Stochastic Dominance for Project Screening and Selection under Uncertainty
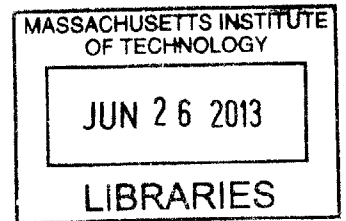
by

Adekunle M. Adeyemo

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2013
[June 2013]

Author . . . .
Department of Chemical Engineering
February 27 2013

Certified by. . . . .
Gregory J. McRae
Hoyt C. Hottel Professor of Chemical Engineering (Emeritus)
Thesis Supervisor

Accepted by . . . .
Patrick S. Doyle
Chairman, Department Committee on Graduate Theses

Stochastic Dominance for Project Screening and Selection under Uncertainty

by

Adekunle M. Adeyemo

Abstract

At any given moment, engineering and chemical companies have a host of projects that they are either trying to screen to advance to the next stage of research or select from for implementation. These choices could range from a relative few, like the expansion of production capacity of a particular plant, to a large number, such as the screening for candidate compounds for the active pharmaceutical ingredient in a drug development program. This choice problem is very often further complicated by the presence of uncertainty in the project outcomes and introduces an element of risk into the screening or decision process.

It is the task of the process designer to prune the set of available options, or in some cases, generate a set of possible choices, in the presence of such uncertainties to provide recommendations that are in line with the objectives of the ultimate decision maker. Screening and decision rules already exist that do this but the problem with most of them is that they add more assumptions to the structure of the preferences of the decision maker, or to the form of the uncertain distribution that characterizes the project outcome, than is known at the time. These challenges may lead to the screening out of viable alternatives and may ultimately lead to the selection of inferior projects.

This thesis aims to demonstrate the applicability of Stochastic Dominance as method that can overcome these obstacles. Stochastic Dominance has been shown to be a general method for incorporating risk preferences into the decision-making process. It is consistent with classical decision theory, it makes minimal assumptions of the structure of the utility functions of the decision makers and of the nature of the distributions of the uncertainty and under certain conditions can be shown to be equivalent to the other objectives.

In this work, an up-to-date review and an implementation framework for Stochastic Dominance is presented. The performance of the method relative to some of the other screening and decision objectives is examined in the light of three case studies: the design of a reactor-separator system for the production of a chemical, the selection of a crop for biomass production and the design of a biomass to liquids process. The limitations of the method are also discussed together with suggestions for how they can be overcome to make the method more effective.

3

Thesis Supervisor: Gregory J. McRae
Title: Hoyt C. Hottel Professor of Chemical Engineering (Emeritus)

Acknowledgments

It is said that it takes a village to raise a child. I have found, in the past 5+ years, that it may take a bit more to produce a PhD. A lot of people have, over the past five-and-a-half years, been essential to the successful completion of my thesis and I'd like to take some time to express my sincere appreciation for their investment in my life.

First I'd like to thank my adviser, Gregory McRae, for teaching me how to think like a scientist and a problem solver. I appreciate the patience he showed me during the learning process and I'm grateful for all the wisdom and insights on life I got in our numerous non-research conversations.

I thank my committee members, Richard de Neufville and Charles Cooney, for always having time on their calendar for me and for the insightful questions and feedback they had on the direction of the work. This work is so much better because of their input.

I'd like to thank my family for the support and love they showed me throughout the process. Mummy, Oyin and Yinka - thank you for constantly reminding me that I am loved and cherished, regardless of the outcome of research (and life, for that matter) and for believing in me especially in the times when circumstances caused me to really doubt myself. Thank you for the phone calls, the emails and text messages, and for being quick to forgive my failure to respond on time.

Of the many friends I made at MIT, four stand out above the rest for the sheer depth and contribution they made to my life in graduate school - Debo Ogunniyi, Po-Ru Loh, Sam Perli and Stephanie Lam. Thank you so much for sharing your lives with me and for showing me what loving and serving another person looks like. To the Ogunniyi family - thank you for welcoming me into their home and their lives. I can now say I have family here in the US and I thank you for being that family to me.

The MIT Graduate Christian Fellowship, the Sidney-Pacific bible study group and Cambridgeport Baptist Church together formed my faith home for the past five and-a-half years and I am grateful for the many friends I made who kept me focused on the more important things in life, especially during frustrating periods of research. I enjoyed the Friday night talks, the Toscanini's visits, Sunday Brunch, the long talks into the night, the dinners and lunches, retreats at Toah Nipi, the planning of events and, most of all, the warmth, joy and the beauty of a place of belonging in a group of people with whom I shared not just an academic experience, but a faith as well. Many of the conversations we had were played a big role in my growth and transformation and I appreciate everyone who was a part of these groups.

I made good friends among the my ChemE cohort - the 2007 entering class - and I appreciated the times we spent together, both in the early years when we were navigating the quals, practice school and adviser selection and in the later years of surviving the mid-PhD slump. I want to especially thank Jaisree Iyer, Emily Chang, Woo Keung and Pedro Valencia for their companionship throughout the time. In addition to the cohort, I also appreciate Tatyana Shatova, Jonathon Harding and Matt Stuber for being great co-TAs in 10.302 and 10.551 and for making the experience a

fun one.

I am the last of a host of graduate students privileged to have been a part of the McRae Group and I want to thank the few of them that I was fortunate to share the experience with - Ken Hu, Bo Gong, Arman Haidari, Anusha Kothandaraman, Ingrid Berkelmans, Carolyn Seto, Mihai Anton, Alex Lewis, Dr. Duan and Chuang-Chung Lee. Thank you for the fun days, group meetings, good conversations and the camaraderie that arose from sharing the basement of 66 - especially during the floods! - all of which formed a core part of my graduate research experience. I thank our assistant, Joan Chisholm, for her care and attention, most especially when I became the sole member of the group. Joan, I wish you all the best in your retirement.

Richard Braatz and his group occupied the lab shortly after my adviser retired and he was kind enough to let me keep my cubicle for as long as I remained a graduate student. Beyond that, I appreciate his including me in many of the activities of the group - outings, seminars and group meetings. The rest of the Braatz group - Xiaoxiang Zhou, Lifang Zhu, Mo Jiang, Ashley Ford-Versypt, Joe Scott, Kwang-ki Kim, Ali Mesbah, Jouha Min, Mark Molaro and Lucas Foguth - became my daily companions and friends and I also thank them for filling the void that the graduation of the rest of my previous group made, and, more than that, adding a colour of their own to my experience.

The congenial nature of the department, is, in my opinion, largely unrivalled at MIT and it owes a lot to the staff, present and past, many of whom became friends. Suzanne Maguire, Joel Dashnaw, Fran Miles, Katie Lewis, Iris Chen, Mary Wesolowski, Christine Preston, Alina Haverty, Alison Martin, Jean Belbin, Jim Hardsog, Melanie Miller, Rebecca Hailu, Rosangela Santos, Gwen Wilcox, Barbara Balkwill and Beth Tuths - thank you all so much for the dedication to the department that you display daily, and for making time for random graduate students with constant requests (I won't mention any names).

I was a program assistant for the MIT Summer Research Programme (MSRP) for two summers and I got to meet and know an amazing group of people in the Office of the Dean of Graduate Education. I want to especially thank the coordinators of the programme, Christopher Jones and Monica Orta, both of whom became good friends and were always willing to provide a listening ear and/or give an encouraging hug and in general, be resources and wellsprings of care and support in school.

I spent all but the last 4 weeks of the programme living in Ashdown House, the oldest (and newest) graduate dorm at MIT. My housemasters, Ann and Terry Orlando, as well as the house manager, Denise Lanfranchi, were more than I could have asked for. My roommates, Sam and Wilt, made coming back to the apartment evening feel like coming home (and not just to a place to sleep) and our 'other roommates' Jen and Albert, also contributed to the atmosphere we enjoyed.

Graduate education is not cheap and a number of sources provided the funding that made it possible. I am grateful to the department of Chemical Engineering, the MIT Energy Initiative, Eni S.p.A (the Italian energy company) and George M. Keller, the Haas Family and the Practice School for fellowships.

Different technologies contributed significantly to my learning at MIT. In particular, I'd like to thank Google Inc. for their amazing search engine that always provided

a (million) starting point(s) for my inquiries into any given subject; Wikipedia for (often) providing a readable summary of the area/field I was investigating and helping me identify key terms that I needed to know to better grasp the ideas. My adviser was long-distance for the duration of my PhD and so I am grateful for VoIP technology definitely made the costs of advising much cheaper as we could talk for hours without having to think about costs.

Last and most of all, I'd like to thank God Almighty for his grace and goodness to me. Thank you for saving me, making me yours, giving me the opportunity of a lifetime to study at such an amazing school and making the journey more than just an academic one, but a spiritual one too. My life belongs to you and I look forward to walking the rest of it with and following hard after you.

Contents

List of Figures

13

14

List of Tables

19

Chapter 1

A New Approach to Project Selection and Design

1.1 Introduction

At any given moment, engineering and chemical companies have a host of projects that they are trying to select from for implementation. These choices could range from a relative few to a large number. For example, the choice of expanding the production capacity for a chemical usually consists of two alternatives - increase the capacity of an existing plant or build a new one. One the other hand, for every drug that makes it to eventual market, pharmaceutical companies often begin with a large number of possible molecules to investigate. In both these cases, the firms are often constrained by a budget, expectations of their stockholders and the market (for publicly held firms), or the desires of owner/board that makes the choice problem real and precludes them from the pursuing all the options placed before them.

However, perhaps the biggest challenge in making project decisions is the uncertain nature of the outcomes of the projects. The success of a new drug for example depends on its efficacy versus existing treatments, passing clinical trials and FDA approval and while that of a new carbon-reducing technology on other things like future price of crude oil, uncertainty in demand etc. This uncertainty introduces an element of risk - the possibility of failing to meet set goals - into the decision process and managing this risk effectively by wisely selecting (or designing) projects for execution that are in line with the overall goals and preferences of the firm and its stakeholders is what successful firms learn to do in practice.

Project design and selection is not an new problem and as such several approaches have been developed to address the issue. Firms use expected value measures, key performance indicators, weighted average cost of capital and a host of other measures that are created to help with this decision. The problem with many of them though is that they come with assumptions on preferences of the decision-maker or the structure of the uncertainty present in the projects for them to make good choices among different uncertain options. These assumptions are often unstated/implicitly assumed when the rules are used and so can give rise to a mismatch between the desires

and attitudes of the decision maker and the final project choice, thus increasing the likelihood that poor project decisions will be made to the possible detriment of the firm.

This thesis presents a new approach to project selection/design under uncertainty that reduces/eliminates this danger. This new approach is able to incorporate preferences of decision makers to the degree that it is known and to use these to effectively screen projects and ultimately help select among them. It doesn't make assumptions on the nature of the distributions of the uncertainty as well and as such has a broader application scope than many of the currently used methods.

With this approach, it is our hope that decision-makers will now be able to better explore the decision options available and can make quality decision, improving resource allocation within the firm that ultimately benefits their stakeholders and possibly the community at large.

1.2   Motivating Problem

Increasing concern about the rate of use of fossil fuels to power today's global economy as well as the security issues that go hand in hand with the increase in fossil fuel use has led a lot of countries to invest heavily in the development of renewable, greener sources of energy in the near future to replace the ubiquitous fossil fuels.

Consider an energy firm considering the portfolio of projects to invest in over the next project period. Such a firm faces a set of options that can be categorized into three large groups as follows:

- The choice of raw/starting materials

- The choice of chemical/material transformation

- The choice of final product(s) to manufacture

Examples of these are illustrated in Figure 1-1. In this structure, a project choice will correspond to a path through the network. For example, the construction of a solar farm will be grouped in the nature-mechanical/electrical-electricity category, the production of industrial chemicals from crude oil using microbes (find reference?) will be classified in the fossil fuels-biological-fuels/chemicals category and production of ethanol and other fuels using gasification and gas-to-liquids technology will be classified in the biomass-thermochemical-fuels category.

Even within each category just described, it is possible for there to be yet another level of choice. If the thermochemical transformation of biomass described above is selected, for example, there are several different decisions that have to be made regarding the particular choice (or choices) of biomass feed, gasifier type, catalysts for the Fischer-Tropsh reactors as well as the product mix that could result in vastly different design structures. These different options are highlighted in Figure 1-2 where a typical process diagram for a biomass-to-liquids plant is illustrated together with a list of possible choices for the different sections.

22
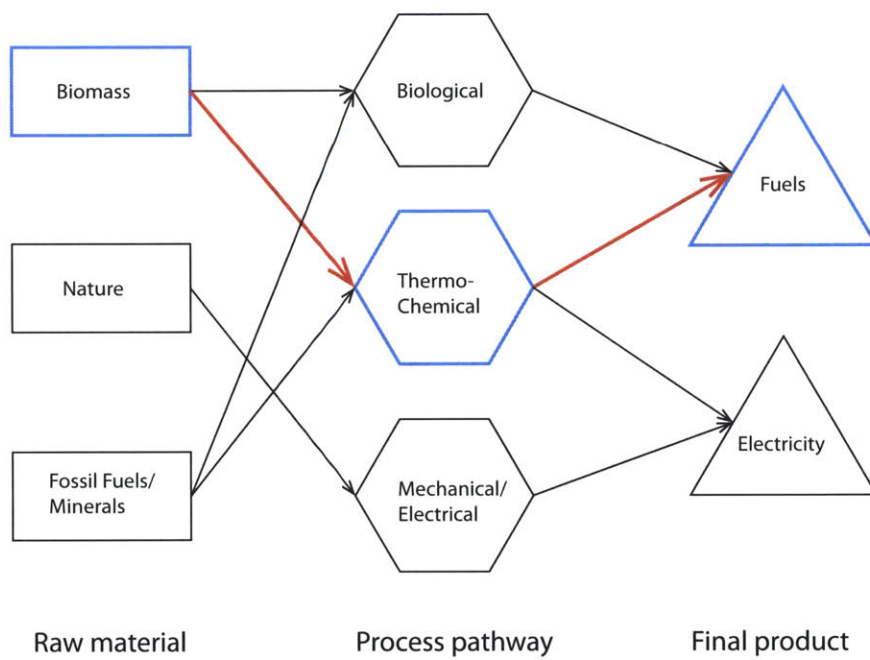
Figure 1-1: An illustrative network for investment in energy. The leftmost column gives examples of raw materials, the middle for process transformation options and the rightmost column for possible products
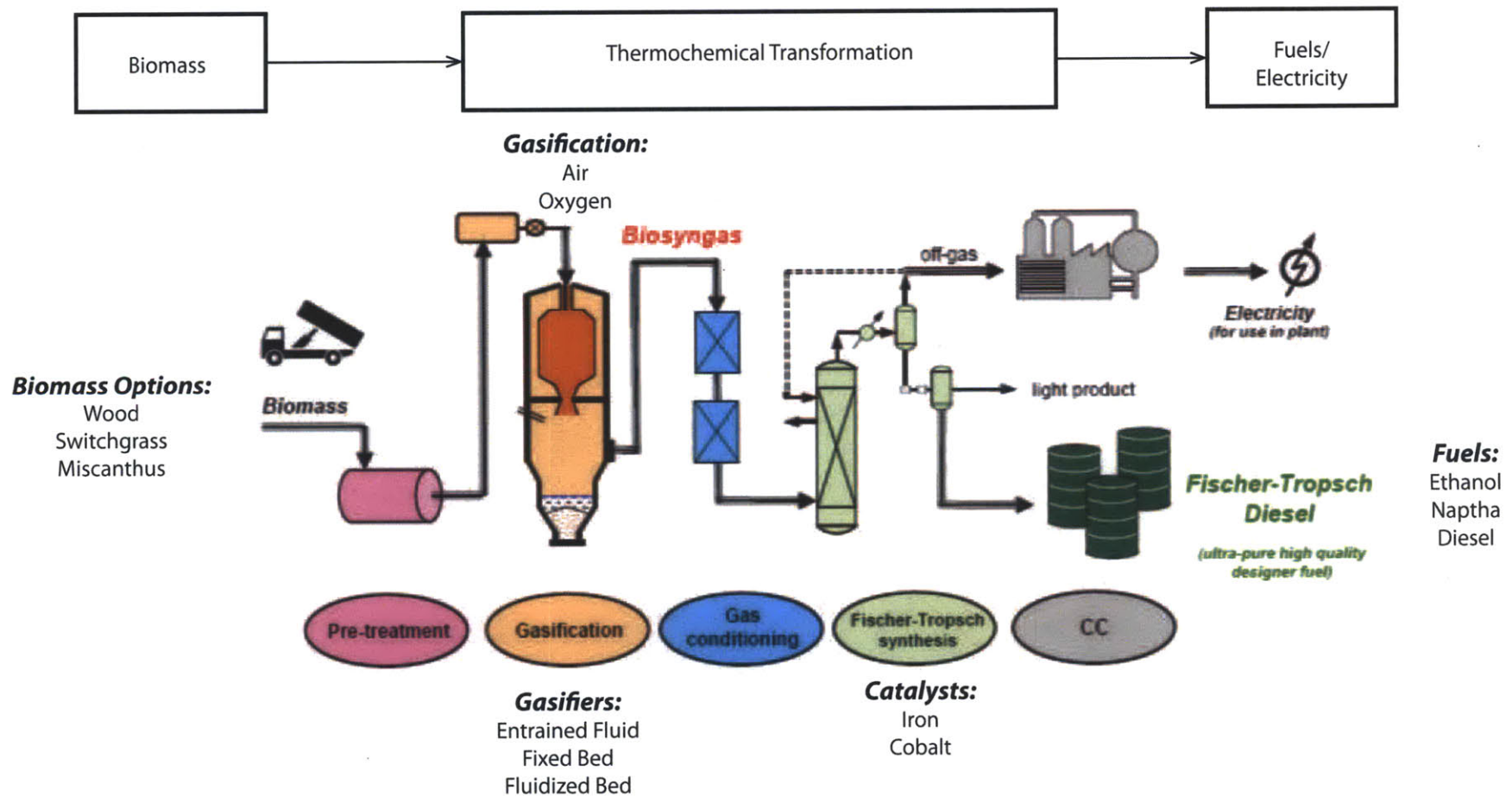
Figure 1-2: Expansion of project choices from Figure 1-1. Text shows the different options embedded in each choice of raw material, process pathway and final product selected. Adapted from [17]
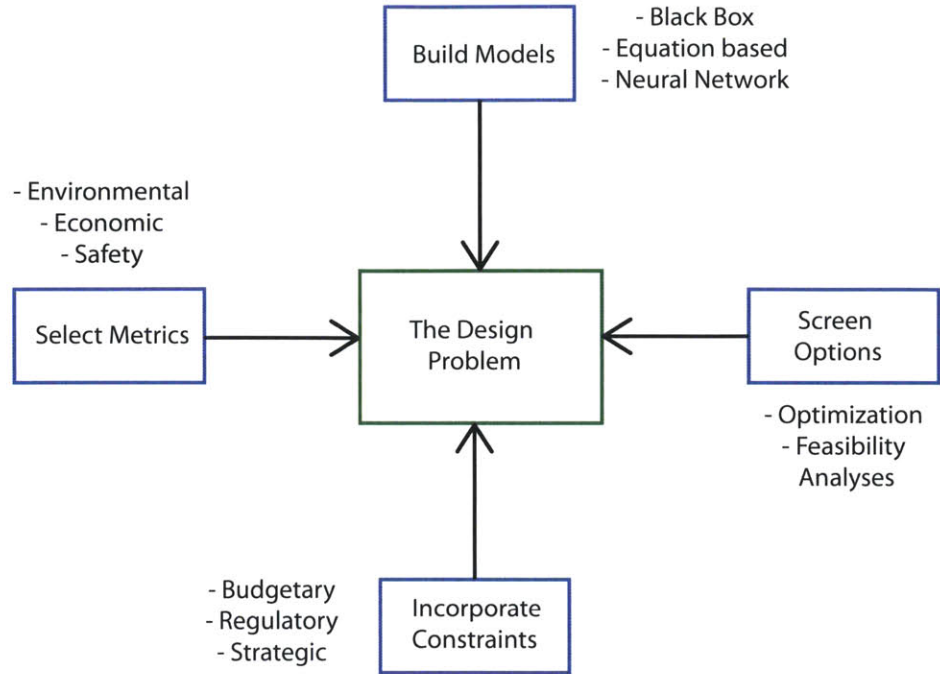
Figure 1-3: The main elements of the design problem

1.3 Process Design and Uncertainty

Process design is essentially the art of systematically examining these alternatives to eventually recommend a project or a set of projects for either further investigation or for investment. We breakdown the design problem into into four main sub-problems outlined below and illustrated in Figure 1-3:

1. Building models/pilot plants

2. Selection of metrics

3. Evaluation and comparison of alternatives

4. Incorporation of constraints

To fully understand the processes that we ultimately intend to build/implement, it is necessary to understand them and this is often achieved through the use of models/pilot plants. The understanding of the science and engineering behind the common processes in Chemical Engineering - heat transfer, separation, fluid transport, etc - together with an increase in the availability of cheap computing power has led to the growth and popularity of mathematical models to represent chemical processes and today, special software like Aspen [3], help us to quickly build process flow diagrams representing complex chemical systems.

Evaluation and comparison of the numerous possible configurations that can result from a design is a big challenge. To do this efficiently, a *superstructure* containing all

the possible project choices (cite or display image) is often constructed and effective algorithms [28], [16], [41] created for the efficient examination of these options.

Decisions among the many possible configurations requires some objective criteria for ranking and the traditional metrics of choice have usually been economic [116], with an eye towards process reliability and safety [105]. However, increasing concern about global warming and sustainable energy has led the environmental considerations to be an important metric in process design [21], [30] and it is the job of the designer to use these using different (and sometimes incommensurable) metrics to make decisions.

Finally, the designer has to be aware of different regulations and constraints that guide the design of the plant as these often constrain the set of possible designs. There are often economic constraints like the budget, and the minimum acceptable return for the investment; environmental/regulatory constraints that determine emission levels hazardous waste materials and land use or other zoning police constraints that affect the siting of plants as well as technological constraints (second law efficiency, equilibrium concentrations in reactors etc) that limit the extent of what can be obtained.

All of these have comprised the classic challenges that have affected process design and are largely addressed in some fashion in many of the standard books on chemical engineering design today [112], [116], [103]. In the presence of a single metric (or a weighted combination of a number of different problems), the design problem becomes an optimization problem which can be represented in general as

$$
\begin{aligned}
& \min f(x) \\
s.t.\ & g(x) \leq 0 \\
& h(x) = 0 \\
& x \in D
\end{aligned}
\tag{1.1}
$$

where D is the feasible set of the design problem and can incorporate the fact that X includes integer and continious variables. Graphically, the optimization problem in two dimensions is illustrated in Figure 1-4. The area enclosed by the blue curves is the feasible region and the red curves illustrate the (non-linear) objective function which is increasing in the north-eastern direction. The optimum occurs where there is no increase in the objective function that results in a feasible point. This point is marked by the red dot and is the design point selected. More discussion on optimization and techniques for finding optimal solutions can be found in the appendix.

