Large availability of funds tier
S-2	5,000	Medium availability of funds tier
S-3	3,200	Small availability of funds tier
Large case example		
L-1	40,000	Large availability of funds tier
L-2	20,000	Medium availability of funds tier
L-3	10,000	Medium to small availability of funds tier
L-4	5,000	Small availability of funds tier

We ran the heuristic for mutation probability equal to zero for the small case and mutation probability equal to 0.1 for the large case. In the small case, the averages are based on 20 iteration of the heuristics and for the large case, the averages are based on 5 iterations. Figure 33 and Figure 34 illustrate the results for the small and large case under different values of B0. As it can be seen, generally when the initial budget is relatively large, both heuristics perform more or less the same. The reason is that in this case, most projects can be implemented and there is no need to first implement the ones that come sooner to build funds for those which become available later. In smaller budget tiers, the order becomes important and heuristic 2 outperforms heuristic 1.

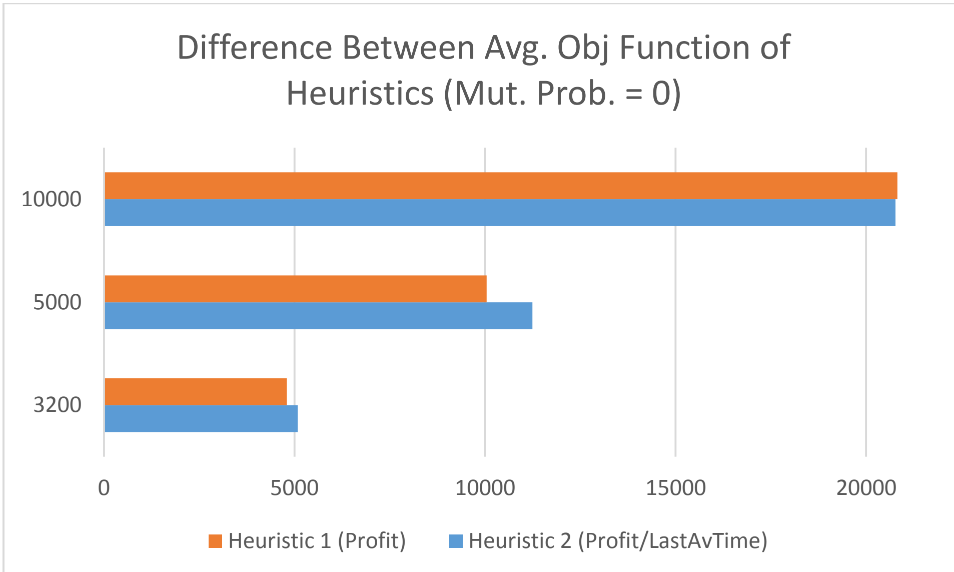

Figure 33 Difference between two heuristics - small case (different values of B0)

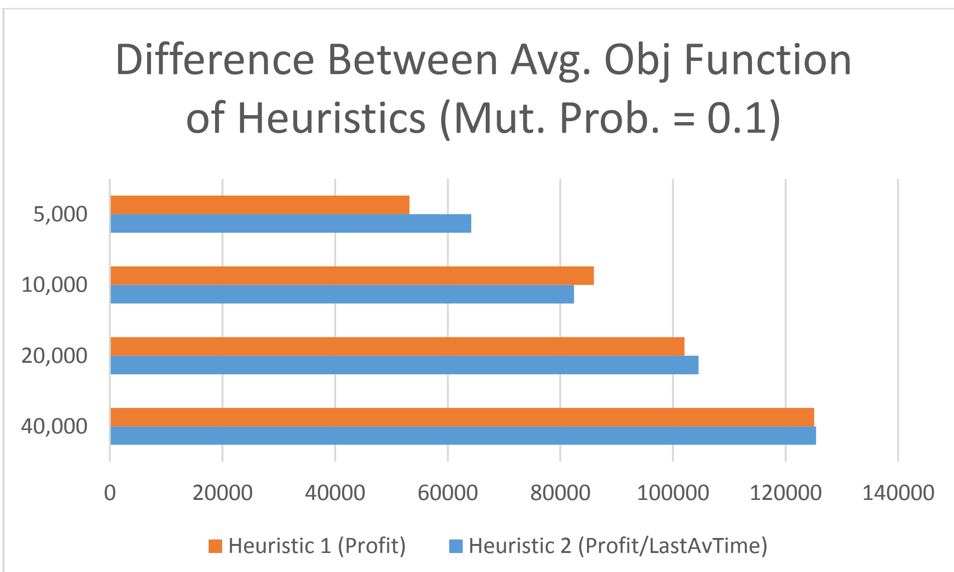

Figure 34 Difference between two heuristics - large case (different B0 values)

Based on these results, we decide to mostly use the second heuristic.