However, one major challenge that gets little treatment in many of these is the presence of uncertainty in the design. The incorporation of uncertainty into the design process affects all the elements mentioned above as is highlighted in Figure 1-5 and ultimately expands the design problem in each of the elements. With respect to building models, uncertainty in our knowledge of the process kinetics poses the challenge to design adequate experiments to narrow the range of our uncertainty, the

26
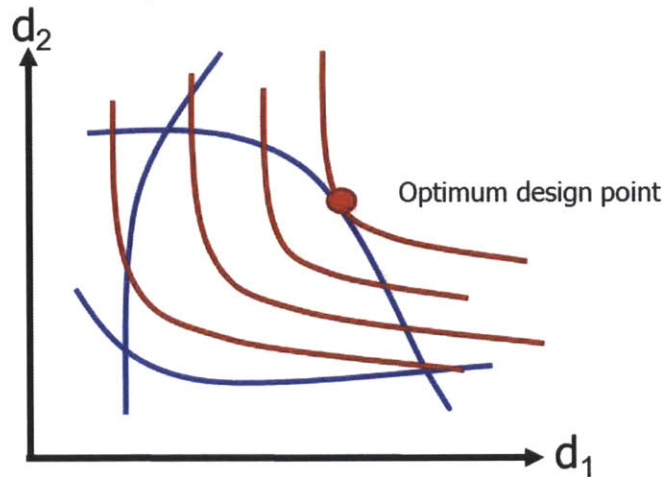
Figure 1-4: Graphical illustration of optimization problem for classical design

selection of the best (family of) models that incorporate the most of what we know about the project at the present time and the ability to tune this models to best represent the cases at hand by the proper determination of the models parameters.

Uncertainty in the model inputs and the model structure in turn tends to affect the model outputs that we care about (profit, safety and sustainability) and to quantify the extent to which this is the case, in addition to determining which variables/parameters have the largest effect on the outputs, we need to be able to propagate the uncertainties through the models (uncertainty analaysis) and decompose the uncertainty to the contributions of the different inputs via a sensitivity analysis. In addition, as mentioned earlier, uncertainty introduces risk into the decison process and it is necessary that the selection metrics are constructed so that they properly reflect the riskiness of each project.

Finally, uncertainty impacts the ability to meet regulations and constraints that projects have to meet to be implemented. Failure to do this may result in financial, failure to operate or, even more serious, the loss of life. Designs that fail to properly account for this by taking uncertainty into are likely to be unsuccessful or, if implemented, disasters to the firm.

With respect to the solution of the design problem earlier illustrated, the presence of uncertainty makes the solution of an optimum now ill-defined as is seen Figure 1-6. As can be seen, the objective functions and the constraints are no longer unique and a range of optima will now be identified as opposed to a unique optimum in the earlier case.

In tackling this problem, there was the introduction in literature of what we will henceforth in this thesis refer to as **objectives**, new metrics that condense the uncertain problem to the form in 1-4 that can now be solved via traditional techniques. An example of such an approach is the use of expected value where the uncertain design problem is rewritten with expectations of the objective function and constraints used to formulate the problem.
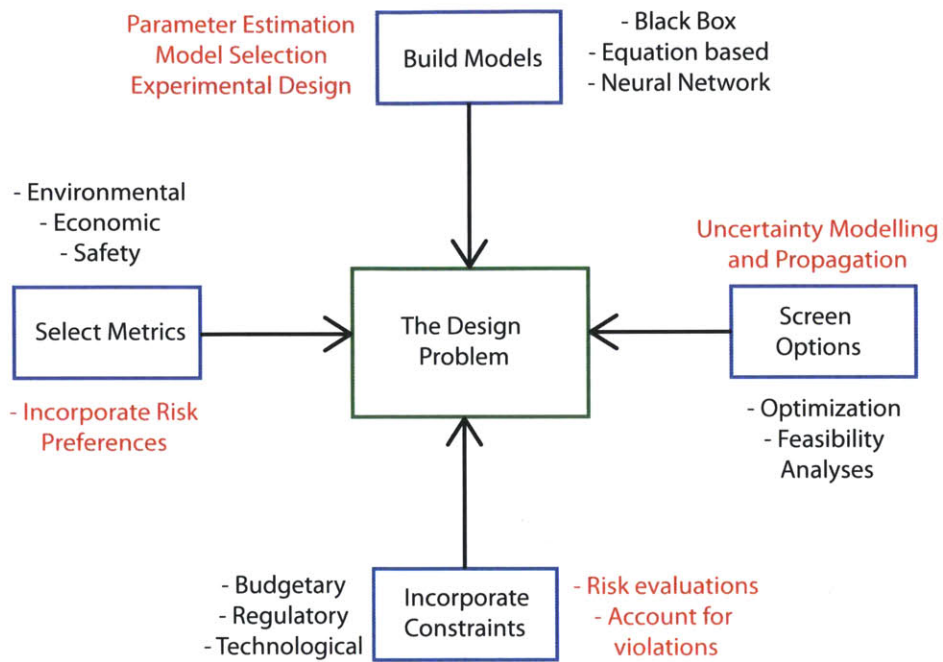
Figure 1-5: The main elements of the design problem with the introduction of uncertainty
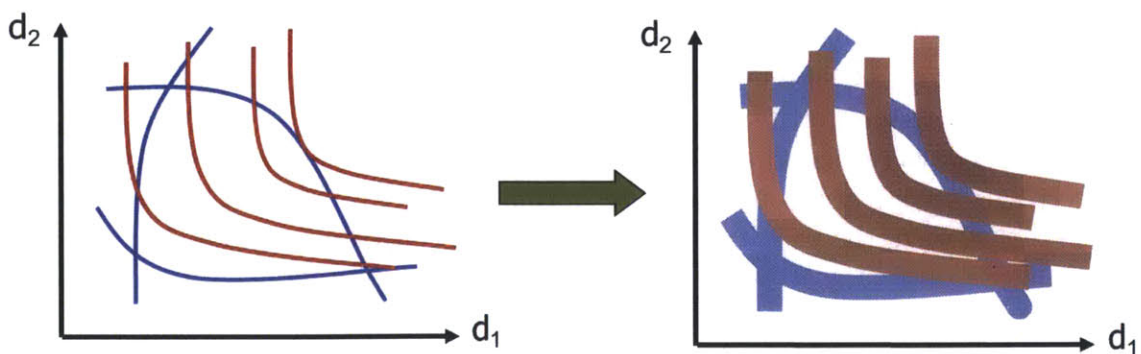
Figure 1-6: The effect of uncertainty on the design problem
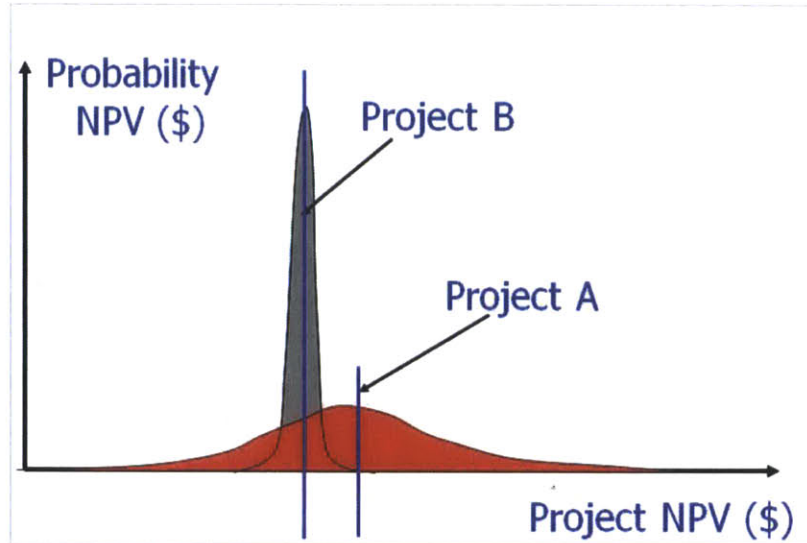
Figure 1-7: An illustration of two projects with different distributions

The big challenge of these 'objectives' are the assumptions they layer on the preferences of the decision maker as well as the requirements they sometimes have on the nature of the distribution of the uncertain outcome for them to be optimal. Take the two projects illustrated in Figure 1-7.

Project A will be selected in a choice between these two using the expected value metric (expected values are indicated by the vertical blue lines running through the middle of both projects). However, if real decision makers are presented with this choice, the selection will not be so clear cut as there will be some that will tend to prefer project B even though it has a smaller expected value. This is because it has a smaller spread of outcomes and as such Given these reasons, the overall effect of uncertainty on the design problem is huge and multi-faceted and research has gone into tackling many of these areas, much of which will be highlighted in subsequent chapters in the thesis. The main focus of this work however is in the proper incorporation of risk and risk preferences into project selection metrics (in particular, profitability) and this is described some more in the next section.

1.4   Representing and Incorporating Preferences

Once all the relevant uncertainties mentioned above and introduced into the design/selection problem, it is often the case that the decision maker now has to select from project choices that are represented not by point estimates but by a distribution of outcomes.

It is the task of the process designer to prune the set of available options (or in some cases, *generate* a set of possible choices) in a way manner is in line with the objectives of the ultimate decision maker. As mentioned earlier, decision rules have been in use that do this and are very much in use in industry as a method for selection. The problem with these is that they usually add more assumptions to

the structure of the preferences of the decision maker, or the form of the uncertain distribution that characterizes the project outcome, than is known at the time of the decision. Take, for example, the expected value measure assumes that the decision maker is unconcerned about risk, or the traditional mean-variance rule, embodied by the capital-asset-pricing-model, assumes that uncertainty in the project is either normally distributed or the decision makers have a very specific type of risk preference. For the decision-maker who is concerned about accounting for risk in his project selection choice, the expected value measure will prune out decisions that will likely have been more preferred by him, in the end offering him a suboptimal project. The same goes for the decision-maker faced with projects that are non-normal in their outcome distributions.

To use preferences in decision making, it is necessary to know them and be able to represent them in a form that is amenable to analysis. With complete information, it is possible to represent the particular preference of the decision maker by a function that can be used in doing decision analysis. This level of information is however often lacking or difficult to obtain/measure and we end up with only partial information usually of a general nature. The framework we develop in this thesis is adapt to increasing levels of knowledge of the decision maker's preferences and can use more knowledge to further prune the set of available decisions. In every case, the assumptions being made are transparent so that our rules are only as sharp as they need to be - no more, no less.

1.5   Thesis objectives

The objectives of this thesis are as follows:

1. **To present Stochastic Dominance as a viable framework for project selection and process design under uncertainty.**

2. **Illustrate how it relates to other known objectives and adequately handles risk preferences in decision making under uncertainty .**

3. **Demonstrate its application via the use of a number of case-study examples and compare with other metrics presently in use.**

4. **Demonstrate ways in which it can be further refined to give unique decisions when needed.**

1.6   Scope of Thesis

The alternatives that are analyzed in this thesis are mutually exclusive and so portfolio combinations of these alternatives are not examined. Dynamic decision making, where modifications to the project can be made after initial implementation and uncertainty is realized is also not considered . The proposed method at present only handles discrete options, so continuous variables in the problem structure are discretize before

any analysis is carried out. Uncertainty is modelled via discrete distributions obtained via sampling and we deal with parametric uncertainty while assuming the model structures are certain. All the models are equation based models and simulation methods[1] are used for solving the selection problem that arises from implementing the framework. Our focus is also limited to the financial evaluation of projects and we don't include other objectives although it is our belief that the framework can be adapted for such use.

1.7   Thesis Outline

After the introduction in this chapter, we continue in Chapter 2 with a review of literature of process design in the Chemical Engineering literature. We examine some of the techniques and approaches used to deal with uncertainty and discuss their shortcomings. We summarize the areas where we see opportunities for contributions and developments and highlight which are going to be developed in this thesis.

In chapter 3, we take a deeper look at uncertainty and how it affects decision making. We given an overview of different frameworks that exist for modelling uncertainty but settle on probability theory for this thesis. We also examine ways to select among different distributions as well as how to use sensitivity analysis to determine key input uncertainties.

Chapter 4 discusses the elements of decision theory and presents utility theory as the key to understanding and modelling decision makers preferences with a view to incorporating them into the decision problem.

In chapter 5, we introduce the framework of Stochastic Dominance for decision making under uncertainty. We show how it relates to different classes of preferences within utility theory and explain the mechanism for its implementation and also discuss improvements on the framework - like Convex and Almost stochastic dominance - that improve the refinement of the efficient set. We also briefly introduce other ways to reduce the efficient - like the use of alternative metrics and the iterative approach to using Stochastic Dominance by identifying and reducing the uncertainties in the key input variables.

Chapters 6, 7 and 8 are the case-studies that we use to fully examine the methodology. In chapter 6, we look at the design of a reactor-separator system wher ethe reaction rates and product prices are uncertain and how these, coupled with preferences of the decision maker, can affect the choice of the capacity of the system. Chapter 7 examines a subset of the Biomass-to-Liquids problem introduce earlier in this chapter where the focus is on the selection of an adequate biomass crop as feed for the plant. Chapter 8 looks at design of the larger BTL system, investigating choices in raw materials, process technologies and possible product choices. In all these cases, we compare the results from Stochastic dominance with those from the other metrics and examine more closely the differences and similarities where they exists to further illuminate the method's advantages when compared to others.

---

[1] Simulation methods as opposed to mathematical programming techniques

In chapter 9, we summarizes the key findings and lessons of the thesis and highlight a numnber of different directions that one could embark upon with the work presented. An appendix on mathematical programming essentials as well as one that contains the essential algorithms used in the methodology are included for the reader.

Chapter 2

Project Selection: Background and Review

We begin this chapter with a review of current wisdom in project selection and design and how it is carried out in practice. The outcome of our analysis is to present identify some of the key aspects where improvements are needed, laying the ground work for our discussion of stochastic dominance later in the thesis.

2.1 The Design Process

In the chemical industry (and the larger engineering industry at large), projects that finally get implemented go through different phases in the development process. These phases can, in general be classified into the following stages:

1. Conception and definition

2. System-level design

3. Detail-design

4. Refinement and optimization

5. Implementation

While there are variations in the names and number of stages (see [116], [117], [103], [112], the descriptions of the overall process is similar. An illustration of the process as outlined by [117] is given in Figure 2-1

In the first stage, the scope of the project, the market, the kind of product and a brief outline of its characteristics, or the basic elements of the process are the key aspects of the design that may be defined. In the second stage, a high-level, input output model is constructed that gives a fair description of the process/product output in terms of the input. For example, a process model will relate the key inputs (like raw materials, energy and costs) into the product output (production rate, revenues) to a good degree of accuracy.
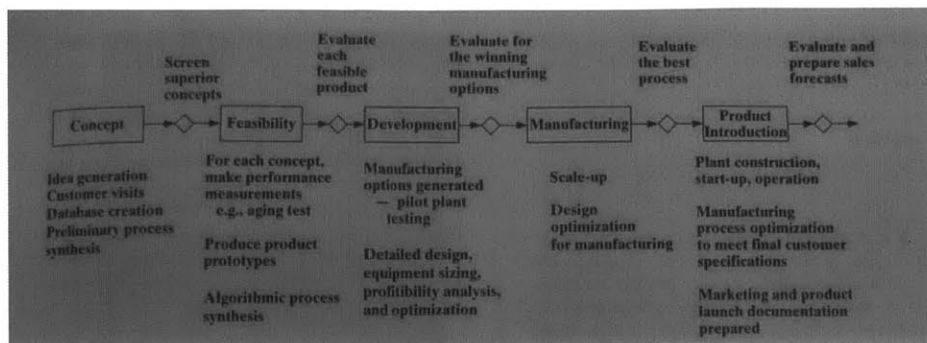
Figure 2-1: Illustration of the product development process. Taken from [117]

The third stage involves more detailed design of some of the elements of the system. So for a chemical process, this is where more accurate models for the separation, heat-transfer and other elements of the process. In the fourth stage, more attention is paid to the interactions between the now more-detailed subsystems in the process while in the final stage, the basic process is optimized around the parameters selected in the previous stage and the finalized design is then pushed towards implementation which usually involves engineering, procurement and construction.

2.2   Stage-gates and the Design Process

Companies begin the design process with a large number of products or processes under consideration in the development pipeline. For example, for every drug that comes to the market, pharmaceutical companies reportedly begin the development process with about 10,000 molecules [93]. It is important the that the design process is able to effectively manage the complexity of such a process.

In addition, as products/projects advance through each stage, there is an increasing cost associated with their evaluation, with a corresponding reduction in the overall uncertainty in the numbers. This is illustrated in Figure 2-2. To prevent development costs from ballooning, we have to be able to prune the feasible set of project options early on.

The stage-gate process is the approach used to manage development cost and complexity in product and process design. All feasible projects start out at the beginning of the process. They are analyzed using a procedure and metrics identified at the outset and at the end of the analysis, they encounter a gate where decisions about which projects advance to the next stage of analysis or are pruned out of the list are made. This sequence is repeated/iterated over until a final decision about the project(s) to be implemented are reached and the process is terminated.

At the end of each stage one of three decisions is made regarding each of the projects that is still in the pipeline up to that point

1. Proceed

2. Terminate

34
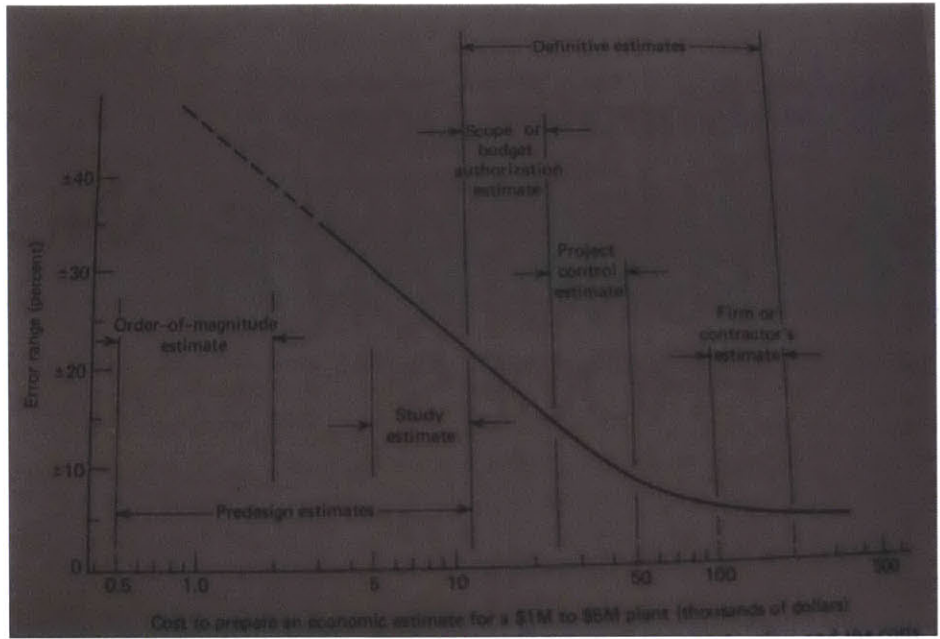
Figure 2-2: Illustration of the increase in project cost with design stage and the reduction in cost associated with it. Taken from [116]
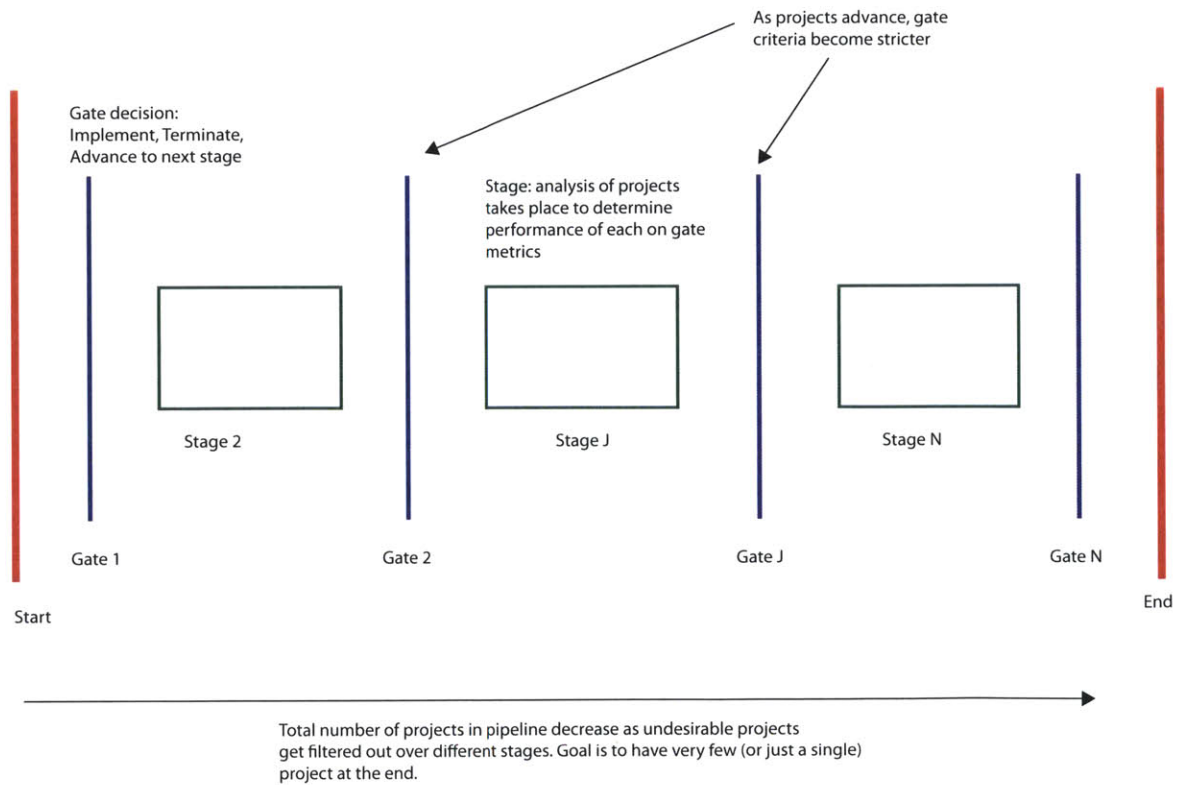
Figure 2-3: An illustration of the stage gate process

3. Revise

These decisions are based on the outcome of an evaluation on each project. These evaluations are sometimes based on clear and measurable metrics on which each project is scored. However, these scores are often subjectively combined to produce an overall recommendation that generally lies in one of the threee classes above. Projects that are set to proceed often have satisfied all/most of the significant criteria and are due for evaluation at the next stage. Projects that are terminated fail on all/most of the critical dimensions and resources of the company will be better utilized if diverted elsewhere. An example of a project that fails will be one that does not turn a profit in the long-term, or another that violates safety or environmental standards. Projects that are often revised are those that meet some but not all criteria and are potentially viable if few modifications are made to the present form.