5.4.3 The effect of mutation probability – tuning the heuristic parameters

In order to verify the performance of the heuristic and tune the mutation probability parameter required for step 2 of the heuristic, we ran the heuristic for the small case example, which we were able to find fast optimal solutions for it. To illustrate the dependency on the mutation probability, we ran the second heuristic for the small case for various values of mutation probability and various values of initial investment funds (B0). For each of this pair of inputs, we ran the heuristic 20 times (20 iterations). The average of all these 20 iterations are used as a measure of performance. Figure 35 illustrates the summary of the results. As it can be seen, on average as we increase the mutation probability, the average objective function value decreases. There is only one exception to this and that is for B0=10,000. The best performance for the high budget tier is with mutation probability 0.2. It is worth noting that the optimal objective function value was not found only when the mutation probability was 1. Based on these results, we have concluded to use a mutation probability of 0.1 for the large case.

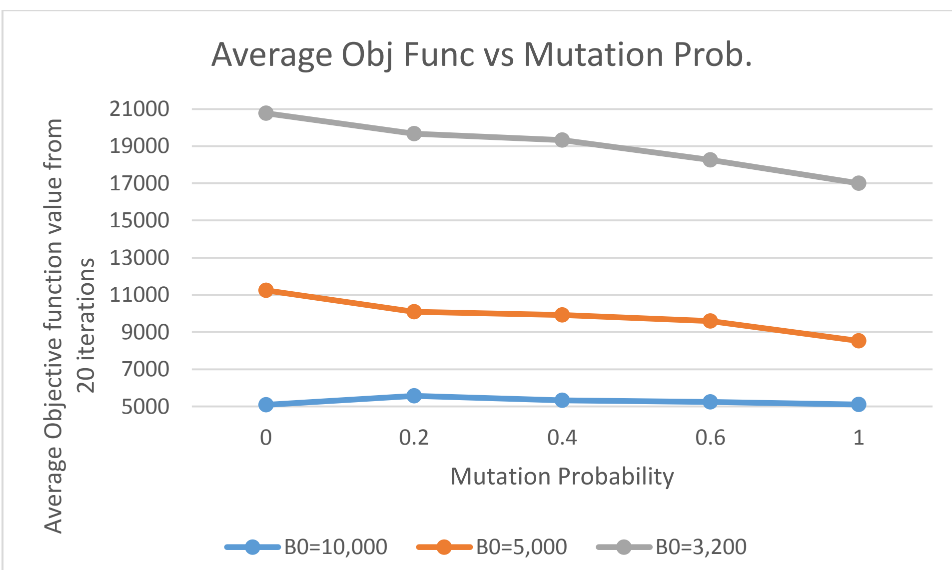

Figure 35 Average objective function value vs mutation probability for different cases of B0

5.4.4 The overall performance of pre-processing and heuristics – large case

The purpose of this section is showing the performance of the pre-processing step and the heuristic. To better illustrate the performances, we compare the results from different models/methods summarized in Table 26 for each availability of funds tier.

Table 26 Different methods/models used for illustrating the effects of pre-processing and the heuristics for the large case

Model name
Solving MADASD using Xpress w/o preprocessing (XW)
Solving MADASD using Xpress w preprocessing (XP)
Solving MADASD using Heuristic 1 (Favorability measure: Profit) (H1)
Solving MADASD using Heuristic 2 (Favorability measure: Profit/LastAvTime) (H2)

Figure 36 summarizes the results for 5 different iterations of the heuristics 1 and 2 and the results from XPRESS with and without pre-processing when the initial available budget (funds) is relatively high (40,000). Each iteration result for the heuristics are independent. The Xpress results are also not proven to be optimal as we have stopped the run after a minimum of 2 hours. The objective function value illustrated is the best solution found by XPRESS within the time limit imposed. As it can be seen, the heuristics (H1 and H2) almost always beat XW. However, throughout the five iterations shown, H1 only beats XP once. When you take into account the running times, for each iteration and the running times of XW and XP, the solutions found using the heuristics are very fast good solutions. Based on the stats in Table 27, the average running time for the heuristics are approximately 1 minute as opposed to the 5 hours used for finding the best solution found for XP and XW. A fairer comparison for H2 was to run the heuristic at least 300 times (approximately 5 hours) and compare the best solution among the 500 iterations to the one found from XP (127,453). To verify that H2 can beat XP, we ran H2 for more iterations. H2 was able to find a better solution after 5 more additional iterations. At iteration 10, the solution found by H2 was 129,358.