Within each stage though, the decision is often similar: given the set of options presented, which ones should be selected for further analysis, revised or terminated. The techniques we develop in this thesis are largely applicable to the projects we want to select to advance in the stage-process and those we want to reject and we don't do much by way of suggestions for revisions. Furthermore

2.3 Metrics for project selection

When projects are selected to advance from a given stage or when they are terminated, they are often evaluated by scoring them on a number of metrics to objectively compare them with other possible options. Determining these metrics can be challenging and care needs to be taken to select the right ones.

This thesis is concerned with financial performance as one measure in which projects can be evaluated and even this has a variety of measures associated with it. Some of the more popular of these measures include the following: total annual cost, net present value (NPV), internal rate of return (IRR), payback period, return on investment (ROI). profitability index (benefit-cost ratio) and the hurdle rate. These different criteria are often grouped into three general criteria - quantitative criteria, which includes the total annualized cost, total profit, etc, qualitative criteria that include measures like the payback period and the internal rate of return and compromise criteria which includes the net present value [86]. We discuss a number of these metrics below:

The net present value of a project is a total value of the project over its lifetime, with the time value of money accounted for. Its value is given by

$$NPV = \sum_{i=1}^{N} \frac{CF_i}{(1 + r_D)^i} - I_C \qquad (2.1)$$

where CF is the cashflow (revenues net expenses) in the particular year, i, r_D is the discount rate (assumed constant here otherwise indexed by the year), I_C is the total capital investment in the plant and N is the lifetime of the investment. The net

present value is the terminal value of the present value of the project, which is just the cumulative value of a project up to a particular time point i.e.

$$PV_k = \sum_{i=1}^{k} \frac{CF_i}{(1+r_D)^i} - I_C \qquad (2.2)$$

The internal rate of return (IRR) is the value of the discount rate that makes the net present value of the project equal to zero i.e.

$$I_C = \sum_{i=1}^{N} \frac{CF_i}{(1+IRR)^i} \qquad (2.3)$$

where the equation often has to be solved iteratively to obtain the value of IRR.

The payback period is the length of time it will take for the capital investment to be recovered from the project returns. The equation to determine the payback period is similar to the IRR but this time, we are solving for the value of N where the discount rate is now fixed at a particular value. In the plot of the present value of a project over its lifetime, the payback period is the time it takes for the present value to reach zero.

The return on investment is simply a ratio of the profit from the project (before taxes) to the total capital invested in it. Over the lifetime of the project, this can be defined as

$$ROI = \frac{\sum_{i=1}^{N} \frac{CF_i}{(1+IRR)^i}}{I_C} \qquad (2.4)$$

The payback period is the amount of time it takes for the project to generate enough cashflow to repay its investment costs. The discounted payback period is similar but the cashflows are discounted to account for the time-value of money. Hurdle rates are minimum internal rates of return that a project must generate to be considered investment worthy. Projects that have IRRs less than the hurdle rate are rejected while those greater are accepted. The profitability index is a ratio of the present value of cashflows to the initial capital investment required for the project.

Of the metrics listed, internal rate of return and the net present value are the most common metrics selected by Chief Financial Officers [39]. Novak and Kravanja [86] carried out a survey of the objective functions used in the synthesis and design of process flowsheets over a given year and found that only 10% of the papers used the net present worth (value) as a criterion, with the most dominant being minimization of cost based metrics which accounted for more than half the 64 papers surveyed.

The reason for the importance in the choice of metric is that it can eventually affect the choice of project selected. [86] and [53] demonstrate this with different examples. In [86], the design of a heat-exchanger network using different criteria. They found that quantitative measures favour large projects with smaller (annual) costs and smaller cashflows, qualitative criteria favour projects with smaller cashflowsh with larger profitability while the compromise criteria strike an adequate balance between both.

The net present value is the best criteria for selecting between mutually exclusive projects [18]. Bagejewicz [9] discusses this metric in greater detail for the capacity planning problem and in particular compares it to other metrics like the ROI and the IRR and like [86] and [53], examines conditions under which the metrics lead to different optima.

Thus, even before considering the effect of uncertainty in a project, it is important to realize that the choice of metric can have an effect on the problem. The pros and cons of the different metrics should be weighted appropriately in order to find the one that corresponds most closely with the context under which the projects are evaluated before one is selected. In this thesis, we have selected the Net Present Value as the metric with which the project will be analyzed. As mentioned, this corresponds with both the wisdom in economic theory [18] as well as the view of practitioners in the field [39]. While the survey highlighted earlier in Chemical Engineering suggested that this measure still has some way to go, it is encouraging to see its advocacy in the more recent undergraduate texts in design [112] and [103].

2.4   Uncertainty and Risk

In the absence of uncertainty, project selection reduces to a problem that consists of

1. Selecting the appropriate metrics to use to evaluate each project

2. Carrying out the evaluation itself for each metric

3. Combining the different scores in a way that can give a ranking among the different projects

In the case of a single metric - say a financial objective for example as discussed in the previous section - then that ranking is implicit in the scores assigned to each project and the one that gives the highest score on the metric (i.e. that maximizes the metric) is easily selected. Thus, for example, if we choose to rank projects based on a financial metric, our first task will be to select an appropriate metric. If we choose it to the be the Net Present Value (NPV), our next task will be carrying out the evaluation on all the options. This is generally done using cost and revenue models built for each option. Since it is a single objective, the final values of the calculated NPVs gives a ranking of the projects and the one with the highest number will be picked. If a multiple ranking measure was used - say we also considered environmental safety - we would next either have to devise a means of combining these different metrics into one or finding an alternative approach for using them to somehow make a decision. We will discuss this idea of combination in more detail in Chapter 5 of this thesis.

In the presence of uncertainty however, even for single metric choices like the NPV, the matter becomes a bit more complicated because each project doesn't give a unique outcome that can be used to rank them. Instead, each project now yields a distribution of possible outcomes/scores on the given metric - each of which can
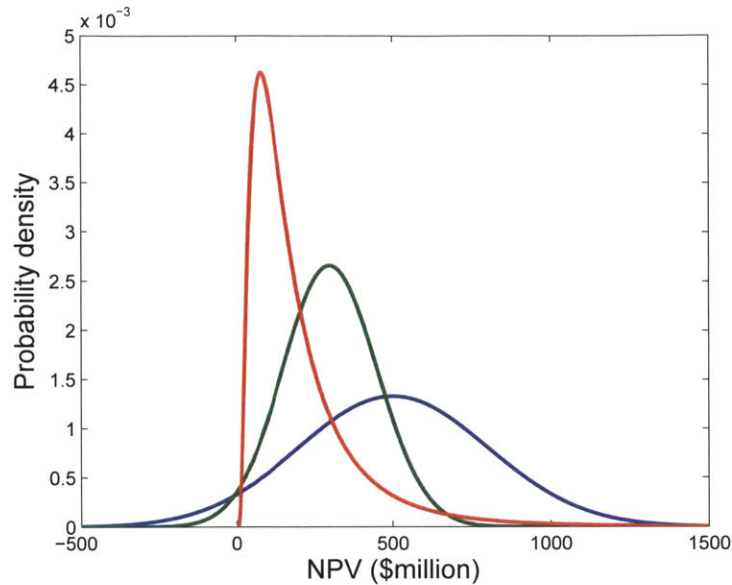
Figure 2-4: An illustrative example of the outcome of uncertain projects

be associated with a particular probability (or in the continuous case, a probability density). An example of this is given in figure 2-4.

In this example, we assume the decision maker has to choose between three project with different return profiles as shown in the diagram above. In this example, the metric of choice has been fixed to the net present value (NPV). In the presence of uncertainty, the projects, instead of giving a single output, now produce a distribution of outcomes that is characterized by a probability density (in the continuous case illustrated) or a probability mass function (in the discrete case). Given that each project now gives a range/distribution of outcomes, the challenge in decision making under uncertainty is determining how to select the "best" project from among the potential choices. In the next section, we review some of the current methods in use in resolving this problem.

2.5   Literature Review

We briefly examine current literature under these different groupings and highlight show how they are used to make decisions under uncertainty.

2.5.1   Single Objectives

By far the most common approach to project selection in the presence of uncertainty is ranking based on the **expected value**. The expected value of a uncertain quantity is the probability weighted average of the outcomes. Mathematically, this is described as:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx \qquad (2.5)$$

39

for the continuous case and

$$E[X] = \sum_{i=1}^{N} p_i x_i \tag{2.6}$$

for the discrete case. Thus to use the expected value to rank and select among the distributions given, the value of $E[X]$ is calculated for each of the projects and the one that gives the highest value is selected.

The **probabilistic objective** is another single objective approach. Here, the metric used is

$$C(X) = P(X \leq A) = \int_{-\infty}^{A} F(x)dx \tag{2.7}$$

where $F(x)$ is the cumulative distribution and (come up with the discrete representation).

$$C(X) = \frac{1}{N} \sum_{i=1}^{N} 1_{X \leq A} \tag{2.8}$$

where the function $1_{X \leq A}$ represents the indicator function that takes on the value of 1 when the expression $X \leq A$ and 0 otherwise. This objective minimizes the probability that the outcome of an event falls below some fixed number A. This formulation of the objective was first introduced by Roy [96] and for constrained optimization problems by Charnes and Cooper [24]. The value, A functions as a parameter in this problem and will be investor dependent.

Yet another objective is the **Worst-case/Robust Formulation**. Here each uncertain variable is characterized by its worst outcome and this outcome is used to rank the outcomes. Mathematically, this is calculated as

$$W(X) = \min_{D} X \tag{2.9}$$

where D is the support of the outcome X. This approach is used when X has a finite support i.e. the set D is bounded. The robust formulation is one of the more common engineering objectives and is used primarily in control applications as well as design and a detailed reference for the approach is given in [14].

Finally, we have another metric introduced by Aseeri and Bagajewicz [8] called the **Risk Area Ratio (RAR)** which is a ratio of the *opportunity area* to the *risk area*. Figure 5-9 pictorially illustrates these concepts. The Opportunity Area (OA) is defined as the area over which the cumulative distribution of the prospect, F_X, is above that of the project with the maximum NPV, F_Z. The risk area is the area over which the converse holds i.e. where $F_Z \geq F_X$. As we see, and unlike the other objectives so far introduced, the RAR is benchmarked by the decision variable with the highest expected value and is not a stand-alone objective. We calculate the RAR

40
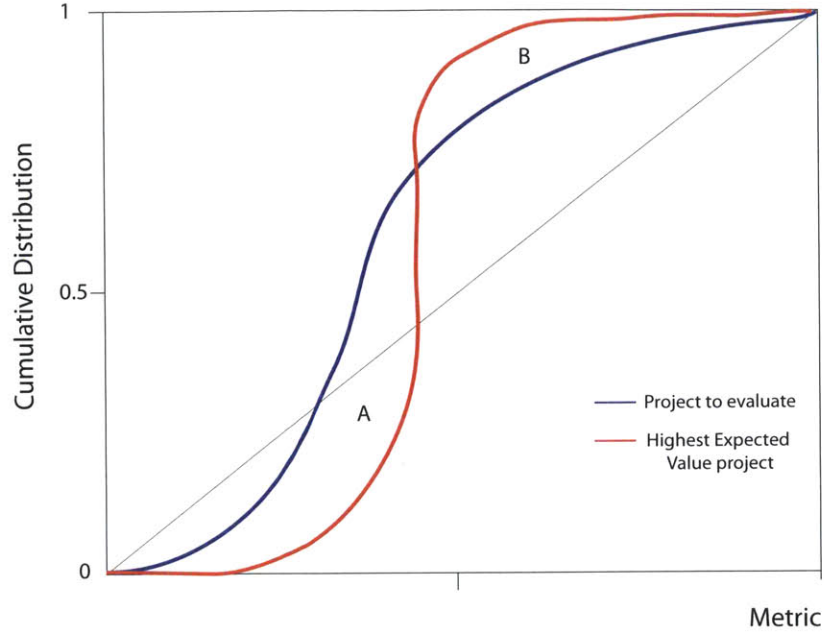
Figure 2-5: Illustration of Risk Area (region labelled A) and Opportunity Area (region labelled B)

as follows:

$$RAR = \frac{Opportunity\,area}{Risk\,area} = \frac{\int_{-\infty}^{\infty} \psi^+}{\int_{-\infty}^{\infty} \psi^-} \tag{2.10}$$

$$\text{where } \psi^+ = \max\{F_X - F_Z, 0\} \tag{2.11}$$

$$\text{and } \psi^- = \max\{F_Z - F_X, 0\} \tag{2.12}$$

2.5.2 Risk and Multiple objectives

Markowitz [68] rejected the claim that investors cared (or should care) only about the expected return of a distribution (the principal approach used at the time) and claimed that they should care about both its expected return as well as the variance (or risk) and by doing so, they could make better investment decisions. Thus, he ushered in the reward-risk framework for portfolio/project selection in the presence of uncertainty, and his proposal was a simpler alternative to the expected utility paradigm that was introduced by Bernoulli and axiomatized by Von Neumann and Morgenstern [121].

The expected value of a distribution has long been the standard measure of reward. Several risk measures however have however been used as a measure of risk to make decisions. We list a few of them below:

41

Variance

The variance of a distribution is the mean square deviation of the distribution from its mean value. Mathematically, it is defined as

$$Var(X) = E[(X - \hat{X})^2] \tag{2.13}$$

$$= \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x) dx \tag{2.14}$$

$$= \sum_{i=1}^{N} p_i (x_i - \hat{x})^2 \tag{2.15}$$

where $\hat{x}$ is the mean value of distribution and we have written out the form for both the continuous and the discrete cases.

Semi-Variance/Semi-deviation

Markowitz [69] introduced the semi-variance as an alternative measure for risk because it didn't penalize upsides like the variance. Semi-variance is a type of the more general semi-deviation measures where risk is measured primarily as falling below some aspiration level. So for a general level, d, the semi-deviation function associated with a random variable, X, is given as

$$r(X, c) = \begin{cases} (x - d) & \text{if } x \leq d \\ 0 & \text{if } x \geq d \end{cases} \tag{2.16}$$

And with this, the mean square semi-deviation is defined as

$$SD = \begin{cases} \int_{-\infty}^{\infty} r^2 f(x) dx & \text{if } X \text{ is continuous} \\ \\ \frac{1}{N} \sum_{i=1}^{N} p_i r_i^2 & \text{if } X \text{ is discrete} \end{cases} \tag{2.17}$$

When d is the expected value of the distribution, then we have it known as the semi-variance.

Value-at-Risk (VaR)

Value-at-Risk (VaR) was proposed by Jorion [51] as a direct measure of risk. Value at Risk is defined as the lowest amount, A, such that with probability beta, the loss over a given period of time, does not exceed A. There are as such three elements to uniquely specifying VaR for a distribution

1. A specified probability level, α

2. A loss amount, A

3. A time period

Since we are currently dealing with single-stage decision making, the last element doesn't matter (single period) and as such the VaR is mathematically defined as

$$\alpha VaR = \inf_{A \in R} \{P(X \geq A)\} \leq \alpha \tag{2.18}$$

where X is the loss amount. Because, in this study, we generally think of X as a profit like variable i.e. a variable where we want positive values and dislike negative values, then we redefine the VaR to reflect this (Krokhmal et at. (2010)) as

$$\alpha VaR = -\inf_{A \in R} \{P(X \leq A)\} \geq \alpha = -F_X^{-1}(\alpha) \tag{2.19}$$

where $F_X^{-1}(\alpha)$ is the inverse of the cumulative distribution of X. Note that in relation to the probability level,α, the relation '>' was used instead of '$\geq$'. Defined this way, we have the definition of VaR expressed in terms of the cumulative distribution as the α-quantile. Another expression of VaR is sometimes used, one defined as the 'lower' α-quantile (Krokhmal et al (2010)) and it is given as

$$\alpha VaR^- = -\inf_{A \in R} \{P(X \leq A)\} \geq \alpha = -F_{-1}(X, \alpha) \tag{2.20}$$

where the expression $F_{-1}(X, \alpha)$ is defined as

$$F_{-1}(X, \alpha) = \inf\{\eta | P(X \leq \eta)\} \geq \alpha \tag{2.21}$$

Conditional Value at Risk (CVaR)

The Conditional Value-at-Risk can be defined as the expected value of the random variable conditioned on it being larger than the value at risk i.e.

$$\alpha CVaR(X) = \alpha CVaR^-(X) = -E[X|X \leq -\alpha VaR(X)] \tag{2.22}$$

Slightly different to the Value-at-Risk metric, we have another definition of the conditional value-at-risk and it is the 'upper' CVaR, $\alpha CVaR^+(X)$ given by the following equation

$$\alpha CVaR^+(X) = -E[X|X < -\alpha VaR(X)] \tag{2.23}$$

where the difference is in the conditional relation of X with $-\alpha VaR(X)$ whether '$<$' or '$\leq$'.

Krokhmal et al (2010) reference also other definitions of CVaR - one given by Rockefellar and Uryasev [92] as

$$CVaR_\alpha(X) = \lambda_\alpha(X)VaR_\alpha(X) + (1 - \lambda_\alpha(X))E[X|X < -\alpha VaR(X)] \tag{2.24}$$

with $\lambda_\alpha(X)$ defined as

$$\lambda_\alpha(X) = (1 - \alpha)^{-1} F_X(-VaR_\alpha(X)) \tag{2.25}$$

43

and another by Acerbi (2002) as

$$CVaR_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha VaR_\lambda(X)d\lambda \qquad (2.26)$$

where we have redefined $\alpha VaR(X)$ as $Var_\alpha(X)$ to make the equations above more readable.

2.5.3 Multiple objectives and Dominance

With single objectives, the scores assigned to the different projects also serve as a basis for ranking and selection. For example, with an expected value metric, the project with the highest expected value (when scoring is done such that higher outcomes correspond to increased preference) is selected. With the Risk-Area ratio on the other hand, lower values correspond to more preferred outcomes so the project that gives the smallest value will end up being picked.

With the reward-risk framework, it is no longer as simple as we get a vector of scores which cannot be usually be ranked as easily.

Example 2 As an illustration, imagine we have a set of projects that yield the following scores for two different metrics: [3,4], [0,5], [1, 1], [4, 2] and [2, 2]. We tabulate the scores on both metrics in Table 2.1. We assume the two metrics to be 'postive' metrics - meaning that higher scores on each metric are desired. With a risk metric (where lower scores would be favoured), this can easily be rectified by taking the negative.

Project	Metric 1	Metric 2	Rank, Metric 1	Rank Metric 2
A	3	4	Second	Second
B	0	5	Fifth	First
C	1	1	Fourth	Fifth
D	4	2	First	Third
E	2	2	Third	Third

Table 2.1: Scores on different metrics for different projects

As the table shows, the rankings of the projects depends on the metric of choice. Project B, for example, ranks first on Metric 2 but fifth on metric 1 and project D ranks first on metric 1 and third on Metric 2. Thus if we compare project B and D, we can't say which is unilaterally superior as, depending on the strength of preference for metric 1 or metric 2, a decision maker may choose one or the other. B and D will then be said to be in the same *efficiency* set - the set for which no option is unilaterally superior to another. However, if you compare project A and C for example, for both metrics, project A is superior to project C as it scores higher on either metric. In this case we can say that A is unilaterally superior to C as every decision maker when presented with A or C, so long as they desire more to less of both metrics, will pick A. In this case, project A is said to *dominate* project C.

Efficiency and domination are the rules used to classify projects when multiple objectives are involved and nothing else is known about the relative preference between metrics for decision makers. They are also called *Pareto* efficiency and *Pareto* dominance, after Vilfredo Pareto, an Italian economist who used the idea in his studies of economic efficiency and income distribution (cite Pareto Efficiency Wikipedia Article).

Keeney and Raiffa (1976) more formally define dominance as follows: for a pair of projects F and G with components f_i and g_i for i = 1 to N metrics, F dominates G whenever

$$f_i \geq g_i, \quad \forall i \tag{2.27}$$

$$f_i > g_i, \quad \text{for some } i \tag{2.28}$$

2.5.4 Illustrative Example

We demonstrate the concepts that we have been discussing so far with an example problem. Assuming the firm is faced with 6 investment opportunities illustrated in Figure 2-6. To simplify analysis, we assume that each project costs the same to implement so there are no extra constraints to consider. To aid comparison, we also plot the distribution of all the projects together in Figure ??. The projects are identified by numbers 1 - 6 with project 1 being topmost left and 6 bottom right. We count across and down. The distributions corresponding to the projects are given in Table 2.2[1]

Project	Distribution Type	Parameters
1	Normal	$\mu = 600, \sigma = 200$
2	Normal	$\mu = 300, \sigma = 150$
3	Mixed Normal	$\mu_1 = 500, \sigma_1 = 40 \ \mu_2 = 200, \sigma_2 = 120$
4	Mixed Normal	$\mu_1 = 850, \sigma_1 = 40 \ \mu_2 = 150, \sigma_2 = 200$
5	Uniform	a = -200, b = 1000
6	Lognormal	$\mu = 600, \sigma = 200$

Table 2.2: Information on project distributions in illustrative example

For the single objective functions discussed earlier, Table 2.3 shows the ranking given to the set of projects by the each objective. For each case, we assume that the project with the highest selected value will be chosen.

For the multi-objective functions discussed, we split the projects into efficient and dominated sets and the table below indicates the assignment for each of the objectives discussed.

The value-at-risk and the conditional-value-at-risk were both set at values of α of 0.1. As we see from the table, the decisions for the three objectives are all different

---

[1]Mixed normals represent a probabilistic mixing of two normal distributions in order to yield a multi-modal distribution. To generate samples, a sample is drawn from one distribution according to a given probability and from the other in the complementary case

45
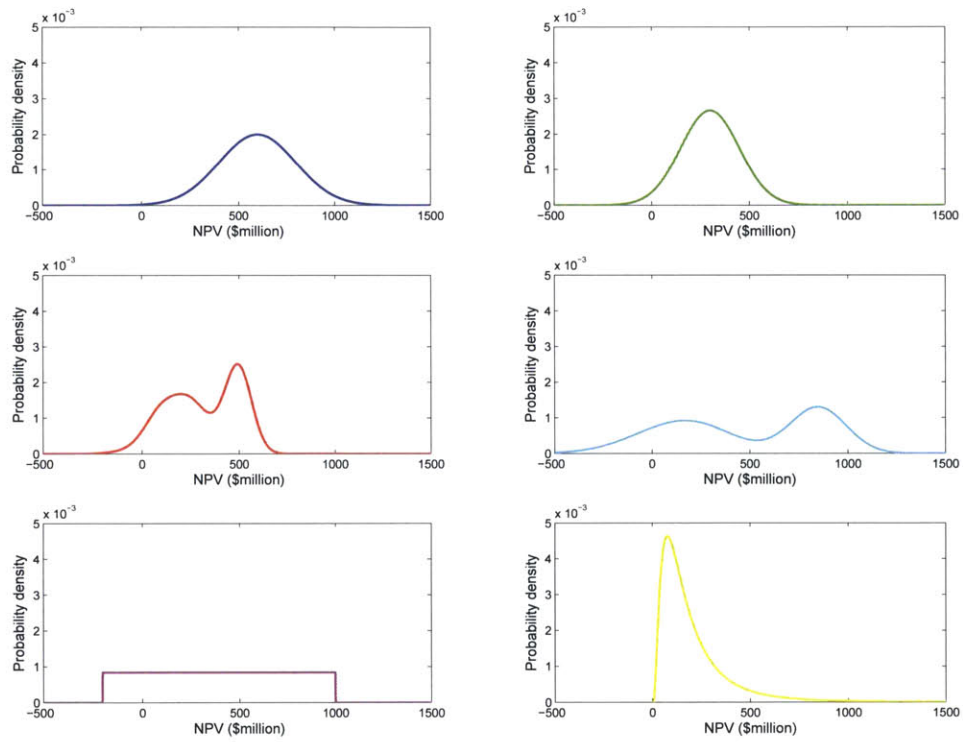
Figure 2-6: Projects Options for Example. Project 1 is topmost left, 2 is topmost right, 3 is middle left, 4 is middle right, 5 is bottom left and 6 is bottom right

Objective Function	Selected Project
Expected Value	1
Worst-Case	6
RAR	4

Table 2.3: Project selections with the different single objectives

Performance Measure	Risk Measure	Efficient Set
Mean	Variance	1, 2, 3
Mean	SemiVariance	1, 2, 3, 6
Mean	Value-at-Risk	1
Mean	Conditional VaR	1

Table 2.4: Project Rankings with the different Single objectives

and the efficient sets of the multiple decisions are also different. As a result there are decisions that will be either be omitted from the choice set of the decision maker or will contain decisions that wouldn't be picked by him, resulting in an inefficient screening process.