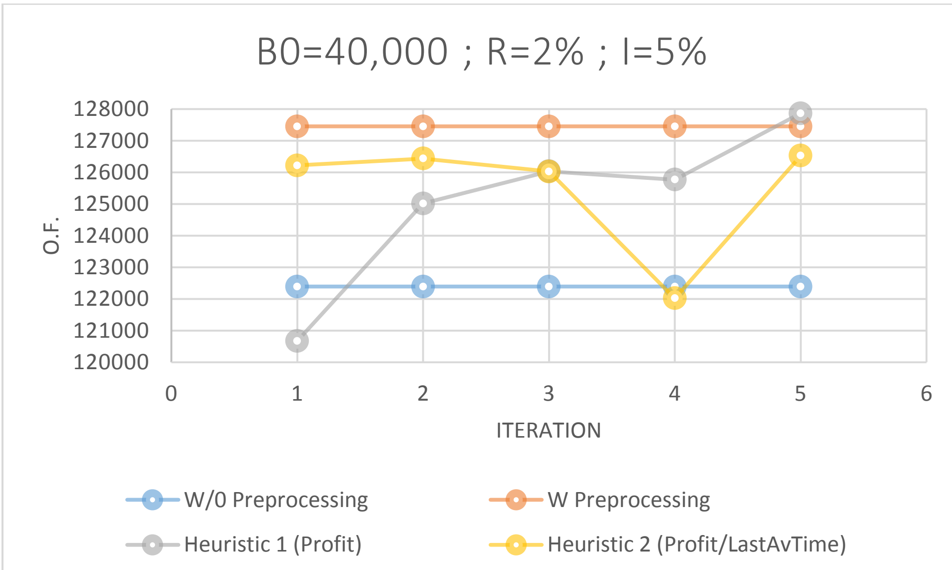

Figure 36 Performance of different methods for B0=40,000

Table 27 Average execution time (s) for finding the reported best solution of different methods (B0=40,000)

Method	**Avg. Running time (time reported best solution found) (secs)**
Xpress without Preprocessing (XW)	18,042.20
Xpress with Preprocessing (XP)	17,659.49
Average running time for H1	64.59
Average running time for H2	56.64

Figure 37, Figure 38, and Figure 39 illustrate the performance (best solutions found) of the different methods for different tiers of available budget at time 0, $B0 = 20{,}000$, $B0 = 10{,}000$, and $B0 = 5{,}000$, respectively. Again, both H1 and H2, beat XP and XW at least during one of the 5 iterations of the heuristics. H2 has its best performances in the lowest available budget tier (Figure 39).

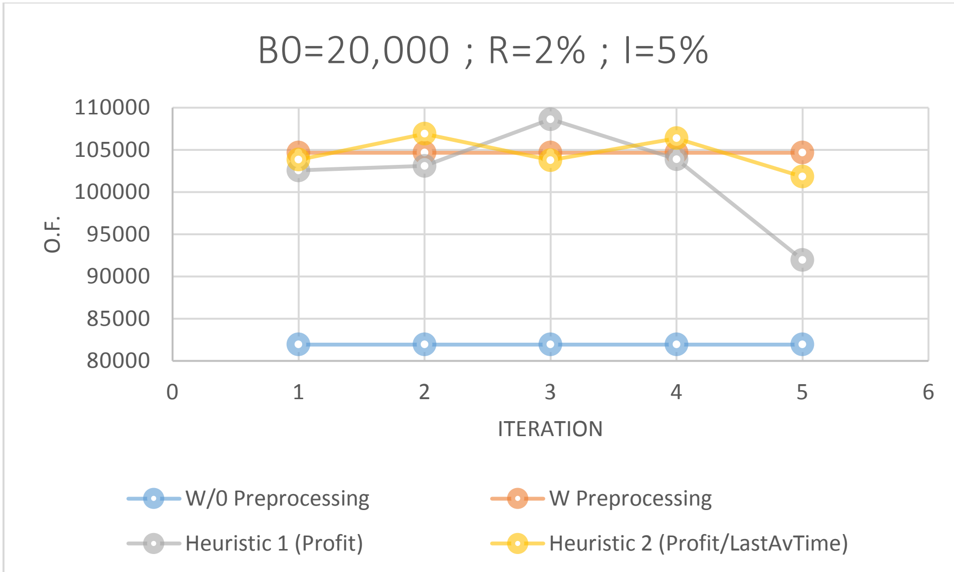

Figure 37 Performance of different methods for B0=20,000

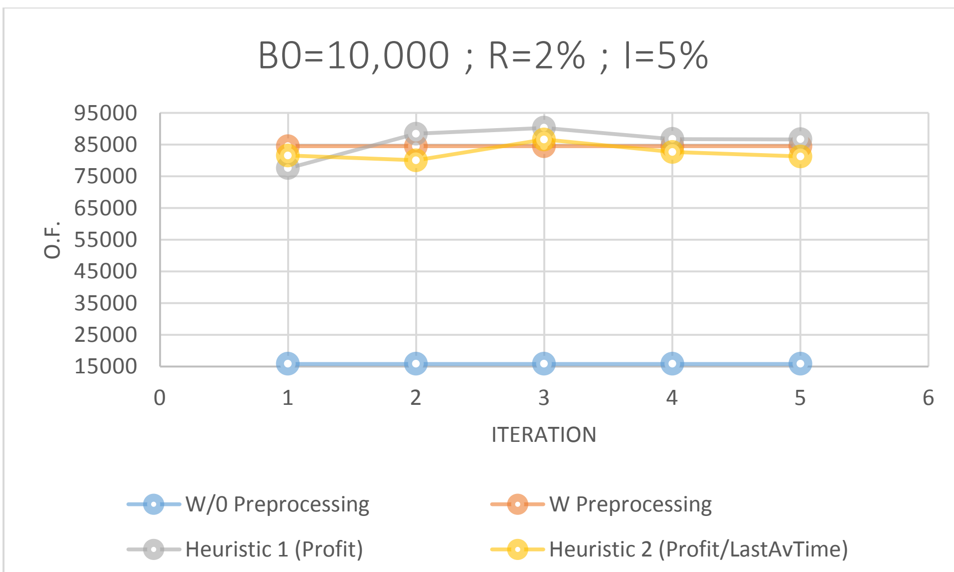

Figure 38 Performance of different methods for B0=10,000

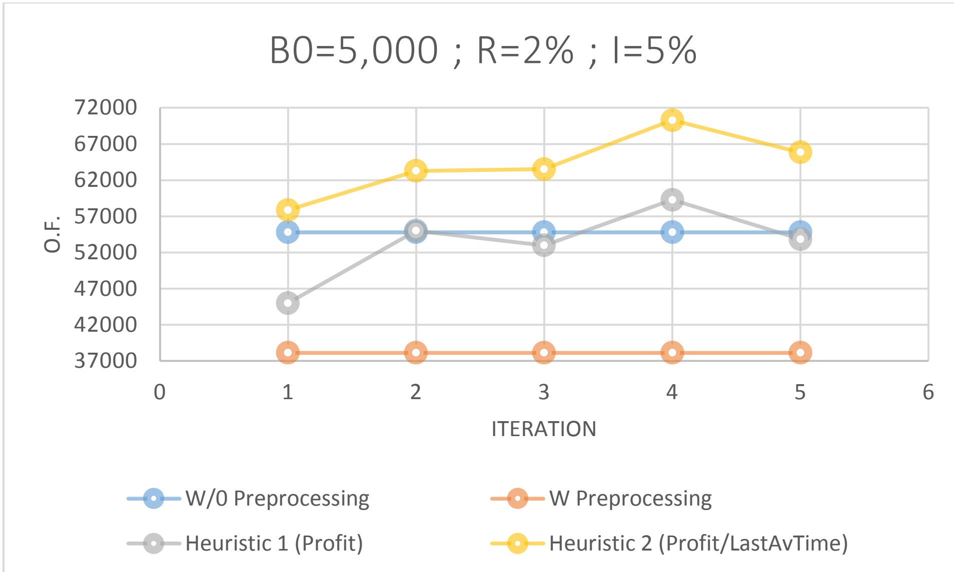

Figure 39 Performance of different methods for B0=5,000

The average running time of the iterations and the running time for XP and XW are summarized in Table 28. Note that XP and XW are not solved to optimality and the solutions reported are the best solutions found at the time stated in Table 28.

Table 28 Time in which best sols are found for different methods and different B0s - large case

B0	Method best sol time (s)			
	XW	XP	H1	H2
B0=20,000	7,455.4	7,164.0	73.5	48.8
B0=10,000	6,899.4	7,161.5	68.4	40.9
B0=5,000	7,411.0	6,032.7	57.86	39.2

By comparing the results from Figure 36-Figure 39, we can see the estimated gain in the objective function value from additional funds. For example, the best objective function value for $B0 = 5{,}000$ is 70,275.5. The best objective function value for $B0 = 10{,}000$ is 90,314.1. This means that an increase of 5,000 in the initial budget can cause an estimated increase of 20,000 in the objective function (marginal profit of 15,000). Also, the best result for $B0 = 20{,}000$ is 106,921. Therefore another additional 10,000 increase in initial budget would only increase the objective function by about 15,000 (marginal profit of 5,000). Lastly, the best solution for $B0 = 40{,}000$ is 129,358. This is only an estimated 22,000 increase in objective function (2,000 marginal increase). Therefore, we can conclude based on the available budget, a small amount of loan can potentially increase the profit by a lot if the initial budget is small. If the initial budget is larger, the marginal benefit lessens.

Chapter 6: Summary, Conclusions, and Future Directions

6.1 Summary

In this dissertation, we presented different models for finding the optimal grouping of phases and scheduling of phased investment projects. Each model is tailored for a specific situation. The first model that is the backbone of all other models as well, is the Deterministic Single-project Phased-investment Problem (DSPP). DSPP captures cost and time interdependencies among phases of the project. The model is used for finding the optimal decisions for a parking garage construction example. We show that via sensitivity analysis, managers can gain insight about how to improve the overall system. Some of the contributions of the novel MIP in this part of this dissertation are: Considering time and cost interdependencies among the phases of the project; considering construction duration (implementation duration) for each phase; and accounting for the initial infrastructure size and cost required for future development of phases.

The second model is a multi-stage scenario based MIP for the Stochastic Single-project Phased-investment Problem (SSPP). The uncertainties are modeled using the trinomial lattice framework. Due to the modeling framework for uncertainty, the problem size increases exponentially with the number of stages. We present two solution methods for finding a good value for the here and now decision variables for the SSPP problems. The first method is finding the here and now decision variable for all scenarios using the Solve-Search-Delete (SSD) algorithm. The second method is finding the here and now decision variable value using simulation combined with optimization. We apply the SSPP model to the parking garage construction example. The contributions of this part of the problem are: Presenting a novel multi-stage stochastic optimization model for phased investment that accounts for time and cost interdependencies among phases, and initial infrastructure requirements; Solving for every scenario of the huge stochastic problem using a novel solve-search-delete algorithm and also finding optimal solutions with given confidences by solving the problem using the combination of simulation.