2.6 Issues with current approaches

We briefly review below the main drawbacks of many of the methods we just outlined.

Expected value maximization does not lead to diversified portfolios [96] - unless of course, there are non-linear interactions in project choices. For linear and independent investments, only the project with the maximum expected value will be chosen. An intuitive reason for this can be seen from Gilboa's description [37] of expected value as a certainty generator. In the infinite limit, the expected value is certain (has zero variance) and as such the expected value problem is, in effect, transformed to a deterministic one - and diversification is not optimal in such problems [2]. The only time an appearance of diversification can be obtained is if there are budget constraints on the each investment option, independent of the total available budget. In that case, the investments will take place in a rank order from highest to lowest investment until the budget is exhausted. However, arbitrary caps on projects lead to suboptimal solutions from the optimality principle [11].

The **Worst-case** approach suffers primarily from being too conservative as it implies that no matter how good an alternative looks, it can be discarded on the basis of one really bad outcome - regardless of how improbable that outcome is. It is rarely ever used for financial evaluations though it can be used to identify potential arbitrage opportunities

---

[2]In the actual deterministic case, the future is known for certain and so the optimal decision will be to invest all resources in the project that gives the highest returns per dollar

47

The Risk Area Ratio fails the Expected Utility analysis demonstrated in a later chapter. Basically, under certain conditions, it can be shown to choose projects that are universally determined to be inferior to others in the same pool. Our new framework - stochastic dominance - is able to identify such projects and is also able to vet the performance of other metrics too.

The **Mean-Variance** model selects alternatives based on, as the name implies, the mean and variance of the decision outcomes. The goal will be to maximize the mean and minimize the variance of the alternative. These are two independent objectives and as a result, rather than a single solution value, a set of values that make up a frontier can be optimal depending on the trade-off between the objectives that the decision maker is willing to tolerate.

Weisman and Holzman [123] show that mean-variance is equivalent to EU theory when the utility function is either quadratic or the uncertainties follow the normal distribution (it might be interesting to illustrate the results here with equations). We reproduce their derivation below Typically we begin by assuming that the decision maker is risk-averse everywhere i.e. the second derivative of the utility function is always negative (Keeney, 1976). This type of utility function can be parameterized by the function.

Semi-variance suffers from the setback that it is more difficult to handle than variance [12] although algorithms exist for determining the efficient frontier under this metric ([12] [45]). Under symmetric conditions however, Markowitz [69] shows that it yields the same results as the traditional mean-variance approach.

Value-at-Risk is a risk measure used in the banking industry today widely popular because of its simple yet intuitive meaning. However, Artzner et al. [7] demonstrate that it lacks coherence as a risk measure - in particular, under certain conditions, it fails to encourage diversification as an approach to reducing risk. Also, it doesn't give an idea of how much stands to be lost. So a 2week, 5% VAR of $1million says that over a two-week period, the company has a probability of 5% of losing a million dollars or more. Depending on the tail of the distribution, this could be as large as a billion dollars for example - enough to possibly wipe the firm out. As a result, it should be used with caution.

Conditional Value-at-Risk The conditional value at risk is a consistent risk measure [97], [7] and has only relatively recently begun to enter the main stream of risk [91] even though it has been around for a while (see [33]). It improves over the value-at-risk framework in two ways - it doesn't just give a lower bound for risk as it also gives an expectation of the loss that can potentially be incurred. It also satisfies portfolio diversification requirements.

2.7 Applications of Selection Objectives in Engineering

Mulvey et al.[77] introduced Robust Stochastic Programming in an attempt to account for the riskiness of alternatives. They use variance as a surrogate measure of

risk and attempt a multi-objective approach of maximizing mean and minimizing the variance of the return. While this improves over the pure expected value, it penalizes both upside and downside deviations symmetrically whereas investors only care predominantly about downside variance.

Cheng et al. [25] propose multi-objective, two-stage approach to the design problem in a dynamic programming framework for the analysis of the design of a reactor with the possibility of the appearance of superior technology in the horizon. They select three metrics - expected profit, downside risk and process lifetime - as a basis for evaluating their decision and develop a multi-objective dynamic programming framework to solve the problem. They do not however expand in much detail the basis of selection of their choice of risk metric. The compare their solution with the pure expected value formulation but do not examine different formulations of risk to see how their decisions change with them.

Aseeri and Bagajewicz [8] propose a new metric - the risk area ratio in an aim to move from the point-like measures that characterize metrics like mean, variance etc, to one which aims to capture information on the distribution as a whole. This ratio is however flawed from a preference perspective and it will be illustrated in detail in a later chapter. They however do a good job of reviewing other metrics (VaR, CVaR, mean-variance, etc) and the fact that their approach begins to consider ways to utilize the entire distribution as a basis for decision making is a step in the right direction.

In the Chemical Engineering process design literature, Turton, Bailie et al. [112] were the first proponents of the stochastic dominance approach to selecting projects in the presence of uncertainty. Stochastic dominance is a distribution free approach and relies on very broad and general assumptions about the utility functions of the decision maker. However, they don't discuss the use of stochastic dominance for decisions with continuous components as one can only make comparisons across discrete choices of projects.

Most of the articles reviewed above have focused primarily on the design of a single project and on ways to manage risk within such a project. Very few have focused on design/project selection as one that needs to be done at a portfolio level. This is important because diversification of investments has been long recognized as a way of reducing and managing risks [68]. Sometimes this is done in an arbitrary fashion using artificial budget caps and constraints - where predetermined levels of investments across different projects are arbitrarily decided. While this can indeed provide some level of diversification, it is not guaranteed to be optimal. By using a more quantitative/objective portfolio based analysis, an optimal allocation of resources that maximizes return and minimizes risk can be obtained.

Rogers et al. [93] are an exception however as they offer one of the more comprehensive approaches in the process engineering literature. In addition to considering an options-based approach to value the pharmaceutical investments, they also embed this in a portfolio framework where they recognize that real project decisions should be carried out in the context of assembling an investment portfolio from the perspective of the company. Again however, behavioural assumptions are not discussed and variance is assumed to be a sufficient measure of risk without any discussion of the

basis upon which the metrics lie.

2.8 Elements of the decision problem

In examining the design/selection problem so far, we have explicitly assumed that we have a number of things that make it possible to carry out the analysis. Some of these we need from the decision maker and others from the designer. We highlight and briefly discuss them below.

From the decision maker and stakeholders we need:

1. Metrics for evaluation such as whether to use net present value, internal rate of return, or a number of other different metrics. As discussed in the first section of this chapter, this choice also has the potential to significantly affect the project decision, even in the absence of uncertainty and should be examined in detail

2. Constraints such as environmental and safety regulations, budgetary restrictions, strategic investments direction of the firm towards a particular industry, constraints and budgetary constraints play a huge role in the determination of feasible designs and often time

3. A description/knowledge of their general risk preferences/attitudes

From the design engineer, we expect

1. Mathematical/simulation models for the project options which translate input decisions to output desirables

2. A framework for the representation and propagation of uncertainty through the generated models

3. The ability to model the preferences of the decision maker in a manner that generates a usable objective function or provides a rule for refining the feasible setto appropriately use for project selection

The ability to solve the decision problem relies on the intersection of both groups of expectations and sometimes, it is information that is iteratively communicated and used to solve each problem.

In subsequent chapters, we will examine some of the elements highlighted above in greater detail. The next chapter examines ways in which uncertainty is quantified (modelled and propagated) in process systems. We utilize the probabilistic framework because it is the most widely used approach for representing uncertainty. Uncertainty propagation characterizes the uncertainty in the output from a model given knowledge of the structure of the model and the form of the input uncertainties. Once we are able to quantify uncertainty in the different projects, we are able to run some of the analysis that we reviewed earlier in this chapter.

To analyze and incorporate preference, we use utility theory as the framework for representation, culminating with the development of expected utility as the prescriptive approach for to decision making under uncertainty within utility theory.

Chapter 3

Uncertainty Analysis

The presence of uncertainty in project outcomes plays a huge role in shaping how projects are ultimately selected. However it is often impossible to tell, a priori, the uncertainty in project outcomes as all we often possess is the knowledge of the input uncertainties as well as a model of the project that links the inputs to the outputs. In order to determine output uncertainties, we need to be able to:

1. Adequately model the uncertainties in a form that makes them amenable to analysis

2. Efficiently propagate the input uncertainties through the models in order to determine the output uncertainty

3. Quantify the effect of each uncertain input on the output uncertainty via a sensitivity study

While the last point is not necessary for uncertainty propagation, it is very useful in outlining steps to embark upon after analyzing the decision problem and help improve the next round of decision-making. In this chapter, we discuss the three objectives in greater detail. First, we highlight examples of the different frameworks
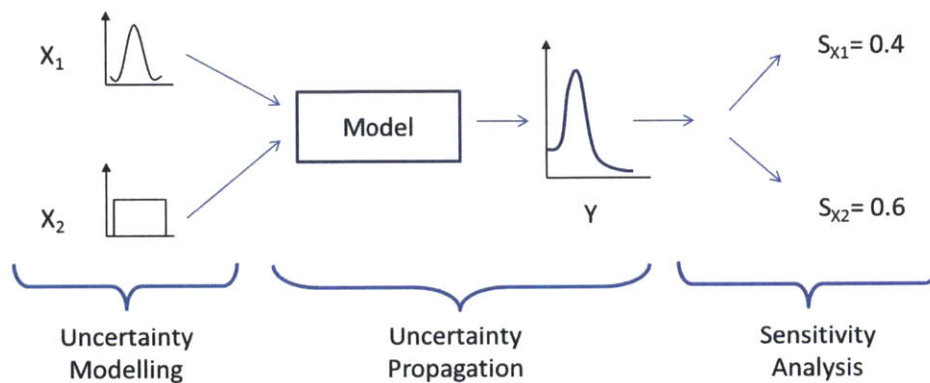
Figure 3-1: An illustration of uncertainty analysis

for modelling uncertianty, but focus primarily on the use of probability theory since it is the most widespread framework in use.

The second part of the chapter gives an overview of the different methods of propagating uncertainty through models - via monte-carlo methods and the more advanced polynomial chaos expansion approach (insert reference). Advantages and disadvantages of both approaches are highlighted and a simple guide for when to use one or the other provided.

Uncertainty analysis is usually restricted to mean the first two objectives but in this thesis we expand it to include the third objective, traditionally referred to as sensitivity analysis. Knowledge of sensitivites is important for action - we can use this in an experimental design to figure out the variable to study that gives us the best 'bang for our buck' in reducing the uncertainty in the output.

3.1   Introduction

Process design is carried out under an uncertain environment and the ability to effectively handle the presence of uncertainty in process design is the biggest challenge that the field faces today. The uncertainty arises from sources internal and external to the process. Examples of internal uncertainties are technological performance parameters (e.g. reaction rate constants), and external uncertainties include raw material yields, prices, market sizes for products, regulations etc.

The traditional way uncertainty was handled in process design included using nominal values for the uncertain variables and then using a (local) sensitivity analysis to check for optimality. While this represents an improvement over the deterministic case, it still fails to capture the full range of outcomes that decision-makers are exposed to when they need to make decisions. The goal of uncertainty modelling and propagation is to quantify the effects of the uncertainties in the different inputs of a process or a system on its output(s) and to determine the key drivers of the output uncertainty where possible.

Uncertainty is often classified into two major types - epistemic and aleatory uncertainty [82]. Epistemic uncertainty refers to the types of uncertainties where it is possible for us to ultimately know the true values either through increased observations or through the use of more accurate measuring tools. Physical constants such as the acceleration due to gravity, the speed of light in a vacuum, etc all fall under this category. Aleatory uncertainties on the other hand are those that arise from underlying random processes and whose uncertainties cannot be reduced by increased observations. These variables are truly random and the best that can be done is the knowledge of the particular random process that describes the behaviour of such an uncertain variable. Examples of such are the positions of subatomic particles, the behaviour of stock prices, etc.

There is sometimes some overlap between the two types of uncertainties. Take, for example, the modelling of the behaviour of stock prices. Assuming we believe that we can model it via a Weiner process, then, the problem typically becomes one of trying to determine the appropriate parameters that describe the distribution. With more

historical data we may believe we can more likely converge on the exact parameters that describe the distribution, even though the underlying variable remains truly random. Knowing which kind of uncertainty we are dealing with is quite important as it can help inform our actions for dealing with the uncertainty. Where epistemic, it may be worth experimenting to know the true value of the parameter, thus eliminating the *source* of the uncertainty, whereas, with aleatory uncertainties, we are, at best, only able to eliminate the *effects* of uncertainty via the use of contingency plans

Within mathematical models, uncertainty is often also divided into two areas - parametric uncertainty and structural uncertainty. Parametric uncertainty refers to the lack of knowledge of particular parameters that are used within a model. Take for example, the rate of the chemical reaction described below:

$$A \longrightarrow B \tag{3.1}$$

$$r_A = \frac{dC_A}{dt} = k_1 C_A \tag{3.2}$$

In the above, we have assumed the reaction proceeds via a first order mechanism. The parameter to be determined is the reaction rate constant, k_1. This parameter is unknown and often has to be determined from experiments which usually involve running the reaction over a period of time, determining the concentration of reactant, A, at intervals and then fitting the first order curve to the data to determine the rate constant. Since there will be errors in measurement (due to a host of reasons), there will be a corresponding uncertainty in the value of this determined parameter. This uncertainty is the parametric uncertainty. Structural uncertainty however refers to the uncertainty regarding the specific form the mathematical model takes. For example, we could also have hypothesized that the reaction proceeded via a second order mechanism, leading to the reaction

$$r_A = k_2 C_A^2 \tag{3.3}$$

In this case, when trying to determine the parameter k_2 from the experiments, the results will likely be quite different. Given that the parameters for these models are obtained with measurements that assume one model over the other, model uncertainty certainly influences the parametric uncertainty. Model uncertainty is particularly important in the area of model selection and discrimination. While in this thesis, our particular focus will be on parametric uncertainty, interested readers are referred to [15] for a more in-depth analysis of model uncertainty in process design.

To handle uncertainty, one is in need of two things - a mathematical framework that can be used to represent it and an efficient way to use this representation in computations required for making design decisions. We discuss these in the subsections below.

3.2 Uncertainty Modelling

In order to incorporate uncertainty in the design and selection process, we need a mathematical framework for modelling and representing it. Tatang [108] outlines four frameworks for representing uncertainty and they are listed below (you can try to find another reference if that works better)

- Interval arithmetic

- Fuzzy theory

- Demnpster Schafer theory

- Probabiity theory

While he discusses the pros and cons and field of applicability of the different approaches, he ultimately concludes that probability theory is the most widespread and most theoretically developed framework for modelling uncertainty and that is the framework we will be utilizing in this thesis. Interested readers in the other frameworks are referred to his thesis work [108] and related references.

In addition to a theoretical framework, we need to a model that describes the way the particular uncertainties that we are concerned with vary. In probability theory, this model is the probability distribution and it can be one of many forms e.g. uniform, normal, exponential etc. Knowledge of the underlying generating process for the random variable can suggest a particular form for the uncertainty.

Often times, we know some statistics of the underlying variable like its mean, variance, range, mode etc. Information theory can be used to suggest models that minimize the extra amount of information that the model adds to what we know. For example, given knowledge of the mean and the variance of a random variable, the Gaussian is the distribution that adds the least amount of extra information to what we already know. The principle of maximizing entropy is the mathematical backbone for the determining such distributions and it is elaborated on in [49] and [27].

Figure 3-2 from Gong [38] summarizes the maximum entropy distributions that arise from different scenarios of available information.

3.3 Uncertainty Propagation

Besides modelling uncertainty, one also needs to be able to propagate it through the process system. This is necessary in order to quantify the overall uncertainty in the outputs as a function of the inputs.

Monte-Carlo methods are currently the simplest and most widespread means for propagating uncertainty through models today. These methods were invented in the 1940s by Metropolis and Ulam during the Manhattan project for the development of the atomic bomb. It was named after the Monte Carlo region in France, which was a popular gambling destination The basic principles behind the method are:

1. The selection of random samples from the probability distributions of the inputs

Available Information	Maximum Uncertainty Distribution	Probability Density Function
None	Uniform	$f_X(x)=\begin{cases}\dfrac{1}{b-a} & x\in[a,b]\\[4pt]0 & otherwise\end{cases}$
[a, b] $E[X]=\mu$	Truncated exponential	$f_X(x)=\begin{cases}c\exp(-kx)^{\,\cdot} & x\in[a,b]\\ 0 & otherwise\end{cases}$
$\text{Mode}[X]=c$	Triangular	$f_X(x)=\begin{cases}\dfrac{2(x-1)}{(b-a)(c-a)} & x\in[a,c]\\[6pt]\dfrac{2(b-x)}{(b-a)(b-c)} & x\in[c,b]\end{cases}$
[0, +∞] $E[X]=\mu$	Exponential	$f_X(x)=\dfrac{1}{\mu}\exp\left(-\dfrac{x}{\mu}\right)$
$E[\ln X]=\mu$ $E[\ln(X^2)]=\sigma^2+\mu^2$	Log-normal	$f_X(x)=\dfrac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right)$
[-∞, +∞] $E[X]=\mu$ $E[X^2]=\sigma^2+\mu^2$	Normal	$f_X(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right)$

Figure 3-2: Distributions corresponding to the respective available information via the principle of maximum entropy. From [38].

2. The deterministic solution of the model to obtain the process outputs and

3. The eventual aggregation of the outputs into a histogram to get the output distribution

Selection of random samples from the probability distribution can be done with the use of appropriate pseudo-random number generators. Standard numerical methods for solving systems of equations are then used to generate model outputs from the inputs and these can be made into a histogram using appropriately sized bins to collect the outputs.

Monte-Carlo methods were initially developed to solve deterministic problems like determining the expectation of a function where the analytical calculations would have proved to be intractable. The advantages of the Monte-Carlo method are that it is very easy to use and as more and more samples are taken, convergence to the "true" output distribution is guaranteed [47]. Its biggest disadvantage however is that the rate of convergence is quite slow and typically many samples are required to typically improve the convergence of the output to desired levels. Regardless of the dimension of the problem, Monte-Carlo converges at a rate of $N^{-1/2}$ where N is the number of samples. Thus to reduce the error of a monte-carlo estimate by half, one needs to quadruple the number of samples and to improve the accuracy of an estimate by 1 decimal place, you need to take a hundred times more samples - hence its computational intensity.
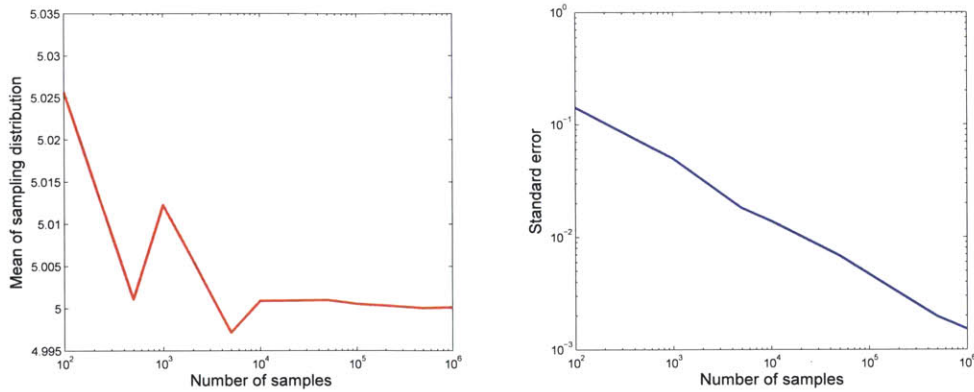
Figure 3-3: Illustration of monte-carlo convergence for the example. The figure on the left shows how the mean converges to the value of 5 while that on the right shows the convergence of the estimate

Example   We illustrate the convergence above with an example where we try to estimate the mean and standard deviation of a normal random variable whose mean and standard deviation we know to be 5 and 2 respectively using a range of sample sizes from 100 to 1 million. For each sample size, we estimate the mean of the samples taken. This is a sample mean for the particular sample size. We repeat this 100 times in order to get a 100 sample means for the particular sample size. We can now use this to determine the mean and standard deviation of the sampling distribution for each sample size. The standard deviation of each sampling distribution is the standard error. We then plot standard error against the number of samples and confirm the convergence relationship. We plot the means of the distributions also to display the convergence to the true value of 5 as we increase the number of samples.

In an attempt to improve the convergence of the Monte-Carlo method, better sampling techniques have been developed - techniques like the Latin Hypercube sampling, importance sampling, control variates etc. The overarching theme behind these techniques is to use our knowledge of the uncertain variable to spread out the samples so that fewer samples are used to obtain the same distribution. The trade-off imposed by these methods is the increasing complexity of both the implementation and the analysis of the resulting statistics. In a number of cases this trade-off is often worth it - although we will be using simple monte-carlo for uncertainty propagation in this work.

3.4   Sensitivity Analysis

Sometimes, all we need to know for deciding on the project to implement is the knowledge of the output uncertainty as a function of the inputs and the parameters. With this knowledge, many of the techniques reviewed in the previous chapter are able to select promising projects from among a group of candidates and eliminate those that are outside the preference of the decision maker.