The third and last model is the Deterministic Multi-project Phased-investment Problem (DMPP). This model simultaneously solves the project selection and scheduling problem and also considers interdependencies among the phases. Re-investment is also permitted – similar to DSPP and SSPP. To solve this problem, we present a pre-processing step that helps reduce the size of the problem. We also propose a heuristic that proceeds by adding projects one by one and adjusting the available budget at different times accordingly. The order of adding projects to the list of accepted projects depends on the selected measure of favorability. Some of the contributions of the DMPP are: Presenting a novel mathematical model that can assist managers in making simultaneous selection and scheduling decisions for a pool of projects built from some potentially sequential phased projects that require initial infrastructure investments and have synergies among phases. The model allows transfer of funds between different stages of time. It also allows for abandonment of projects already implemented. In addition the starting time of the projects' implementation are allowed to be flexible. The other contributions are related to the process of finding solutions: A preprocessing step is presented that can reduce the number of variables and constraints without any compromises in terms of the objective function value; A novel heuristic is developed that optimizes the problem for each project at a time.

6.2 Conclusions and Findings

Each model is applicable to different settings. The DSPP model is most applicable to cases that we are planning for a single project and the planning period is short or the information that we have about the future are accurate and certain. It is also usefeul for cases that we are interested in doing thorough sensitivity analysis for a phased project since the running time of the model is very small. For example, in the parking garage example, under certain set of input environmental parameters such as risk-free interest rate and inflation rate, an increase in the current available budget even though as small as thousands of dollars can cause an increase in profit of about 1 million dollars.

The SSPP model is most useful for planning for a single phased project in which the planning period is long or the environment is very uncertain. In stochastic settings, the problem size dramatically increases. We need to use various solution techniques that in a way decreases the size of the problem. This could be achieved by either decomposition of the overall problem or reducing the number of scenarios through a sampling method.

We show that based on the values of the input parameters, the SSD algorithm could decrease the running time required for solving all scenarios by more than 90%. Also, we show how for the optimization via simulation case, the use of common random numbers (CRN), improves the running time and increases the confidence of finding the approximate solution over the case that we do not use CRN.

The SSD algorithm's speed depends on which scenarios are solved during the solve step. The number of scenarios deleted per each scenario solved are the greatest for the here-and-now decision variable values that are the best possible and worst possible. This is because of the initialization step of the algorithm that we selected the best and worst scenario to solve.

The DMPP is most applicable in settings that we are selecting among a pool of projects that some of which could be projects that could be broken into smaller phases. The problem, even under deterministic settings, is very huge and finding solutions for it is difficult.

Via an example, we show that, by looking at each project individually first, we can reduce the size of the problem. The savings gained by reducing the size of the problem (up to 60% less variables and constraints for large-size problems).

For the heuristic, selection of an appropriate measure of favorability for the projects, can greatly influence the performance of the heuristic. We show that in cases that the initial budget is little, the favorability measure of Profit/LastAvTime performs better than the favorability measure of Profit. For cases that the initial budget is relatively large, both measures perform similarly. The run time for these heuristics is about 1 minute for large-size examples. The solution found is most of the times better than the solution found by commercial solvers after even 2 hours. This saving in time, allows managers to perform multiple sensitivity analysis on the parameters. We show that if the initial budget is small, a marginal increase in the initial budget can greatly improve the overall profits in comparison to the case that the initial budget is medium or large.

6.3 Future Studies

This research opens the venue to many future studies related to phased investment. Some possible research direction are:

- Development of a model for Stochastic Multi-project Phased-investment Projects,
- Considering more sources of uncertainty in the SSPP model
- Improving the efficiency of the SSD algorithm, via parallel processing
- Adding other types of interdependencies to the model and model the interdependencies among the projects as well in the DMPP model
- Perform sensitivity analysis on all the parameters of the models
- Comparing the performance of the heuristic when other favorability measures are considered for the selection step of the heuristic
- Comparing the outcomes of the heuristic to those from Meta-heuristics such as GA, SA, and TS.

References

Badri, M. A., Davis, D. & Davis, D., 2001. A comprehensive 0–1 goal programming model for project selection. *International Journal of Project Management*, 19(4), pp. 243-252.

Bain, R., 2009. Error and optimism biad in toll road traffic forecasts. *Transportation,* Volume 36, pp. 469-482.

Baldwin, C. Y., 1982. Optimal sequential investment when capital is not readily reversible. *The Journal of Finance*, 37(3), pp. 763-782.

Ballou, D. P. & Tayi, G. K., 1996. A Decision Aid for the Selection and Scheduling of Software Maintenance Projects. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on,* 26(2), pp. 203-212.

Bar-Ilan, A. & Strange, W. C., 1998. A model for sequential investment. *Journal of Economic Dynamics and Control*, 22(3), pp. 437-463.

Belenky, A. S., 2012. A Boolean programming problem of choosing an optimal portfolio of projects and optimal schedules for them by reinvesting within the portfolio the profit from project implementation. *Applied Mathematics Letters,* 25(9), pp. 1279-1284.

Beraldi, P. & Bruni, M. E., 2014. A clustering approach for scenario tree reduction: an application to a stochastic programming portfolio optimization problem. *TOP*, 22(3), pp. 934-949.

Bhattacharyya, R., Kumar, P. & Kar, S., 2011. Fuzzy R&D portfolio selection of interdependent projects. *Computers & Mathematics with Applications*, 62(10), pp. 3857-3870.

Black, F. & Scholes, M., 1973. The Pricing of Options and Corporate Liabilities. *The Joural of Political Economy*, 81(3), pp. 637-654.

Brandai, L. E., Bastian-Pinto, C., Gomes, l. & Labes, M., 2012. Govenrment Supports in Public-Private Partnership Contracts: Metro Lane 4 of Sao Paulo Subway System. *Journal of Infrastructure Systems,* 18(3), pp. 218-225.

Brandao, L. E., Dyer, J. S. & Hahn, W. J., 2005. Using Binomial Decision Trees to Solve Real-Option Valuation Problems. *Decision Analysis,* 2(2), pp. 69-88.

Brucker, P., Knust, S., Schoo, A. & Thiele, O., 1998. A branch and bound algorithm for the resource-constrained project scheduling problem. *European Journal of Operational Research,* 107(2), pp. 272-288.

Carazo, A. F. et al., 2010. Solving a comprehensive model for multiobjective project portfolio selection. *Computers & operations research,* 37(4), pp. 630-639.

Carvalho, P. M., Ferreira, L. A. F. M., Lobo, F. G. & Barruncho, L. M. F., 1998. Optimal distribution network expansion planning under uncertainty by evolutionary decision convergence. *International Journal of Electrical Power & Energy Systems,* 20(2), pp. 125-129.

Chan, W. T., Chua, D. K. & Kannan, G., 1996. Construction resource scheduling with genetic algorithms. *Journal of Construction Engineering and Management,* 122(2), pp. 125-132.

Chassiakos, A. P. & Sakellar, S. P., 2005. Time-Cost Optimization of Construction Projects with Generalized Activity Constraints. *Journal of Construction Engineering and Management,* 131(10), pp. 1115-1124.

Chen, W.-N.et al., 2010. Optimizing Discounted Cash Flows in Project Scheduling—An Ant Colony Optimization Approach. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on,* 40(1), pp. 64-77.

Coffin, M. A. & TAYLOR III, B. W., 1996. R&D project selection and scheduling with a filtered beam search approach. *IIE transactions,* 28(2), pp. 167-176.

Coffin, M. A. & Taylor, B. W., 1996. Multiple criteria R&D project selection and scheduling using fuzzy logic. *Computers & Operations Research,* 23(3), pp. 207-220.

Cruz, C. O. & Marques, R. C., Submitted 2013. Exogenous Determinants for Renegotiation of Public Infrastructure Concessions. *Journal of Construction Engineering and Management.*

Cui, Q., Bayraktar, M. E., Hastak, M. & Minkarah, I., 2004. Use of Warranties on Highway Projects: A Real Option Perspective. *Journal of Management in Engineering,* 20(3), pp. 118-125.

Cui, Q., Johnson, P., Quick, A. & Hastak, M., 2008. Valuing the Warranty Ceiling Clause on New Mexico Highway 44 Using a Binomial Lattice Model. *Journal of Construction Engineering and Management,* 134(1), pp. 10-17.

De Neufville, R., Scholtes, S. & Wang, T., 2006. Real options by spreadsheet: parking garage case example. *Journal of Infrastructure Systems,* 12(2), pp. 107-111.

Dekker, S. & Verhaeghe, R. J., 2008. Development of a strategy for port expansion: an optimal control approach. *Maritime Economics & Logistics,* 10(3), pp. 258-274.

Dickinson, M. W., Thornton, A. C. & Graves, S., 2001. Technology portfolio management: optimizing interdependent projects over multiple time periods. *Engineering Management, IEEE Transactions on,* 48(4), pp. 518-527.

Dixit, A. K. & Pindyck, R. S., 1994. *Investment Under Uncertainty.* s.l.:Princeton University Press.

Dorndorf, U., Pesch, E. & Phan-Huy., T., 2000. A branch-and-bound algorithm for the resource-constrained project scheduling problem. *Mathematical Methods of Operations Research,* 52(3), pp. 413-439.

Dupačová, J., Gröwe-Kuska, N. & Römisch, W., 2003. Scenario reduction in stochastic programming. Mathematical programming. *Mathematical programming,* 95(3), pp. 493-511.

Eberly, J. C., 2008. Irreversible investment. In: S. N. Durlauf & L. E. Blume, eds. *The New Palgrave Dictionary of Economics.* 2nd ed. s.l.:Palgrave Macmillan.