However, in some cases, we need to know which of the uncertain inputs have the greatest effect on the output with the goal to possibly reduce the effect that such an input has. And in the case where this uncertain input affects all the projects, reducing its effect may help to better discriminate between good and bad projects. The process of decomposing output uncertainty into the respective contributions of the different inputs is the goal of sensitivity analysis. Sensitivity analysis is often categorized into one of two different groups

1. Local sensitivity analysis where the effects of small pertubations of the inputs around the nominal/design points are used to determine the contribution

2. Global sensitivity analysis where the effects of variaton of the input on the output over the whole range of the input space is used

We review the different techniques for both local and global senstivity analysis below with our primary references being the review text by Saltelli et al [99] and [100] where more detailed references on the different methods can also be obtained.

3.4.1 Local Sensitivity Analysis

Local sensitivity analysis easily traces its roots to the deterministic approaches to modelling, where the earliest means of determining the effect of slightly varying the input on the output could be determined from the derivative of the output function wrt the input i.e.

$$S_i = \frac{dy}{dx_i}\bigg|_{x=x_0} \tag{3.4}$$

where S_i is the local sensitivity, y the output and x_i the particular input variable and x_0 is the nominal value of the parameter (note that when derivatives are taken, unless otherwise stated, we will assume that they are taken at the nominal value, x_0). This technique works quite well when the variables are all within the same order of magnitude and have the same degree of dispersion. When the scale of variation is large enough, it is often good to normalize the sensitivity analysis, which gives

$$S_i = \frac{x_{i_0}}{y_0}\frac{dy}{dx_i} \tag{3.5}$$

so that in this case all the sensitivities are now dimensionless and roughly within the same range of order of magnitude. If the degree of dispersion of the different input vary considerably, it is also important to account for this in determining the sensitivity as given in [100] this way

$$S_i = \frac{\sigma_{x_i}}{\sigma_y}\frac{dy}{dx_i} \tag{3.6}$$

where σ corresponds to the standard deviations of the different subscripts. For very simple models, the derivatives can be obtained analytically while for more complex models (and for black-box type models where the model structure is not observable

to the user), numerical simulations are needed to compute the derivatives with different techiques available depending on the accuracy needed and the complexity of evaluations and such techniques are discussed in [23].

Local senstivity analysis is often done via a 'one-at-a-time' approach, where interaction effects are neglected and the effect of varying only the input of interest on the output is accounted for. Where the models are linear and interaction terms are negligible, the approaches we have considered are sufficient for determining sensitivities. Where this isn't the case, a global approach to sensitivity analsysi is required and this is discussed next.

3.4.2 Global Sensitivity Analysis

Local sensitivity analysis is plagued by a number of factors, notably the inability to include some measure of interaction of the variable of interest with other variables in the determination of sensitivities and the fact, not highlighted in great detail, that the sensitivity computed can be a function of the nominal values used to evaluate the derivative.

Global sensitivity measures (or indices) usually correct for these as they are able to account for interaction and are a measure of the interaction over the entire space of the variable of concern. Different techniques exist for the evaluation of global sensitivities. Examples include Fourier Amplitude Sensitivity Testing [101] [75], polynomial chaos expansions [106], regression based approaches approaches [100] and monte carlo methods [98] [101]. While we will be using the monte-carlo based methods largely because of ease of implementation and how it naturally fits into the uncertainty propagation techniques we are already other methods should be evaluated for situations where they might be more advantageous. We expand on the method below, the discussion largely adapted from [98] and [100].

Saltelli [98] introduced Monte-Carlo based approaches for evaluating Sobol's indices and the ability to minimize computational requirements of the test and to use analysis to get a sense of the total effects, even when single effects cannot be determined. The strength of the difference between first order and the total sensitivity analyses gives a measure of the degree of importance that can be attached to interaction effects.

The approach generates both first order sensitivities, which we represent like the others as S_i and total effect sensitivities, S_i^T which accounts for all the interactions. Both first order and total effect sensitivities are calculated by Monte-Carlo sampling. We assume that we have a model given by

$$Y = f(x_1, x_2, ..., x_N) = f(\mathbf{x}) \tag{3.7}$$

where Y represents the output of the model and x the vector of inputs. In this approach, the first order sensitivity S_i is defined as

$$S_i = \frac{V(E(y|x_j))}{V(y)} \tag{3.8}$$

58

where $V(E(y|x_j))$ is the variance of the expected value of the output (y) when the factor x_j is kept fixed. The total sensitivity, S_i^T is defined as the

$$S_i^T = \frac{V(Y) - V(E(y|x_{-j}))}{V(Y)} \tag{3.9}$$

where x_{-j} is shorthand notation for all other variables apart from x_j. S_i^T can also be described (see[106]) as the sum of the sensitivities of all the terms in $f(x_1, x_2, ..., x_N)$ that doesn't involve variable x_i, i.e.

$$S_i^T = 1 - S_{\tilde{i}} \tag{3.10}$$

where the index $\tilde{i}$ represents stands for all variables except i. In order to calculate each of these parameters, we will first have to generate samples of the input variable. We generate matrices M_1 and M_2 where we define the ith row of M_1 as

$$M_1 = x_1, x_2, ..., x_j, x_{j+1}, ..., x_N \tag{3.11}$$

which is a sample from the joint distribution of all the inputs (which is, for independent random variables, just a collection of samples from each of the individual distributions). In the same manner, the matrix M_2 contains a resampling of the variables in matrix M_1 i.e. the ith row of matrix M_1 contains samples

$$M_2 = x_1', x_2', ..., x_j', x_{j+1}', ..., x_N' \tag{3.12}$$

where the apostrophes are to indicate the resampled values of the variables sampled in M_1. The value of the first order and total sensitivity indices can be calculated as follows:

$$S_i = \frac{U_j - E^2(Y)}{V(Y)} \tag{3.13}$$

where the U_j is defined as

$$U_j = \int E^2(y|x_j = \tilde{x}_j) p_j(\tilde{x}_j) d\tilde{x}_j \tag{3.14}$$

where $p_j(x_j)$ is the probability measure (or probability density function) of the variable x_j and E^2 refers to the square of the expected value of the variable. Since most times we can't compute this analytically, a simulation based approach is used to calculate U_j which is

$$U_j = \frac{1}{Nsim - 1} \sum_{r=1}^{Nsim} f(x_{r1}, x_{r2}, ..., x_{rN}) f(x_{r1}', x_{r2}', ..., x_{r(j+1)}', , x_{rj}, x_{r(j+1)}', ..., x_N') \tag{3.15}$$

The above equation refers to resampling all the variables except x_j and then evaluating the function at those values. In relation to the earlier matrices M_1 and

M_2, we simply create a new matrix N_j where we replace the jth column of matrix M_2 with the column from M_1 and evaluate the function with those values. Once this has been done, we can then proceed to calculate the value of the first order sensitivities. For purely additive models, $\sum_{i=1}^{N} S_i = 1$ thus giving a clear picture of the contributions of the first order indices to the overall variance.

To compute the total sensitivity indices, we use the equation

$$S_i^T = 1 - \frac{U_{-j} - E^2(Y)}{V(Y)} \tag{3.16}$$

where the quantity U_{-j} is estimated via the following equation

$$U_{-j} = \frac{1}{N_{sim} - 1} \sum_{r=1}^{N_{sim}} f(x_{r1}, x_{r2}, ..., x_{rN} f(x_{r1}, x_{r2}, ..., x_{r(j-1)}, x'_{rj}, x_{r(j+1)}, ..., x_N)$$

$$\tag{3.17}$$

where, rather than resample everything, we resample only the value of x_j. Unlike the first order indices, the total sensitivity indices do not need to sum to 1 unless there are no interaction terms, in which case, the values will be identical to the first order indices. As mentioned earlier, the total sensitivity index has included in it, all the interaction terms for variable x_j. Thus for a model with three inputs, the total sensitivity index for input 1 will be the sum of the following

$$S_1^T = S_1 + S_{1,2} + S_{1,3} + S_{1,2,3} \tag{3.18}$$

where $S_{i,...}$ represents the contributions of the interaction terms of input i and the other inputs.

Since the sensitivity indices are determined via random sampling, they are themselves random variables and will have errors associated with their estimation similar to the convergence errors for Monte-Carlo estimation discussed earlier. Homma and Saltelli [46] discuss the quantification of these errors and interested readers should consult it for more detail.

We will implement the sensitivity studies in the case studies we will examine in the later chapters of the thesis as they will inform which inputs to more specifically focus on in the analysis of the different projects.

Chapter 4

Preferences, Utility and Expected Utility Theory

In many firms, the design process often doesn't include the decision maker at the base-level of analysis where preferences need to be included. For example, in selecting projects for investment, the structure of most decision processes look more like that in Figure 4-1.

In general, the decision makers set the strategic direction and other 'system' variables within which the eventual design engineers operate and they make the final project and investment decisions based upon the recommendations of the engineers. Supervisors set specific goals and benchmarks for the design team and report specific options for consideration to the decision makers. The design engineers get goals and targets from the team leads and make details designs which are passed back to the team leaders for recommendations to the ultimate decision makers.

It is the preferences of the ultimate decision-makers that we want to make sure are adequately reflected in the methodology used for screening among the numerous design options that the engineers and designers have to analyze. This is because, the decision maker rarely gets to see the universe of feasible options and if an incorrect screening approach has been used to select the final (set of) project(s) presented to them, it is possible that better projects could either have been screened out of consideration or drowned out by the presence of far too many decisions in the set.

This thesis proposes that better decisions can be made by accurately and transparently incorporating the preferences of decision-makers into the project selection process. But we have yet to define what these preferences are and how we can represent and incorporate them into the decision-making process. Also, in order to be objective in our comparison of alternative approaches to project selection, we need a basis of evaluation that is consistent in our representation of these preferences. We turn to utility and expected utility theory as they provide a sufficiently general framework for addressing these questions.

61
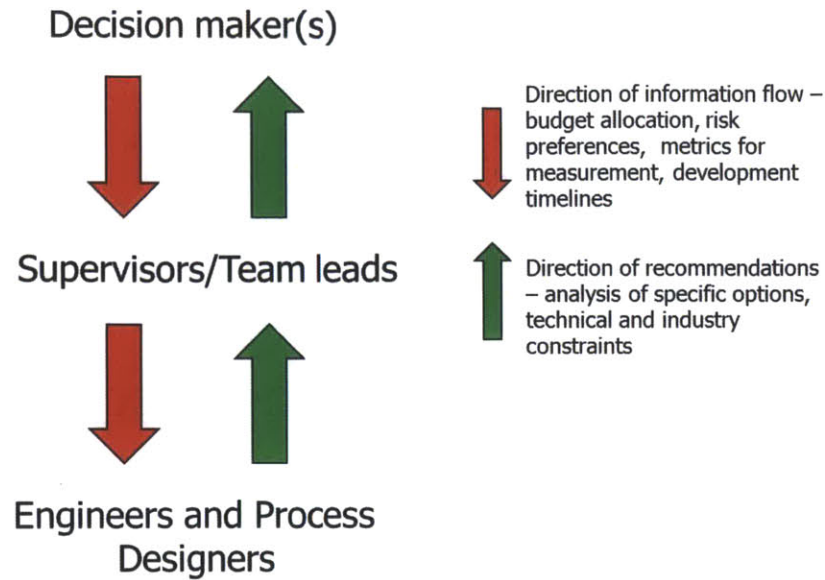
Figure 4-1: Supervision and Information flow during design

4.1 Introduction and History

Uncertainty and the need for making sound decisions in the presence of it has been a key part of human history and most societies have sayings that illustrate some of the accumulated wisdom of how to go about making such decisions[1]. However, it was not until the 17th century that the first steps at establishing prescriptive models for making decisions in the presence of quantifiable uncertainty began. The first model, proposed by Blaise Pascal, was that of making decisions that maximized the **expected value** of the decision outcome under the different possible states of nature - the expected value being a probability weighted average of the different outcomes for each decision.

Thus if a person is confronted with two or more options[2] with different outcomes, he simply calculated the expected value for each option given his knowledge of the outcomes and picked the one with the highest value. Thus for three options described in Table 4.1, the decision with the larger expected value is Y and is the decision that will be recommended under the expected value metric.

Note that we assumed that the states of nature have equal probabilities but it doesn't always have to be the case, the different probabilities can easily be accounted for in the calculation of the expected values in the general equation

---

[1]King Solomon advised to "'Sow your seed in the morning and at evening let not your hands be idle, for you do not know which will succeed, whether this or that or whether both will do equally well"' (Ecc. 11:6) [10]. There is also the old English saying, "'Don't put all your eggs in one basket"', both of which are pieces of advice to hedge against uncertainty in the future

[2]Historically, these options were modelled as lotteries since most of these theories were devised around gambling

State of Nature	Project X	Project Y	Project Z
1	2	5	7
2	4	8	4
3	6	12	9
4	8	20	12
EV =	5	11.25	8

Table 4.1: Outcomes for three different lotteries under different states of nature

$$EV = \sum_{k=1}^{N} p_k x_k \qquad (4.1)$$

where p_k is the probability of state k and x_k its outcome. For projects with a continuous distribution of states and outcomes, we have that EV is obtained as

$$EV = \int_{-\infty}^{\infty} x f(x) dx \qquad (4.2)$$

where $f(x)$ is the probability density function of the distribution.

Daniel Bernoulli, in the 18th century demonstrated a major flaw with the expected value approach to decision making using a now famous example known as the St. Petersburg Paradox decribed by Wakker [122] below:

'Consider the following game: A fair coin will be flipped until the first heads shows up. If heads shows up at the kth flip, then you will receive 2^k. Thus immediate heads gives only $2 and after each tails, the amount doubles. After 19 tails, you are sure to be a millionaire. Think for yourself how much it would be worth to you to play this game'.

We can use the expected value approach of Pascal since we know the probabilities and the value of the outcome and follow Wakker's analysis of the problem. The expected value of the lottery is

$$
\begin{aligned}
EV &= \frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + \frac{1}{16} \times 16 + ... \\
&= \sum_{k=1}^{\infty} 2^{-k} . 2^k \\
&= \infty
\end{aligned}
\qquad (4.3)
$$

Thus the expected value of this lottery is infinite and should be worth more than everything you own to play it. Bernoulli, however, realized that though this was the case, the man in the street wouldn't pay more than the equivalent of $4 to play the game. This led him to formulate the concept of **utility**, the expected value of which people were maximizing, rather than the specific returns. By assuming that the utility to the person of the outcome was the natural logarithm of the returns

rather than the actual monetary value, he was able to derive the value of the lottery as

$$EU = \sum_{k=1}^{\infty} 2^{-k}.ln(2^k)$$
$$= \sum_{k=1}^{\infty} k.2^{-k}.ln(2)$$
$$= 2.ln(2)$$
$$= ln(4) \tag{4.4}$$

Thus the expected utility of the lottery is the same as the utility that a sum of $4 will give to the person, hence the suitable price for the lottery. This attempt at resolving the paradox, was the beginning of the utility and expected utility approach to decision making in the presence of uncertainty.

In 1947, John Von Neumann and Oskar Morgenstern, formalized expected utility theory by delineating a set of four axioms that established the mathematical foundations of the theory and demonstrating how expected utility arises from these axioms. This formalization helped to cement it for most of the 20th century as the backbone of modelling human decision making in the presence of uncertainty.

In the next section, we briefly discuss some of these axioms and illustrate how the decision making behaviour prescribed by expected utility theory and follow that up with how it is used to make decisions. Our discussion will be largely cursory - interested readers who will prefer more rigour are referred to [70], [55] and [94].

4.2   Expected Utility Theory and its Axioms

The axioms that make up Expected utility theory are itemized below and then discussed in further detail:

1. Transitivity

2. Completeness

3. Monotonicity

4. Continuity

5. Independence

While alternative formulations of the axioms of expected utility hold (see [84] and [94] [122]), we are still able to discuss expected utility theory using the above rules.

Transitivity and completeness are necessary axioms needed for consistency from a behavioural point of view of the theory (and decision making in general). Transitivity of decisions simply means that if a decision, A, is preferred to another decision, B,

and B in turn preferred to a third, C, then A should be preferred to C. The absence of this structure on preferences will give rise to a scenario where A is preferred to B, B to C and possibly, C to A leading to cycles of preferences with no clear choice that can be made.

Completeness states that between any two options A and B, there is a clear preference represented by one of three relations: A is preferred to B, B is preferred to A or the decision maker is indifferent between the two. This ensures that a decision can always be made given a set of options - even when the decision implies that any of the choices are good enough (indifference among all options). Without it, it is possible in a set to have a pair of decisions for which nothing can be said about the preference.

Monotonicity, continuity and independence are more directly related to the mathematical structure of the theory. Monotonicity is described by Ross [94] as the following: if a new lottery is preferred to an old lottery if it is simply created by moving probability mass from lower outcomes to a higher outcome. Informally, it is the *dominance* principle highlighted by Peterson [84] as that if under all states of nature, one decision produces an outcome better than another decision, that decision (the first) has a higher utility than the second.

Continuity allows us to express/reduce any outcome for a particular state of nature as a lottery between getting the best outcome and the worst outcome. For example, if a project has 3 possible outcomes with scores 1, 5 and 10 with 1 and 10 being the best (clearly) and 5 being intermediate, continuity states that the utility of outcome 5 can be expressed as a lottery over getting the an outcome 10 and an outcome 1 or, mathematically,

$$u(5) = p \times u(10) + (1 - p) \times u(1) \tag{4.5}$$

In equation 4.5, 'p' represents the probability of getting the best outcome and (1 - p) the probability of getting the worst outcome. Continuity says that for any intermediate outcome between the best and the worst, there should be a value of p that makes the decision maker indifferent to the option to

- either get the original outcome for sure or

- play a lottery where he has the probability , 'p', of getting the best outcome and a probability of '1-p' of getting the worst.

The independence axiom at its heart states that the preference between a choice, A, and another choice, B, should be the same regardless of which other choices there are present in the mix. More specifically, the presence of third decision, C, in the set of options should not affect the preference between A and B. It does not say, for example, that it may not change the choice of the decision selected as C, may, after all, be better than A or B, or both. The preference between A and B however should remain the same.

When all of these axioms are satisfied (or their alternative representations), it can be shown that there exists a utility function that is able to represent the preferences

of the decision maker over the choice set and, in the presence of uncertainty (representable by probability distributions), his decisions will be identical to those made by maximizing the expected utility of the different choices. We won't prove the expected utility theory here but interested readerd are encouraged to look at references [70] [122], [55], [121] and [94] for a more mathematical discussion. Thus given the outcomes of the particular decision over the different states of nature, the probability distributions of those states as well as theutility function over the outcome space, the expected utility of the decision, EU is given by

$$EU = \sum_{k=1}^{N} p_k u(x_k) \tag{4.6}$$

for discrete outcome space and

$$EU = \int_{-\infty}^{\infty} u(x)f(x)dx \tag{4.7}$$

for decisions with a continous space of ourcomes. The choice that will be picked by the decision maker will be one that gives the maximum expected utility. Thus, once we know the utility function of the decision maker, we can use it in making decisions as will be demonstrated in the following section.

4.2.1 Decision making with Expected Utility

To make decisions with expected utility, we need to determine the utility function that represents the preferences of the decision maker and translate each outcome to a utility and then calculate the expectation (the probability-weighted average) to arrive at the expected utility score. This is similar to the expected value approach with one difference - that the expected utility is a function of the utility function of the decision maker and so different projects can have different expected utilities (whereas they all have just one expected value). Take Figure 4-2 with the following decisions whose outcomes are represented by the (continuous) distributions

The outcomes are shown with the expected values of each decision displayed. With the expected value rule, the decision maker will choose project X as it has the highest expected value. Now, let's assume we have three decision makers, each with one of the three utility functions listed below:

$$
\begin{aligned}
U_1(x) &= ln(1+x) \\
U_2(x) &= x \\
U_3(x) &= x + e^{0.07x}
\end{aligned}
\tag{4.8}
$$

These utility functions are displayed in Fig 4-3. We use equations 4.7 and ?? to calculate the expected utilities which are reported in Table 4.2.