Feng, Y. & Ryan, S. M., 2013. Scenario construction and reduction applied to stochastic power generation expansion planning. *Computers & Operations Research,* 40(1), pp. 9-23.

Fine, C. H. & Freund, R. M., 1990. Optimal Investment in Product-Flexible Manufacturing Capacity. *Management Science,* 36(4), pp. 449-466.

Geske, R., 1979. The valuation of compound options. *Journal of Financial Economics,* 7(1), pp. 63-81.

Gonçalves, J. F., Mendes, J. J. & Resende, M. G., 2008. A genetic algorithm for the resource constrained multi-project scheduling problem. *European Journal of Operational Research,* 189(3), pp. 1171-1190.

Growe-Kuska, N., Heitsch, H. & Romisch, W., 2003. *Scenario reduction and scenario tree construction for power management problems.* Bologna, IEEE.

Gupt, S. K. & Rosenhead, J., 1968. Robustness in sequential investment decisions. *Management Science,* 15(2), pp. B-18 - B-29.

Hartmann, S. & Briskorn, D., 2010. A survey of variants and extensions of the resource-constrained project scheduling problem. *European Journal of Operational Research,* 207(1), pp. 1-14.

Hartmann, S. & Kolisch, R., 2000. Experimental evaluation of state-of-the-art heuristics for the resource-constrained project scheduling problem. *European Journal of Operational Research,* 127(2), pp. 394-407.

Heitsch, H. & Römisch, W., 2003. Scenario reduction algorithms in stochastic programming. *Computational optimization and applications,* 24(2-3), pp. 187-206.

Herath, H. S. & Park, C. S., 2002. Multi-stage capital investment opportunities as compound real options. *The Engineering Economist,* 47(1), pp. 1-27.

Huang, X. & Zhao, T., 2014. Project selection and scheduling with uncertain net income and investment cost. *Applied Mathematics and Computation,* Volume 247, pp. 61-71.

Icmeli, O., Erenguc, S. & Zappe, C. J., 1993. Project Scheduling Problems: A Survey. *International Journal of Operations & Production Management,,* 13(11), pp. 80-91.

Iniestra, J. G. & Gutierrez, J. G., 2009. Multicriteria decisions on interdeendent infrastructure transportation projects using an evolutionary-based framework. *Applied Soft Computing,* pp. 512-526.

Jafarzadeh, M., Tareghian, H. R., Rahbarnia, F. & Ghanbari, R., 2015. Optimal selection of project portfolios using reinvestment strategy within a flexible time horizon. *European Journal of Operational Research,* 243(2), pp. 658-664.

Killen, C. P. & Kjaer, C., 2012. Understanding project interdependencies: The role of visual representation, culture and process.. *International Journal of Project Management,* 30(5), pp. 554-566.

Kim, J. L. & Ellis Jr, R. D., 2008. Permutation-Based Elitist Genetic Algorithm for Optimization of Large-Sized Resource-Constrained Project Scheduling. *Journal of construction engineering and management,* 134(11), pp. 904-913.

Kolisch, R. & Hartmann, S., 2006. Experimental investigation of heuristics for resource-constrained project scheduling: An update. *European journal of operational research,* 174(1), pp. 23-37.

Kolisch, R. & Padman, R., 2001. An integrated survey of deterministic project scheduling. *Omega,* 29(3), pp. 249-272.

Kruger, N. A., 2012. To Kill a Real Option - Incomplete Contracts, Real Options and PPP. *Transportation Research Part A: Policy and Practice,* 46(8), pp. 1359-1371.

Lamar, B. W. & Lee, A., 1999. A strategic investment model for phased implementation of flexible manufacturing systems. *International Transactions in Operational Research,* 6(3), pp. 331-344.

Lee, H. W., Choi, K. & Gambatese, J. A., 2014. Real options valuation of phased investments in commercial energy retrofits under building performance risks. *Journal of Construction Engineering and Management,* 140(6).

Lee, J. W. & Kim, S. H., 2000. Using analytic network process and goal programming for interdependent information system project selection.. *Computers & Operations Research,* 27(4), pp. 367-382.

Lee, J. W. & Kim, S. H., 2001. An integrated approach for interdependent information system project selection. *International Journal of Project Management*, 19(2), pp. 111-118.

Lieberman, M. B., 1987. Market Growth, Economies of Scale, and Plant Size in the Chemical Processing Industries. *Journal of Industrial Economics,* 36(2), pp. 175-191.

Liesiö, J., Mild, P. & Salo, A., 2008. Robust portfolio modeling with incomplete cost information and project interdependencies.. *European Journal of Operational Research,* 190(3), pp. 679-695.

Lim, S.-K. & Kim, Y.-D., 1998. Capacity planning for phased implementation of flexible manufacturing systems under budget restrictions. *European Journal of Operational Research,* 104(1), pp. 175-186.

Liu, S.-S. & Wang, C.-J., 2011. Optimizing project selection and scheduling problems with time-dependent resource constraints. *Automation in Construction,* pp. 1110-1119.