From the table, we see that for decision maker 1, option Y maximizes the expected utility while for decision makers 2 and 3, option X maximizes the expected utility. Thus when presented with a choice of all three projects, the decision makers will

66
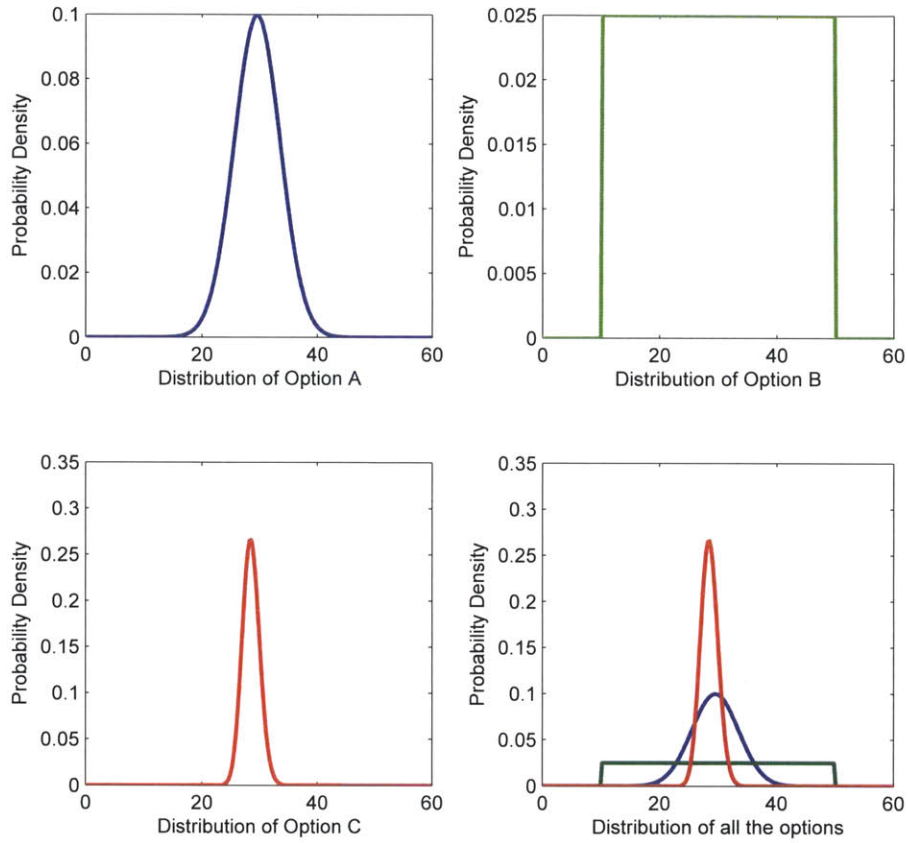
Figure 4-2: Distributions of three different options available to decision makers

Decision Maker	Options X	Option Y	Option Z
1	3.36	3.41	3.38
2	30.06	29.45	28.49
3	41.17	37.64	35.88
Expected value of option	30.03	29.54	28.50

Table 4.2: Expected utilities of the different options for the different decision makers
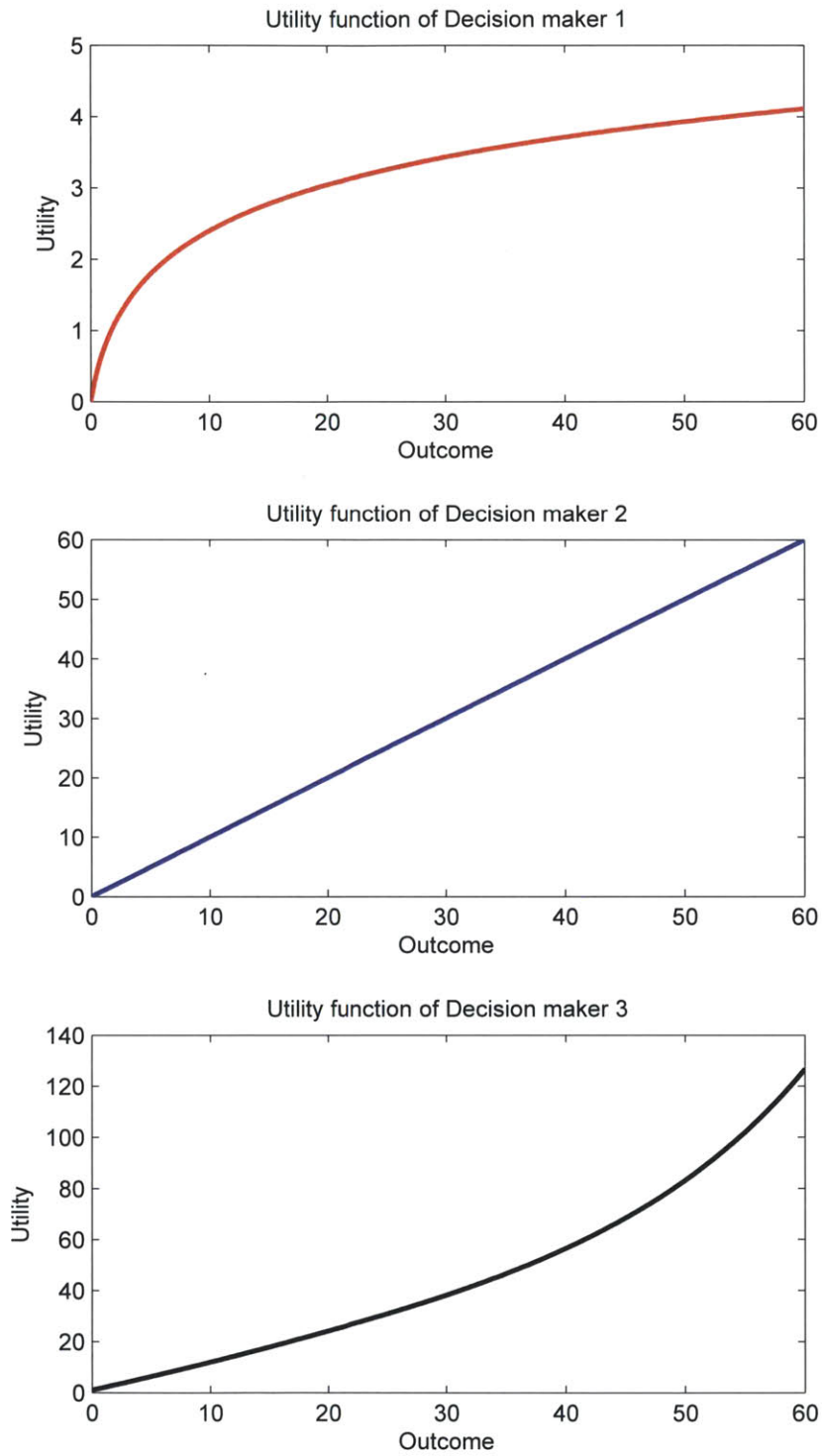
Figure 4-3: Utility functions for the different decision makers
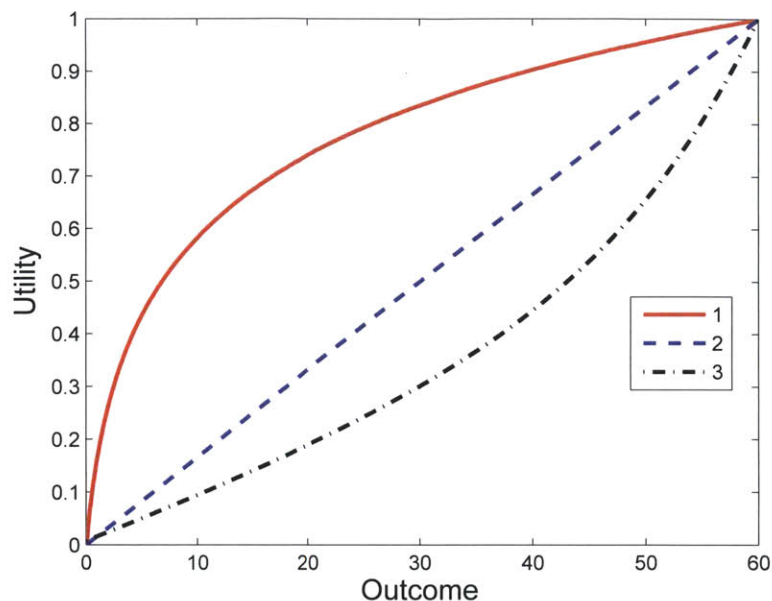
Figure 4-4: Transformation utility functions of the different decision makers

Decision Maker	Option X	Option Y	Option Z	Selection
1	0.854	0.867	0.860	Y
2	0.602	0.591	0.570	X
3	0.497	0.455	0.432	X
Expected value of option	30.03	29.54	28.50	X

Table 4.3: Expected utilities of the different options for the different decision makers

choose accordingly, if they make decisions via expected utility. One thing that comes across clearly in the three different utility curves in Figure 4-3 is that they span different scales - the utility function of decision maker 1 spans from 0-5, that of decision maker 2 from 0-60 and, for decision maker 3, the range goes from 0-200. The absolute values obtained don't have any meaning in particular as all of these utility functions can be rescaled to give utilities that range from 0 to 1 by an appropriate linear transformation, and the choices obtained from using the original forms will be the same as those obtained from the new transformed utility functions. The linear transformation can be obtained by rescaling the utilities[3], giving the minimum outcome a utility of 0 and the maximum outcome a utility of 1. We illustrate the transformed utilities in Figure 4-4 as well as the respective decisions in Table 4.3. As we can see, the choices are identical to those made using the original utility functions.

---

[3]In this case, since all the utilities started from a value of zero, the simplest transformation of scaling by the maximum utility was sufficient.

4.3 Classification of utility functions

Often though, under expected utility, it is possible to classify the general behaviour of the utility function into groups with certain characteristics even when its specific functional form is unknown. There are three main groups that these fall into:

- Risk-seeking utility functions

- Risk-neutral

- Risk averse

These groups are classified with relation to the choice between a lottery and its expected value - which one will be picked. Risk-seeking utility functions represent a group of decision makers who would rather choose a lottery than its expected value for sure. That is, the expected utility of the lottery will exceed the utility of the expected value of the lottery or

$$EU(X)] > U(E[X]) \qquad (4.9)$$

where E represents the expectation operator, U represents the utility function and X the random variable symbolizing the distribution of outcome. Under expected utility, these decision makers' utility functions will everywhere have a second derivative that is positive i.e.

$$U''(x) > 0 \qquad (4.10)$$

Behaviourally, these correspond to people who are more risk seeking and for whom the possibility of a large payoff, even though it is quite slim, dwarfs the downside of a large probability of modest to poor payoffs.

Risk-neutral utility functions represent the group of decision makers that are indifferent between the choice of a lottery or a certain outcome equal to the expected value of the lottery. In other words, the expected utility of the lottery equals the utility of the expected value of the lottery - leading to this indifference. Mathematically, under EU, this is possible only when the utility function is linear which corresponds to a function whose second derivative is everywhere zero i.e.

$$U''(x) \equiv 0 \qquad (4.11)$$

Risk-averse utility functions are the final group and they represent decision makers that will choose the certain expected value outcome over the lottery. Mathematically, under EU, this corresponds to the second derivative everywhere being strictly less than zero i.e.

$$U''(x) < 0 \qquad (4.12)$$

and
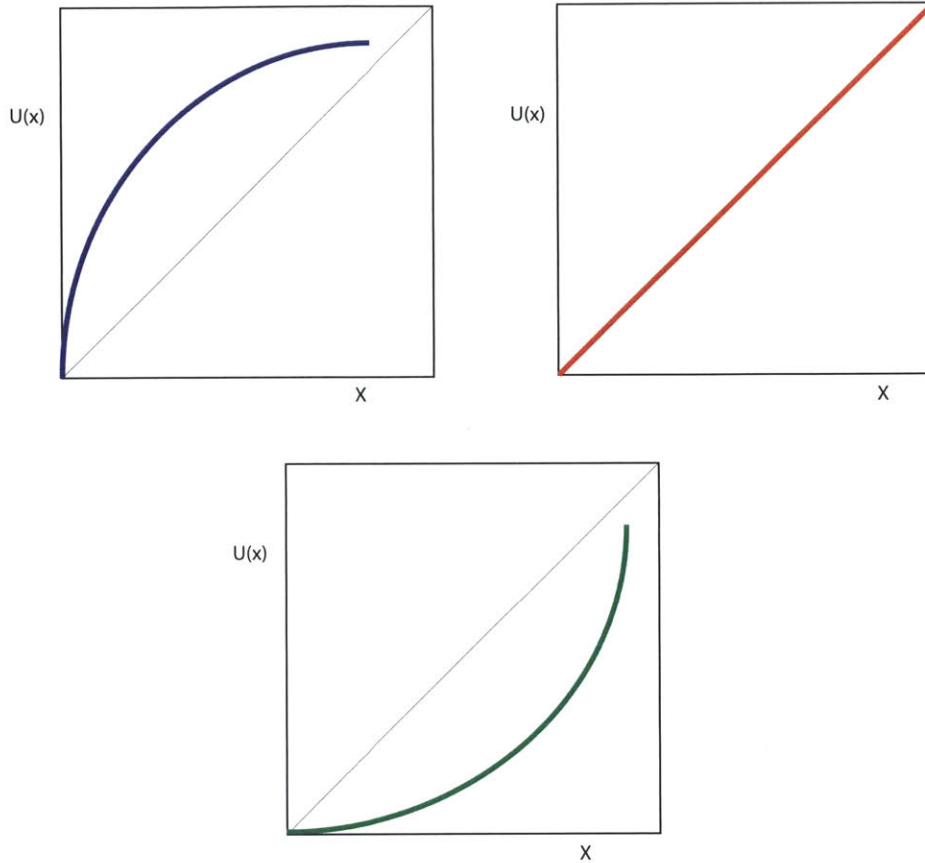
Figure 4-5: Some general shapes of utility preferences. The top left represents the risk-averse decision maker, the top right the risk-neutral and the bottom, the risk-seeking decision maker

$$U(E[x]) > E[U(x)] \qquad (4.13)$$

Most people are risk-averse in making decisions and often all that varies is the degree to which they are risk averse and our focus in this thesis will be on this group. The general shapes of these utility functions are illustrated in Figure 4-5 although there could be variations within each group. For instance the second derivatives don't necessarily have to be monotonically increasing or decreasing (need to illustrate) as long as they satisfy the relationship necessary for the utility function to stay in the same class.

4.4 Decision making without utility functions

Since utility functions help in making decisions, the best way decisions can be analyzed for the decision maker is by extracting his/her utility function and using this function to determine expected utility values. This was the approach taken after the establishment of expected utility as a model for prescriptive choice and differ-

Project	Distribution	Parameters
1	Beta	$\alpha = 3,\ \beta = 6$
2	Lognormal	$\mu = 0.31,\ \sigma = 0.08$
3	Beta	$\alpha = 1.56,\ \beta = 6$
4	Beta	$\alpha = 1,\ \beta = 3.2$
5	Lognormal	$\mu = 0.22,\ \sigma = 0.04$
6	Lognormal	$\mu = 0.27,\ \sigma = 0.12$

Table 4.4: Distributions of projects in Example

ent methods for extracting the utility functions of decisions makers were developed. However, these methods have been fraught with a number of difficulties and has in general had mixed reviews with respect to its use. See Wakker [122] for a discussion of some of them.

Knowing the general class to which the decision maker's utility function belongs is however often much easier to determine - especially within a decision context. within a decision situation what class of utility functions the decision maker's particular utility function will fall into i.e. whether he or she is risk-seeking, risk-neutral or risk-averse.

The big question is this - is it possible to use just the knowledge of the utility class to screen the set of initial options available to the decision maker? Before we attempt to answer this, we introduce another simple example problem that will further illustrate the point.

Example Assume we have a set of project options representable by the following distributions with parameters given in Table 4.4

In addition we pick a family of utility functions represented by the general class of functions

$$U(x) = x^n \tag{4.14}$$

This function has the property that over the outcome range (that is, values of x), we get risk seeking utility functions for $n > 1$ and risk-averse utility functions for $0 < n < 1$ and risk-neutral functions for $n = 1$. Also, for larger n (and conversely, smaller values of $1/n$), we get more strongly risk-seeking (conversely more strongly risk-averse) behaviour, allowing us to explore the different decisions that a range of decision makers with these classes of utility functions will pick. These utility functions are illustrated in Figure 4-6,4-7,4-8.

To make decisions using EU for the different utility functions, the projects outcomes are transformed to utility outcomes using the functions and the expected utility is calculated for each project and for each function. The project that gives the maximum expected utility for each for each utility function is selected as the project to be picked. Table 4.5 presents the decisions selected by each class of decision maker.

The probability densities of the different decisions are presented in Figure 4-9. A number of observations can be made from the above results. First, projects 3
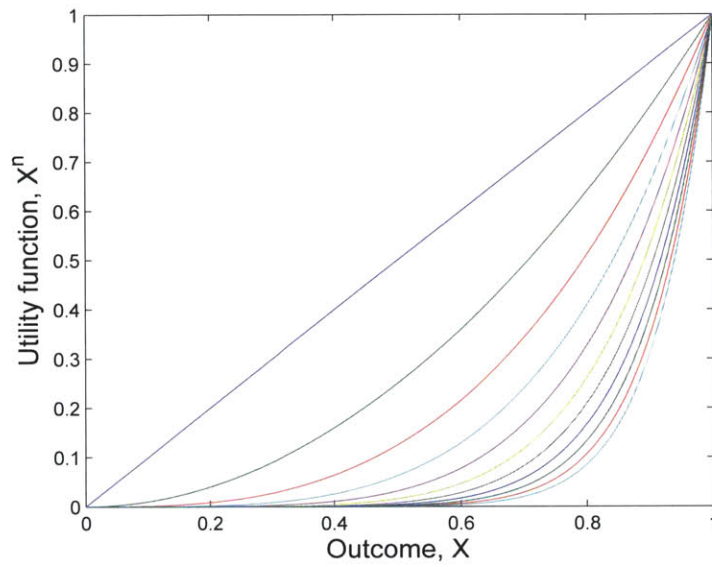
Figure 4-6: Risk-seeking utility functions for the example
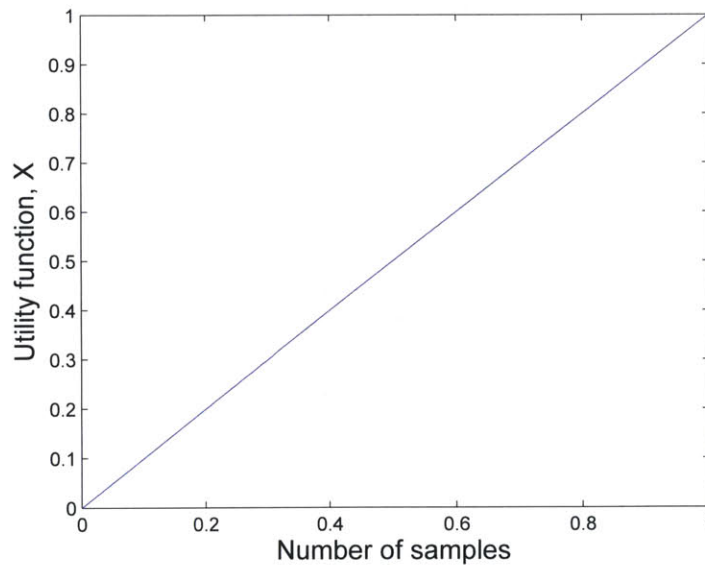
Figure 4-7: Risk-neutral utility function for the example

Utility Class	Selected Projects
Risk-seeking	1, 4, 6
Risk-neutral	1
Risk-seeking	1, 2

Table 4.5: Decisions selected by classes of utility functions
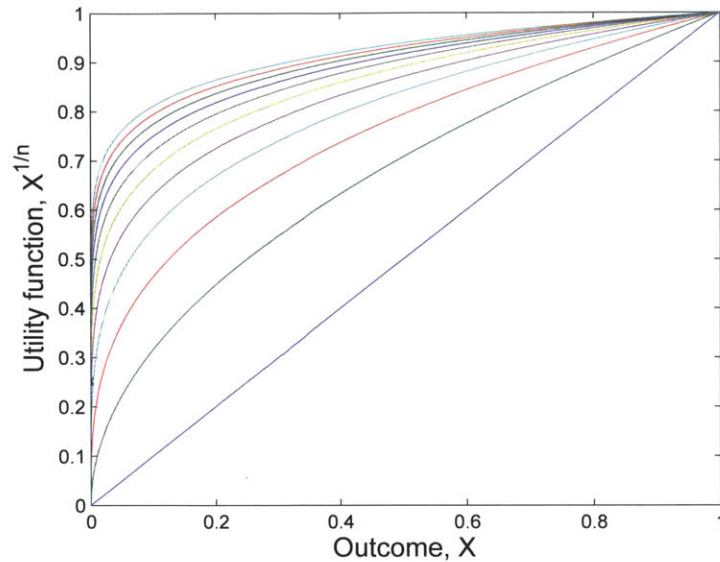
Figure 4-8: Risk-averse utility functions

and 5 are not selected by any of the decision makers. This is because they are each **_dominated_** by another project in the efficient set. This point will be expanded upon in the forthcoming chapter.

For the risk-seeking class, we have that as the decision makers become more risk seeking, they begin by choosing project 1 (least risk-seeking decision) and end up choosing project 6 (the most risk-seeking). This can be more clearly seen in the density functions for each of the projets shown in Figure 4-10. First we should point out that the project means can be ranked as 1, 4, and 2 i.e. project 1 has the largest mean of the three followed by project 4 and then by project 6. But risk-seekers favour projects with higher outcomes rather than means. Thus for the mid-range outcomes (i.e. 0.35-0.7), project 1 gives the probabilities of the highest outcomes so it is the favoured project for some. The more risk-seeking prefer project 4 because beyond 0.7, it gives the largest probability of getting the best reward (and much of the utility function of these decision makers is flat before that as seen in Figure 4-6). The most extreme pick project 6 because, even though the region is not visible in the figure, it has the highest probability of the best outcomes for outcomes really close to 1 - the desire of most risk seekers.

Risk-neutral decision makers care only about the means of the project outcomes so the project that will be picked will simply be the one with the largest mean which is project 1. The distribution around this mean doesn't matter and as such regardless of the spread, the project with the largest mean will be selected.

For risk-averse decision makers, the choice of project starts with project 1 for the less risk-averse decision makers and switches to project 2 as the utility function gets more risk averse. As displayed in Figure 4-11, Project 1 has more extreme outcomes than project 2 - in particular the downside is much worse, with a significant probability of getting outcomes lower than 0.1 (which is virtually non-existent for project 2). Even
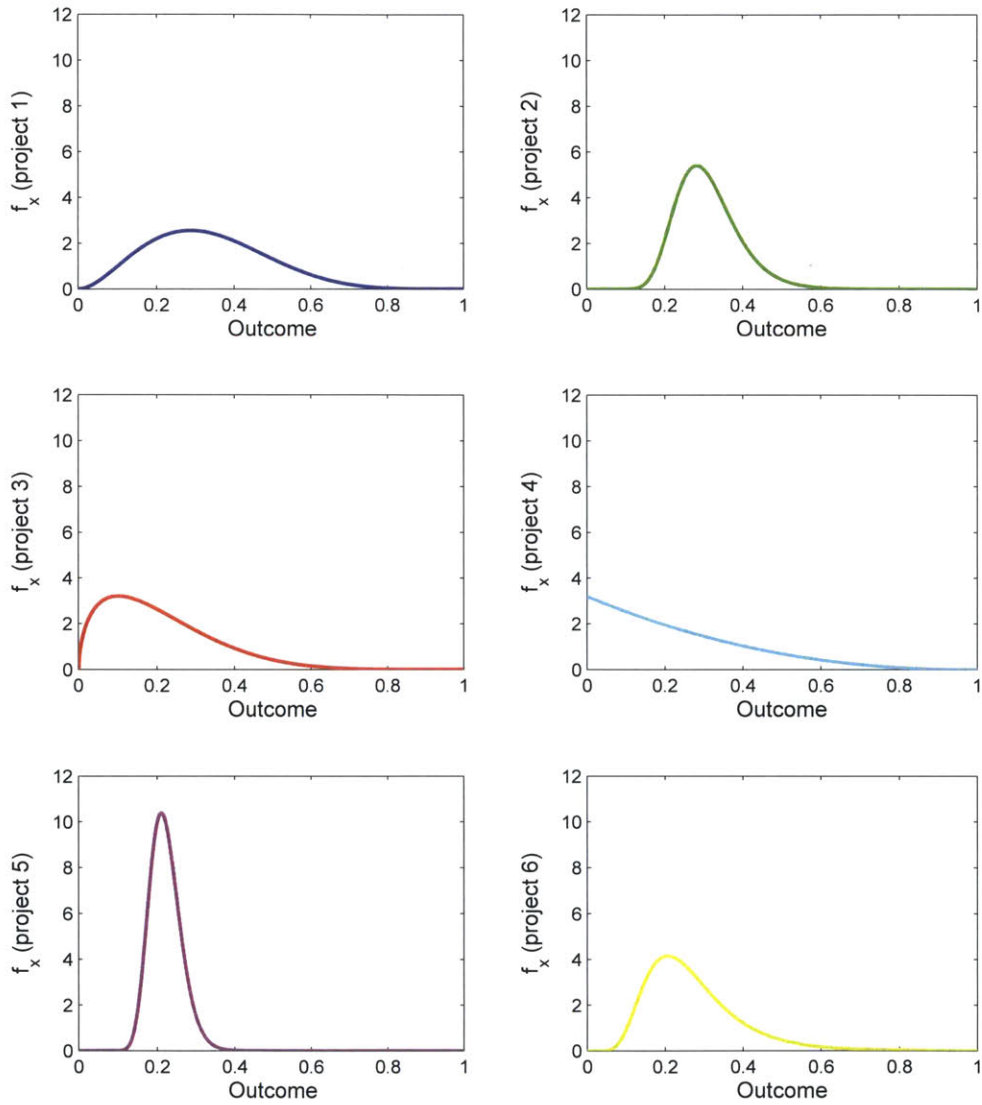
Figure 4-9: Probability densities of the different projects
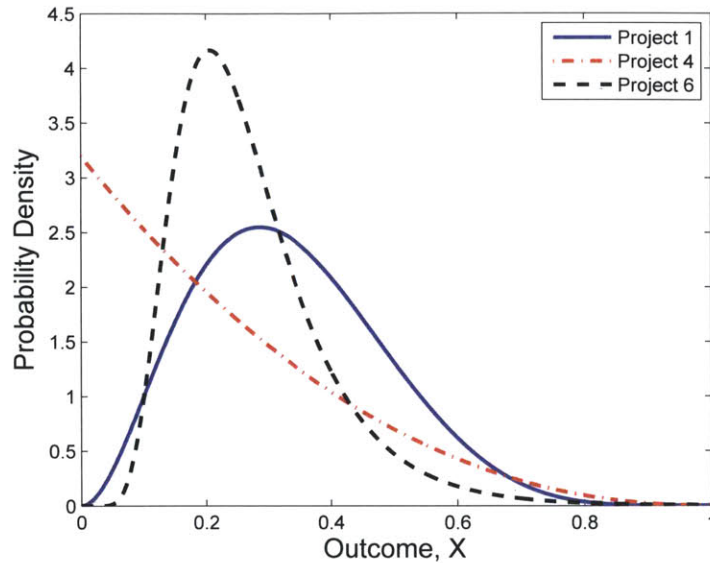
Figure 4-10: Probability densities for the different decisions in the risk seeking decision set

though it also has a better upside as well, the downside matters more as the decision makers get more risk averse and the upside is insufficient to compensate for the risk of the poorer downside.

As we see, the different classes of the utility functions screened the initial decisions appropriately and led to smaller choice sets. However, we were only able to do this by running a large number of utility functions for each class and grouping the decisions. As we mentioned earlier, the particular utility function of the decision maker is often unknown. The set of utility functions of a particular class is also infinite and it will thus be impossible to use the brute force method to appropriately screen the option set to yield the efficient set. A method that allows us to still screen the decisions without having to know utility functions will thus be valuable and we present that method in the next chapter.

4.5   Challenges and improvements to EU Theory

The method we will be presenting in the forthcoming chapter, Stochastic Dominance, largely relies on the fact that utility and expected utility theory provide a good framework for the representation of the preferences of decision makers. This has been the classical view of economic theory so much so that preferences that comply with this framework are called rational preferences. Behaviour that deviated from this model is irrational in the economic sense.

However, since the formalization of expected utility theory, several paradoxes have arisen that challenge its basic axiom. The Allais paradox illustrated that the independence axiom doesn't hold very well in practice for individual decision makers
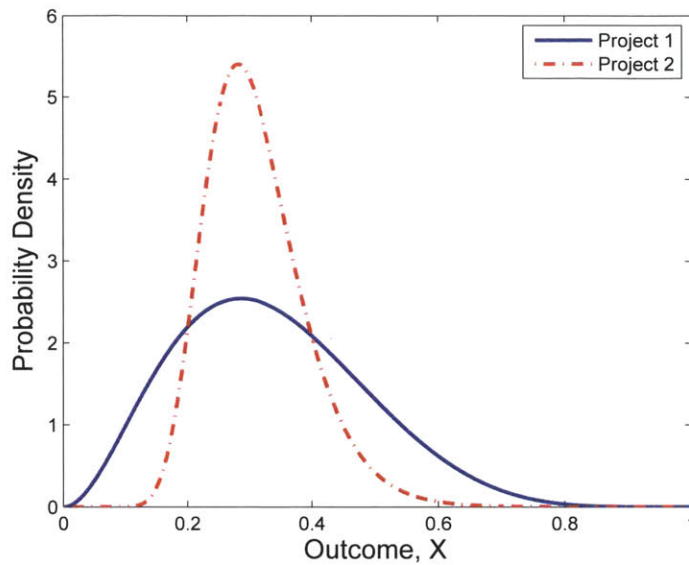
Figure 4-11: Probability densities for the different decisons in the risk averse decision set

because people tend to demonstrate probabilistic sensitivity - overweighting small probabilities and discounting large probabilities [122]. This flaw was rectified by a modified treatment of probability in Rank Dependent Utility Theory developed by John Quiggin [88] where decision weights to the outcomes are not the probabilities of the outcomes as in EU but a weighting function that depends on ranks. This theory and how to apply it is discussed more thoroughly in [89].

Perhaps the biggest challenge to EU though came in the form of Prospect Theory developed by Kahneman and Tversky [52] and represents a comprehensive model for human decision making that is able to incorporate and model irrationality as portrayed by EU, work for which Kahneman was awarded the Nobel Memorial Prize in Economic Sciences [4]. Prospect theory provides two new insights that is ommitted by EU i.e. loss aversion and reference dependence [113] . Loss aversion characterized the fact that in general, humans are more sensitive to losses than gains, and will do more to avoid loss than gain an equivalent amount [114]. Furthermore, what counts as a loss or a gain is a function of the *reference* point from which the change in state is evaluated - the phenomenon of reference dependence [113]. Both of these phenomena lead to contradictions under expected utility theory but are adequately explained by prospect theory.

In prospect theory, decision makers can also have a sigmoidal utility function illustrated in Figure 4-12 - a function that displays both risk-seeking and risk averse behaviour. In line with observations of real decision makers, it is sometimes the case that below a certain threshold, individuals tend to display risk-seeking behaviour

---

[4]Tversky died years before the award and hence could not have been, under the guidelines of the prize, considered posthumously
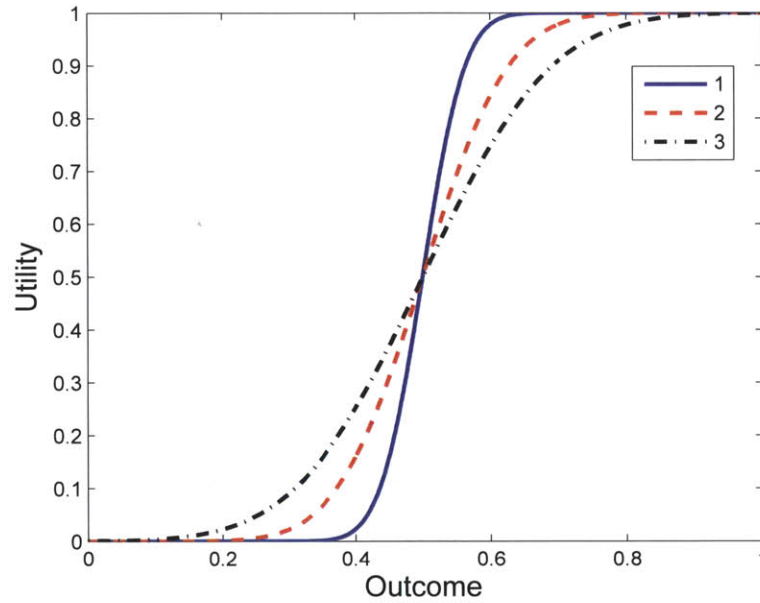
Figure 4-12: Sigmoidal utility functions that display both risk-seeking and risk-averse behaviour

while above that threshold, risk-averse behaviour is observed.

In the figure, we have three utility functions that cross at outcome 0.5. Below this threshold, the functions display risk-seeking behaviour, similar to that seen in Figure 4-6 while above the threshold they display risk-averse behaviour, akin to that observed in Figure 4-8. This joint behaviour is not captured by the utility functions earlier studied and in order for the analysis of different risk preferences to be complete, it is necessary at some point that these functions are included. Levy [64] begins preliminary discussions on how Stochastic Dominance can be applied to utility functions in prospect theory and interested readers are encouraged to look through in more detail. Our analysis doesn't include these utility functions in part because at present, our goal is to come up with a prescriptive approach for how decisions should be made in the presence of uncertainty and prospect theory is still very much a descriptive theory - unlike expected utility. We will thus recommend caution in applying the methods developed in the next chapter in decision situations where expected utility theory fails to adequately describe the decision maker's references.

Chapter 5

Stochastic Dominance

This chapter presents our framework for incorporating knowledge and information (full and partial) of the decision makers into the selection problem in a way that is transparent, that doesn't assume more than is necessary, and can still act as an effective rule for making decisions in the presence of uncertainty.

We begin by defining the approach and the many rules that comprise it. We demonstrate, using the examples from Chapter 2, how it can also be used to make decisions among different projects. After that, we show its link to the preference axioms we highlighted in Chapter 4.

Stochastic Dominance is not without its challenges and, as such, we go further into our analysis by examining them and highlighting how recent developments in the field have been able to overcome some of them.

5.1 Introduction

As demonstrated in the previous chapter, once we know the utility function of a decision maker, it is possible, given a list of options and the uncertainties associated with each, to determine which decision maximizes his expected utility. We did, however, point out that utility functions in general are hard to elicit with known problems associated with trying to do so.

Furthermore, we mentioned that when we lack information about the exact utility function of the decision maker, we often know something about a larger class to which the function belongs. For example, we know that in general utility functions of decision makers are **monotonically increasing**. This results simply from the natural inclination of most people to want more (not less) of a good - especially one, like wealth, that need not be consumed directly (else it suffers from the saturation phenomenon where, beyond a certain point, increasing amounts of something result in a decreasing utility).

As a result, we have the question: is it possible to screen among uncertain projects knowing only the class that the decision maker's utility function belongs to? We can do so and Stochastic Dominance is the general framework that allows us to. But before we discuss the approach in detail, it may be worthwhile to illustrate them with

State	Project X	Project Y	Project Z
1	2	5	7
2	4	8	4
3	6	12	9
4	8	20	12

Table 5.1: Outcomes for three different illustrative projects

Project Y	Project Z
5	4
8	7
12	9
20	12

Table 5.2: Outcomes for three different illustrative projects

some examples. Take a set of three projects X, Y and Z with returns under different states of nature given in Table 5.1.

We assume that every state of nature is equally probable (since there are n states, this give a probability of 1/N). When we compare project X and project Y, we see that even though both projects are uncertain, anyone who prefers more to less (i.e. whose utility function is monotonically increasing) will prefer project Y to X as in every state of nature, the returns of project Y supersede that of X.

Comparing states Y and Z, we observe that the story is a little different i.e. there are states of nature (state 1) where project Y does better than project Z. However, and this difference is crucial, so long as there is no particular preference for states (the utilities capture all the preference), then the scores can be re-ordered as demonstrated in Table 5.2. A look at that table illustrates a similar relationship between Y and Z and X and Y and clearly Y is the superior project.

Thus, knowing nothing other than the monotonically increasing nature of the utility function (as well as state-independence) allowed us to select a project out of three different options. With far more options and more knowledge of the contraints on the class of utility functions, we can do even more, performing a stronger refinement of the utility class and eventually reducing the number of possible options to one that can be selected by the decision maker. The general framework for doing this is what we call **Stochastic Dominance**.

Very simply, stochastic dominance is the application of the dominance principle (discussed in Chapter 2) to uncertain distributions. Stochastic dominance as a decision rule aims to classify decisions into efficient sets using the fact that the utility function satisfies one or both of the characteristics above without assuming any other functional class or form of the utility function. In describing stochastic dominance in detail, we will be comparing two decisions X and Y with distributions $F_X(z)$ and $G_Y(z)$ associated with their outcomes.
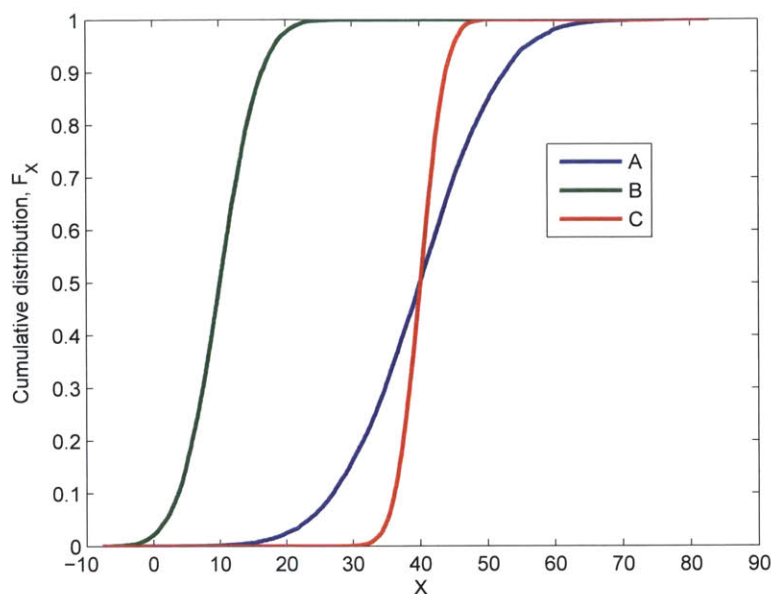
Figure 5-1: Illustration of First order Stochastic Dominance (FSD) using three different projects A, B, C

5.2 Stochastic Dominance orders

There are two principal orderings of Stochastic Dominance: First and Second order Stochastic Dominance. While in principle, there are third and higher orders of stochastic dominance that can exist and theoretically be used (a point that will be expanded upon much later), these two orders are used most often because they retain behavioural intuition about the decision-making process. We discuss each of them below:

5.2.1 First Order Stochastic Dominance

First order stochastic dominance is the rule for all decision makers with strictly increasing utility functions i.e. for all utility functions satisfying $U' \geq 0$, where U' is the first derivative of the utility function, U

For this class of utility functions dominance can be restated in terms of conditions on the distibutions of the outcomes of X and Y. For the FSD, the basic condition is that

$$F_X(z) \leq G_Y(z), \forall z \tag{5.1}$$

where $F_X(z)$ and $G_Y(z)$ are the cumulative distributions of X and Y. In other words, the cumulative distribution of the decision Y, must be to the right of that of decision X. Stated differently, for every possible outcome, z, decision X has a higher probability of generating that outcome, or better, than decision Y. Thus, any decision maker who prefers more to less will choose X over Y.
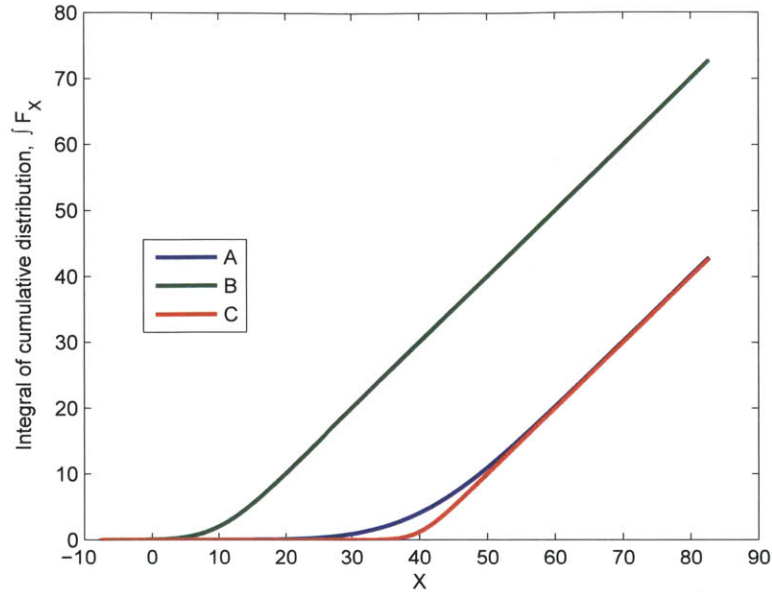
Figure 5-2: Illustration of Second order Stochastic Dominance (FSD) using projects A, B, C

An illustration of FSD is shown in Figure 5-1. Here the cumulative distributions of three different projects, A, B, and C are shown on a single plot. Graphically, FSD corresponds to the cumulative distribution of the dominating distribution lying entirely to the right of the dominated distribution. This is true for A and C with respect to B. Hence A and C dominate B via FSD. However, A and C intersect (near outcome score 33) and so neither lies entirely to the right of the other and so there is no dominance relationship between them. Since there are only three outcomes, A and C thus belong to the FSD efficiency set while B is in the dominated set.

5.2.2 Second order Stochastic Dominance (SSD)

Second order stochastic dominance (SSD) is the rule for all decision makers with risk-averse (concave utility functions). This is the set of utility functions that belong to the set $u \in U_2 : U' \geq 0, U'' \leq 0$.

For SSD, the condition on the distributions is that

$$\int_{-\infty}^{z} F_X(\zeta)d\zeta \leq \int_{-\infty}^{z} G_Y(\zeta)d\zeta \quad \forall z \tag{5.2}$$

i.e. if X dominates Y, then for every outcome level, z, the area under the cumulative distribution of X up to z must be less than that of Y. We again illustrate the SSD relationship using the same projects from the FSD illustration. This is shown in Figure 5-2. As with FSD, SSD graphically corresponds to the integral of the cumulative distribution, (which we will hence refer to as $F^{(2)}$ after Ogryczak and Ruszczynski,
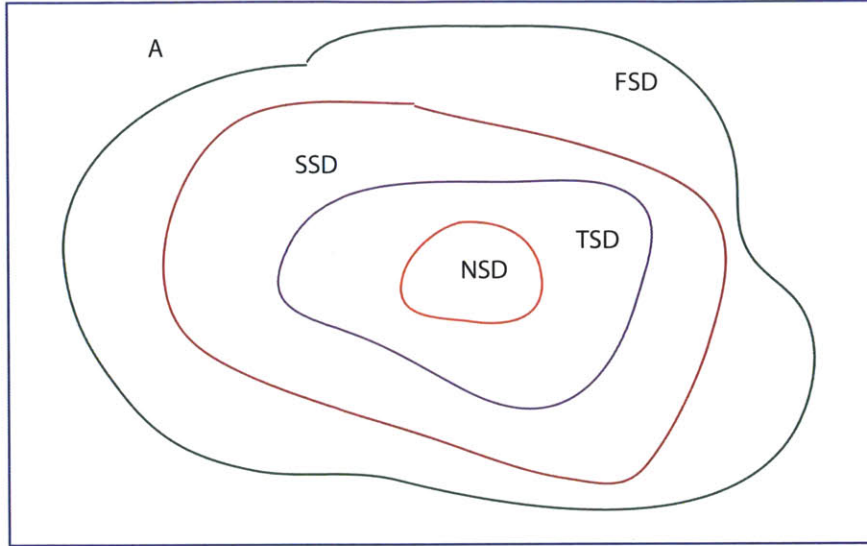
Figure 5-3: Relationships between efficient sets for the different Stochastic Dominance orders

[78] of the dominating distribution lying entirely to the right of the dominated distribution. Again, we see that projects A and C have their $F^{(2)}$ curves lying entirely to the right of project B, illustrating SOSD. This is a result of a larger principle that lower order dominance (in this case, FSD) always implies higher order dominance (here, SSD). In addition, we observe that now project C lies entirely to the right of project A, illustrating dominance by SSD. Thus, the efficient set of SSD in this case contains only project C. This illustrates another general principle that is, in part, a consequence of the first one observed: the efficient sets of higher order stochastic dominance rules are subsets of the efficient sets of lower order rules.

In general, the analytical forms of distributions of X and Y are not known exactly and are obtained by sampling. In that we have a list of outcomes of X and Y and this can be used to construct empirical distributions, $\hat{F}_X(z)$ and $\hat{G}_Y(z)$ which are then used to carry out the comparison. For the rest of our analysis we will be using empirical distributions and won't make any distinction between continuous and discrete distributions.

5.3 Implementation of Stochastic Dominance

To determine whether one random variable stochastically dominates the other, we use the relations established in equations 5.1 and 5.2 in comparing both projects. The algorithms used are presented in Appendix B but the flowchart is presented in Fig 5-4 for easy reference.
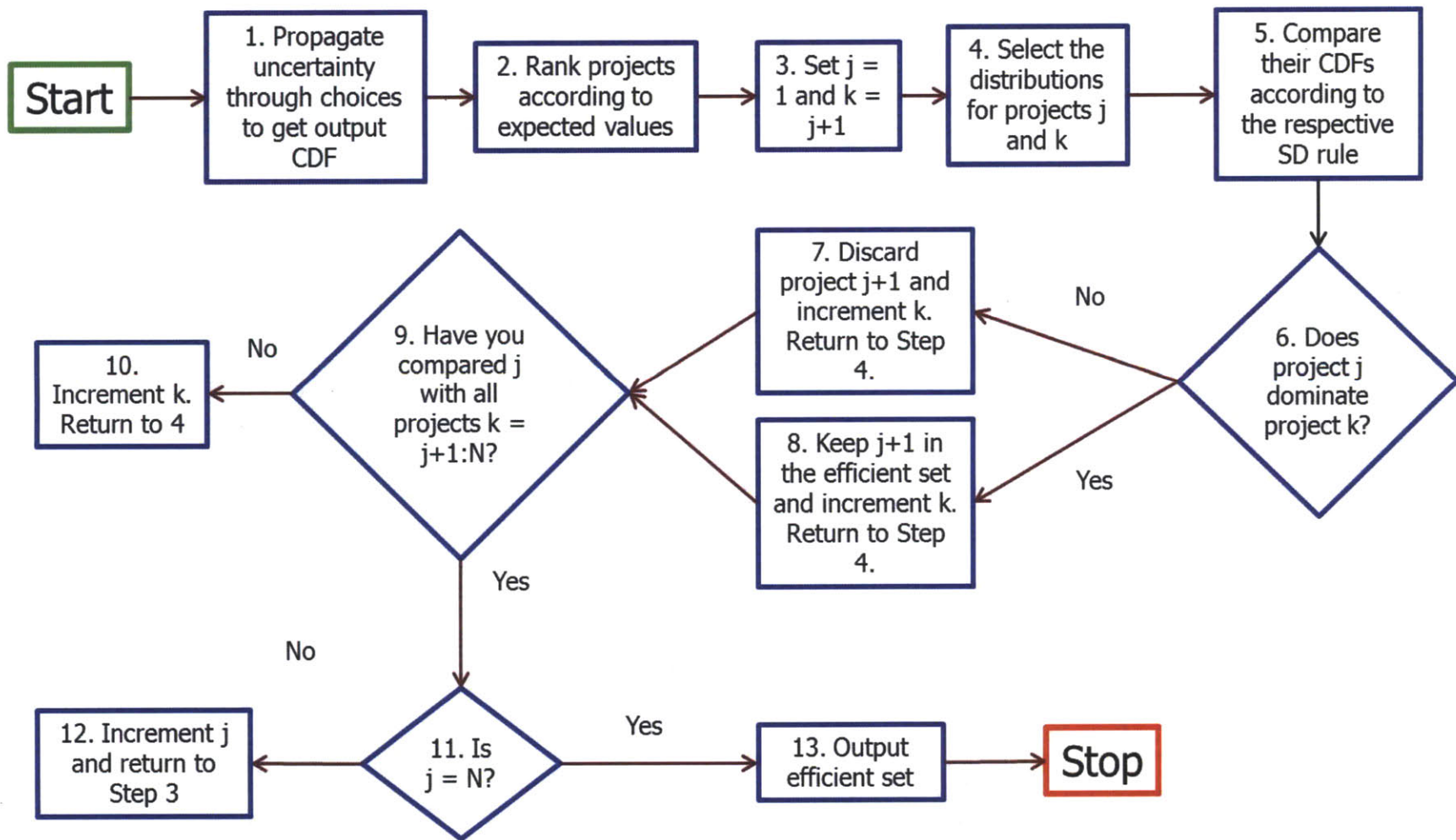
Figure 5-4: Flowchart for Stochastic Dominance Algorithm

Project	Distribution	Parameters
1	Beta	$\alpha = 3$, $\beta = 6$
2	Lognormal	$\mu = 0.31$, $\sigma = 0.08$
3	Beta	$\alpha = 1.56$, $\beta = 6$
4	Beta	$\alpha = 1$, $\beta = 3.2$
5	Lognormal	$\mu = 0.22$, $\sigma = 0.04$
6	Lognormal	$\mu = 0.27$, $\sigma = 0.12$

Table 5.3: Distributions of projects in Example

Stochastic Order	Efficient Set
First Order	1,2,4,6
Second Order	1,2

Table 5.4: Efficient sets for the different Stochastic Orders

If no dominance is observed, higher and higher orders can be used, although as will be discussed in a subsequent section, beyond SSD, the rules don't make much intuitive, economic sense and are largely mathematical artefacts. An example of how Stochastic Dominance is used is illustrated with a simple example below.