Mansini, R., Ogryczak, W. O. & Speranza, M. G., 2014. Twenty years of linear programming based portfolio optimization. *European Journal of Operational Research,* 234(2), pp. 518-535.

Miller, L., Choi, S. H. & Park, C. S., 2004. Using an options approach to evaluate Korean information technology infrastructure. *The Engineering Economist,* 49(3), pp. 199-219.

Ott, S. H., Hughen, W. K. & Read, D. C., 2012. Optimal phasing and inventory decisions for large-scale residential development projects. *The Journal of Real Estate Finance and Economics,* 45(4), pp. 888-918.

Park, H., Shafahi, A. & Haghani, A., 2016. A Stochastic Emergency Response Location Model Considering Secondary Incidents on Freeways. *IEEE Transactions on Intelligent Transportation Systems.*

Park, T., Kim, B. & Kim, H., 2013. Real Option Approach to Sharing Privatization Risk in Underground Infrastructures. *Journal of Construction Engineering and Management,* 139(6), pp. 685-693.

Pedrasa, . M. A., Spooner, E. D. & MacGill, I. F., 2011. *Robust scheduling of residential distributed energy resources using a novel energy service decision-support tool.* s.l., IEEE, pp. 1-8.

Pennings, E. & Lint, O., 2000. Market entry, phased rollout or abondonment? A real option approach. *European Journal of Operational Research,* 124(1), pp. 125-138.

Prastacos, G. P., 1983. Optimal sequential investment decisions unde conditions of uncertainty. *Management Science,* 29(1), pp. 118-134.

Razali, N. M. & Hashim, A. H., 2010. *Backward reduction application for minimizing wind power scenarios in stochastic programming.* s.l., IEEE, pp. 430-434.

Rocha, K. et al., 2007. Real estate and real options - A case study. *Emerging Markets Review,* 8(1), pp. 67-79.

Santhanam, R. & Kyparisis., G. J., 1996. A decision model for interdependent information system project selection. *European Journal of Operational Research,* 89(2), pp. 380-399.

Schmidt, R. L., 1993. A Model for R & D Project Selection with Combined Benefit, Outcome and Resource Interactions. *IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT,* pp. 403-410.

Sefair, J. A. & Medaglia, A. L., 2005. *Towards a model for selection and scheduling of risky projects.* s.l., s.n., pp. 158-164.

Shafahi, A., 2012. *Bidding on projects based on previous works and eminence, A contractors' view point.* s.l.:s.n.

Shafahi, A. & Haghani, A., 2013. *A Linearization Approach for Project Selection with Interdependencies in Resource Costs.* Barcelona, s.n., pp. 230-235.

Shafahi, A. & Haghani, A., 2014. Modeling contractors' project selection and markup decisions influenced by eminence. *International Journal of Project Management,* 32(8), pp. 1481-1493.

Sharma, K. C., Jain, P. & Bhak, R., 2013. Wind power scenario generation and reduction in stochastic programming framework. *Electric Power Components and Systems,* 41(3), pp. 271-285.

Shayanfar, E., Samimi Abianeh, A., Schonfeld, P. & Zhang, L., 2016. Prioritizing Interrelated Road Projects Using Metaheuristics. *Journal of Infrastructure Systems,* 22(2).

Siahkali, H. & Vakilian, M., 2010. Stochastic unit commitment of wind farms integrated in power system. *Electric Power Systems Research,* 80(9), pp. 1006-1017.

Sun, H. & Ma, T., 2005. A packing-multiple-boxes model for R&D project selection and scheduling. *Technovation,* 25(11), pp. 1355-1361.

Tao, X. & Schonfeld, P., 2006. Selection and Scheduling of Interdependent Transportation Projects with Island Models. *Transportation Research Record: Journal of the Transportation Research Board,* pp. 133-141.

Tao, X. & Schonfeld, P., 2007. Island Models for Stochastic Problem of Transportation Project Selection and Scheduling. *Transportation Research Record: Journal of the Transportation Research Board,* pp. 16-23.

Tofighian, A. A. & Naderi, B., 2015. Modeling and solving the project selection and scheduling. *Computers & Industrial Engineering,* pp. 30-38.

Xing, Z., Chen, Y. & Zhang, W., 2006. *Optimal planning by decomposed satisfiability and backward reduction.* s.l., s.n., pp. 53-56.

Zhang, H., Li, H. & Tam, C. M., 2006. Particle swarm optimization for resource-constrained project scheduling. *International Journal of Project Management,* 24(1), pp. 83-92.

Zhao, T., Sundararajan, S. K. & Tseng, C.-L., 2004. Highway development decision-making under uncertainty: A real options approach. *Journal of Infrastructure Systems,* 10(1), pp. 23-32.

Zhao, T. & Tseng, C.-L., 2003. Valuing flexibility in infrastructure expansion. *Journal of infrastructure systems,* 9(3), pp. 89-97.

Zuluaga, A., Sefair, J. A. & Medaglia, A. L., 2007. *Model for the Selection and Scheduling of Interdependent Projects.* s.l., IEEE, pp. 1-7.