5.4 An Illustrative Example

We illustrate the application of Stochastic Dominance to the example in the previous chapter where we had to choose from 6 different projects. The distributions and their parameters are reproduced in Table 5.3

The cumulative distributions are illustrated in Figure 5-5, with all of the distributions shown together in Figure 5-6. We also present the outcome of running the different Stochastic Dominance orders on the decision set in Table 5.4.

A number of things can be observed: First, the SSD efficient set is a subset of the FSD efficient set. This is because the utility functions that characterize the SSD set are a subset of those in the FSD set. This will be discussed in further detail in a later section. Also, if we compare the results in this table to that in Table 4.5 in the previous chapter, we see that the FSD efficient set contains all the decisions selected by all three classes of decision makers. This is not coincidental. All three groups are subsets of the utility functions that characterize the SSD set i.e. they are all monotonic increasing utility functions. Thus since the FSD set will contain all the decisions that can be chosen by any decision maker with a monotonically increasing utility function, it is no surprise that all the decisions make it in.

Another important observation is that the SSD set contains exactly the same set of decisions as those in the set for Risk-averse decision makers. And we obtained this set without needing to test multiple utility functions repeatedly.
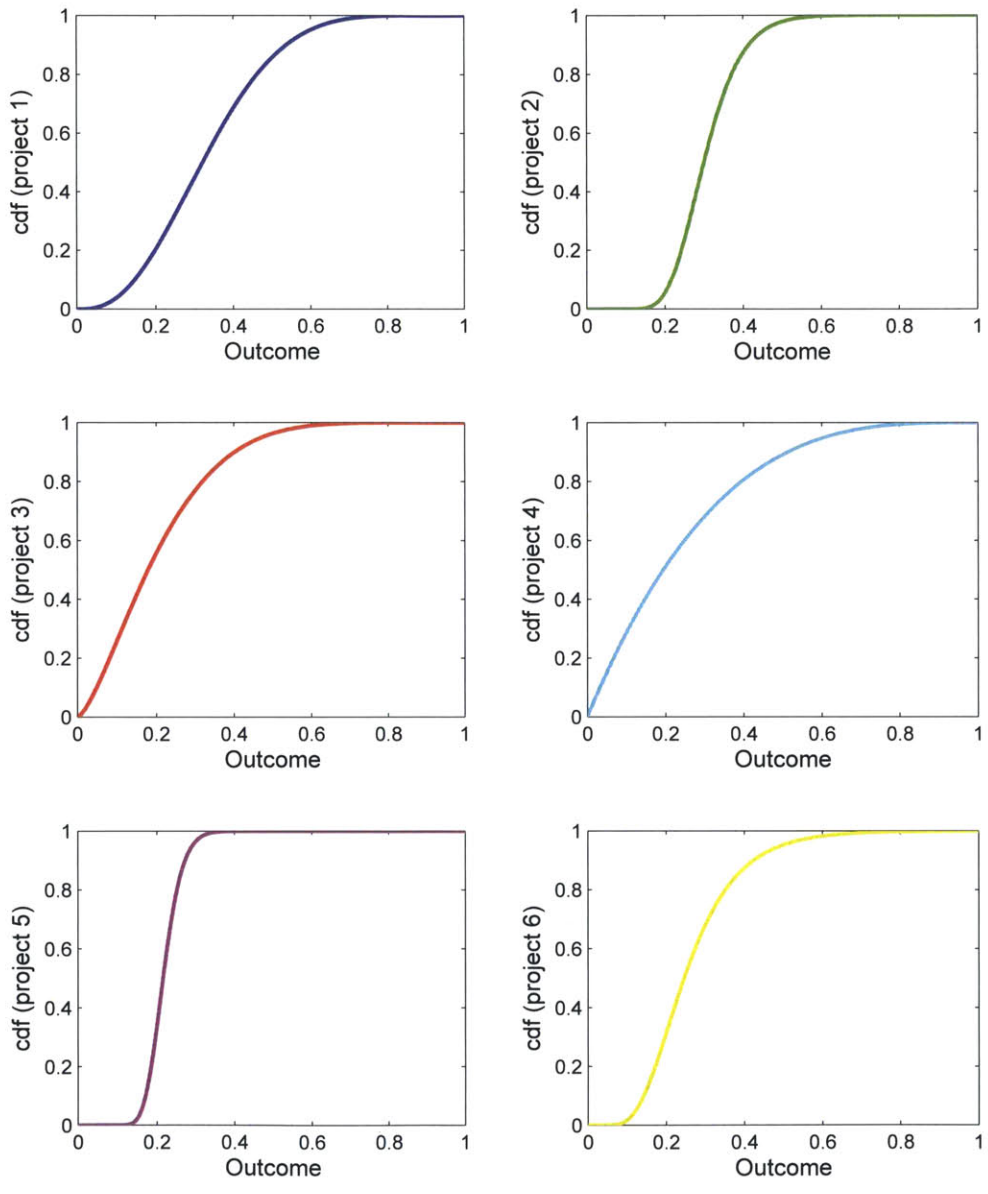
Figure 5-5: Cumulative distributions for the different projects in Table 5.3
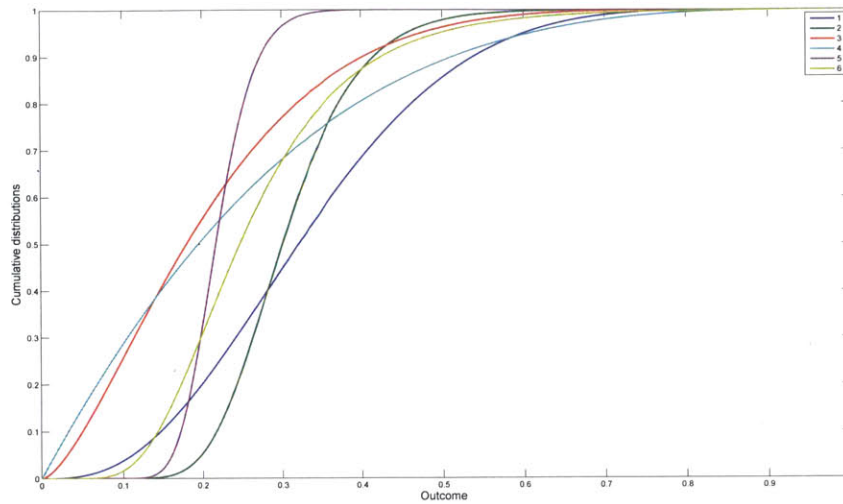
Figure 5-6: Cumulative distribution of all the projects together

5.5 Improvements/Refinements on Stochastic Dominance

Stochastic Dominance In order to address some of the observations, there have been a few modifications to the SD rules. These modifications have been introduced under the heading Almost Stochastic Dominance - ASD (Leshno and Levy [62], Levy [64], Lizyayev [65]). These modifications are meant to eliminate implausible utility functions and prospects that keep SD rules from identifying superior prospects that are more in line with what most decision makers will select.

5.5.1 Necessary and Sufficient Rules for SD

One of the biggest challenges in the implementation of Stochastic Dominance for project decision making is the relative computational time that it takes to carry out relative to the more commonly used metrics like the mean-variance and the expected value formulations. An anecdotal case - for one reasonably sized problem we ran, the Stochastic Dominance framework took about 3 orders of magnitude (close to a 1000 times) longer to compute the feasible set compared to the mean-variance and the oher mean-risk formulations. For larger problems, this difference in computational times may be amplified so that unless better algorithms and developed and more computational infrastructure is devoted to the problem, the SD framework is doomed to be much slower. However, in cases where the cost of infrastructure and computational time is minimal relative to the cost of the implementation of the decision, it may be worth it.

Improvements to speed can be made using necessary and sufficient rules discussed by Levy [64] and these are summarized below:

Necessary and Sufficient rules for First Order SD For first order stochastic dominance of a project F over another project G, there are a number of necessary conditions that must be satisfied

1. The arithmetic mean of F must be greater than the mean of G i.e. $E[F] > E[G]$

2. The geometric mean of F must be greater than the geometric mean of G

3. The smallest value of F must be at least as large as the smallest value in G i.e. $min_F(x) \geq min_G(x)$

Where any of these conditions are violated, FSD of F over G cannot be established and we can avoid the long tedious computation of comparing the cumulative distributions of F and G over their entire outcome space can be avoided.

For FSD, there is essentially only one sufficient rule for establishing the dominance of F over G and that is that the smallest value of F is at least as large largest value of G i.e $min_F(x) \geq max_G(x)$. Note that this is not strict because it assumes that F and G don't contain only one value as this will just be degenerate in both cases.

Necessary and Sufficient rules for Second Order SD The necessary rules for F to dominate G by SSD are that

1. The mean of F must be as large as that of G i.e. $E[F] \geq E[G]$. Note the difference from the necessary rule for FSD.

2. The geometric mean of F must be as large as that of G

3. The smallest value of F must be at least as large as the smallest value in G i.e. $min_F(x) \geq min_G(x)$

The necessary rules in both cases are similar but there is the subtle difference of strict and non-strict inequality in the first two cases. Fishburn [35] shows that if F strictly dominates G by N_{th} degree dominance and both F and G have finite moments through order N, then all the moments cannot be equal and for the first unequal moment, μ_F^k and mu_G^k, the relationship below holds.

$$(-1)^{k-1}\mu_F^k > (-1)^{k-1}\mu_G^k \qquad (5.3)$$

Thus for FSD it is necessary that the means are unequal but it is allowed for SSD. With equal means, dominance by SSD corresponds to inequality in the second moments and by the relationship in 5.3 the second moment of F (the variance) will be smaller than the second moment of G, a relatioship discussed in [95].

5.5.2 Higher order Stochastic Dominance Rules

While our attention in this work will largely be centered around First and Second order Stochastic Dominance, we'll briefly define Third and Higher order stochastic dominance here.

Third Degree Stochastic Dominance Whitmore [125] introduced the concept of Third Order (degree) Stochastic Dominance (TSD) and defines it as the rule for ordering projects for the members of the class of utility functions U_3 for which we have

$$U_3 = \{u : u' > 0, u'' < 0, u''' > 0\} \tag{5.4}$$

and the prospect X being preferred to the prospect Y via TSD if

$$\int_v^a \int_w^a [F(z) - G(z)]dzdy \leq 0 \quad \forall x \in [a, b] \text{ and} \tag{5.5}$$

$$E[X] \geq E[Y] \tag{5.6}$$

TSD is rarely used in practice because of difficulty in getting some economic intuition from it [125]. FSD is the rule for decision makers with monotonically increasing preferences; SSD the rule for those with monotonically increasing and risk averse preferences. The closest preference for which TSD does have some intuition is that for decision makers with a preference for skewness (the section in which [64] discusses it) although in general, it is thought to be of more theoretical interest rather than for practical decision recommendations [125]. Bawa [12] however shows that when comparing prospects with equal means, the TSD rule is the optimal rule to use. Furthermore, the set U_3 listed above is known to contain as a proper subset the class of utility functions in U_2 which also exhibit non-increasing absolute risk aversion [126].

Nth degree Stochastic Dominance Stochastic dominance can be generalized from third to higher orders. First we define as in [35], the recursive relations

$$F^{N+1}(x) = \int_0^x F^N(y)dy \quad \forall x \geq 0 \text{ and } n \in \{1, 2, ..\} \tag{5.7}$$

and the set of utility functions

$$U_N = \{u : (-1)^{i-1}u^i > 0, \quad \forall i = 1, 2, 3, .., N\} \tag{5.8}$$

With these, we have that the prospect X dominates Y by NSD if

$$F^N(z) \leq G^N(z) \quad \forall z \in [0, \infty] \tag{5.9}$$

Since we earlier pointed out that TSD has difficulty in application, it is not expected for the higher order SD rules to be used for decision making except for theoretical interest. However, for completeness sake, it should be noted that they do refine the efficient sets in more detail and can only improve the selectiveness of the final decision that is picked as illustrated much later in Figure 5-3. The challenge however is that the resulting decision will likely be an artifact of mathematical manipulations rather than one that can be supported by economic and behavioural intuitions.

Completely Monotonic Stochastic Dominance (CMSD) So far, we have only analyzed specific classes of utility functions with a fixed number of continuous derivatives - first, second and third. Whitmore [126] introduced yet again a new SD rule for the set of utility functions that have derivatives of all orders i.e the class U_∞ defined as follows:

$$U_\infty = \{u : (-1)^{i-1}u^i > 0, \quad \forall i = 1, 2, 3, .., \infty\} \tag{5.10}$$

To develop the CMSD rule in the manner of the Nth order SD rule, we will have to recursively integrate to infinity according to equations 5.7 and 5.9, a proposition that is quite impossible. Another approach to implementing the CMSD rule will thus have to be found, one which was developed in [125] and is reproduced below:

First [126] presents results from [20] that show that the classes of utility functions that have been of interest have the property of being closed convex cones i.e.

$$\text{For any } u_1, u_2 \in U \text{ and } c_1, c_2 \geq 0, c_1 u_1 + c_2 U_2 \in U \tag{5.11}$$

with the completely monotonic class, U_∞ having said property. With this property, any utility function in a particular class can be written as a positive linear combination of the extremal functions that make up this class. As a result, once we figure out what the extremal functions for each class is, we can derive the necessary conditions for that rule. For example, the extremal utility functions for the FSD rule are the set of utility functions defined as

$$h_a(x) = \begin{cases} 1 & x \geq a \\ 0 & x < a \end{cases} \quad \forall a \in [0, \infty] \tag{5.12}$$

and using this function to calculate expected utilities for a project with distribution F gives

$$E_F[h(a)] = \int_0^\infty h_a(x)dF(x) = 1 - F(a) \tag{5.13}$$

which, when comparing it with another project with cumulative distribution G gives the result

$$1 - F(a) \geq 1 - G(a) \quad \forall a \text{ or} F(a) \leq G(a) \quad \forall a \tag{5.14}$$

which is the more familiar FSD rule that we have discussed earlier. Bernstein's theorem helps provide the extremal utility function of the set U_∞ as

$$h_a(x) = \frac{1 - exp(-ax)}{a} \quad \forall a \in [0, \infty] \tag{5.15}$$

which, when we apply the expected utility operator gives

$$E_F[h(a)] = \frac{1 - M_F(a)}{a} \text{for } a > 0 \tag{5.16}$$

where $M_F(a)$ is the Laplace-Stieltjes transform of the the cumulative distribution, F. CMSD has the advantage that is shows the decision range of the stochastic domi-

90
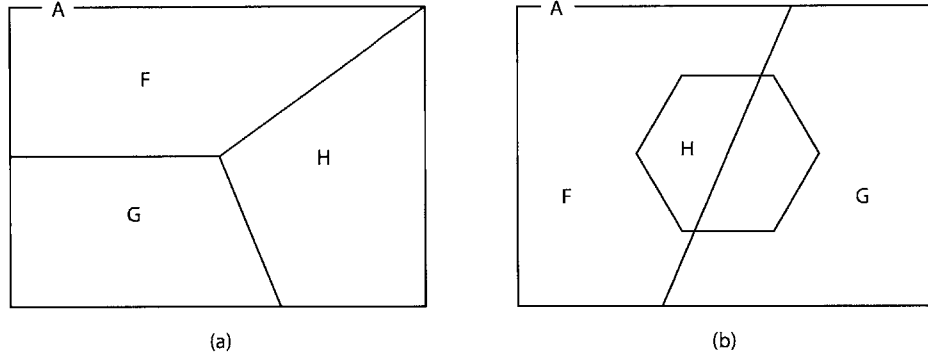
Figure 5-7: Illustration of convex SD

nance rule along a single scale. For when a tends to zero, the rule corresponds to the comparison of the means of the distributions (which is the way a risk neutral decision maker will decide) while as a approaches infinity, the rule becomes the comparison of the worst case scenarios. But as with TSD and NSD, CMSD has largely remained of theoretical interest and has not fully translated to applications in the decision making domain like FSD and SSD.

5.5.3 Convex Stochastic Dominance

Traditional stochastic dominance has largely dealt with admissibility of functions to an efficiency set i.e. the set of decisions that excludes dominated functions determined by partial information of the decision maker's utility function. However, [13] highlights the difference between the admissible set - the set which excludes dominated decisions - and the *optimal* set, which is the set that contains decisions that *will* be picked by a decision-maker with the utility function in the particular class. The idea of such a difference is illustrated in Figure 5-7

The figure is an illustration of a choice between three projects: F,G and H. In the figure, the box, A represents the set of all utility function that belongs to a given class[1], and each labelled region represents the set of utility functions that give rise to a particular choice that any decision maker with a utility function in the class A will pick.

In the figure on the left, F, G and H are disjoint and together cover the entire class, A. When the relevant dominant rule is implemented on each pair of decisions ({F,G}, {G,H} and {F,H}), none will dominate the other because there is a set of utility functions (the region of A covered by the respective choice) for which each decision is optimal. As a result, all the decisions will be admissible i.e. none will be dominated. Also, when presented with a choice of all three, there is a set of decisions where each option will be the optimal choice. admissible set. and optimal efficiency set. Since each decision will be picked, all three decisions furthermore belong to the optimal set

The figure on the right however illustrates a slightly different case. In this scenario,

---

[1]say,for example, the class of all monotonically increasing utility functions·

91

like the one on the left, regular, pairwise dominance will not reveal any dominance because for any pair of decisions, there is a region of A covered by one choice in the pair that is not covered by the other - meaning there are utility functions in A for which one will be superior to the others. But if we take all three together, we see that F and G together cover all of A and it is possible that in the region where H dominates F, G dominates H hence G will be picked and similarly, where H dominates G, F dominates G - leading to F being picked. In that case H will never be picked even though it will belong to the admissible set of decisions. The optimal set (({F,G}) will, in this case, be different from the admissible set ({F,G,H}.

Convex stochastic dominance is a refinement of regular SD that excludes decisions like H, which, though admissible, will never be selected by real decision makers. It was first established by [34] and extended/generalized by [76], [13] and [67].

Like regular stochastic dominance, we can have convex stochastic dominance of different orders. In the following, we give the expressions for when distributions $F_i, i = 1, ..., n$ dominate a distribution, F_{n+1}, (with no loss of generality). For convex first order stochastic dominance (CFSD) we must have a set of λ_i's that satisfy the following [13]

$$\sum_{i=1}^{N} \lambda_i F_i(x) \leq F_{N+1}(x) \forall x \in [a, b]$$

$$s.t. \quad \sum_{i=1}^{N} \lambda_i = 1 \tag{5.17}$$

$$\lambda_i \geq 0, \quad i = 1 \text{ to } N$$

And for convex second order stochastic dominance (CSSD), we have

$$\sum_{i=1}^{N} \lambda_i \int_{a}^{x} F_i(y) dy \leq \int_{a}^{x} F_{N+1}(x) dx \forall x \in [a, b]$$

$$s.t. \sum_{i=1}^{N} \lambda_i = 1 \tag{5.18}$$

$$\lambda_i \geq 0, \quad i = 1 \text{ to } N$$

For the third order stochastic dominance, we have that

Relation	Stochastic order
m = 0	First order
m = 1	Second order
m = 2	Third order

Table 5.5: Equivalence Relations for Lower Partial Moments and Cumulative distributions

$$\sum_{i=1}^{N} \lambda_i \int_a^x \int_a^y F_i(t)dtdy \leq \int_a^x \int_a^y F_{N+1}(t)dtdy \forall x \in [a,b]$$

$$s.t. \sum_{i=1}^{N} \lambda_i = 1 \qquad (5.19)$$

$$\lambda_i \geq 0, \quad i = 1 \text{ to } N$$

As earlier mentioned, our focus will remain first and second order stochastic dominance.

Bawa et al. [13] develop a linear programming algorithm (described below) to determine the first and second order, convex stochastic dominance sets. They do this using the concept of the lower-partial moments defined as

$$L_m(x) = \int_x^a (x - y)^m dF(y) \qquad (5.20)$$

where m is the order of the moment being taken. Integration by parts yields the relations and their equivalence for stochastic dominance presented in Table 5.5

With lower partial moments, the conditions for an option with cumulative distribution F_{N+1} to be eliminated by CFSD or CSSD by options with distributions F_i, for i = 1 to N, can now be generically written as

$$\sum_{i=1}^{N} \lambda_i L_{m-1,i}(x) \leq L_{m-1,N+1} \text{ for } a \leq x \leq b \qquad (5.21)$$

$$\sum_{i=1}^{N} \lambda_i = 1 \qquad (5.22)$$

$$\lambda_i \geq 0, i = 1 \text{ to N} \qquad (5.23)$$

where m is the order of convex stochastic dominance being analyzed (1 for CFSD and 2 for CSSD). It is shown that the inequalites above can be verified by transforming the equations above to a linear program given below

$$Z(m) = \max \sum_{i=1}^{N} \lambda_i \qquad (5.24)$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \lambda_i L_{m-1,i}(x) \leq L_{m-1,N+1} \text{ for } a \leq x \leq b \qquad (5.25)$$

$$\lambda_i \geq 0, i = 1 \text{ to N} \qquad (5.26)$$

If $Z(m) < 1$, then there is no feasible solution to the set of restrictions given by equations 5.23 and as such option F_{n+1} is not dominated by the others via convex stochastic dominance and thus remains in the optimal set. However, if there is a set of λ_i's that solves the equations 5.26 with $Z(M) \geq 1$, then the solution can always be rewritten to satisfy the conditions given by 5.23 by setting $\hat{\lambda}_i = \frac{\lambda_i}{Z(m)}$.

We should also note that to use the linear programming reformulation, it is necessary for the lower partial moment, L_{m-1} to be represented using the discrete points in $[a, b]$. The discrete version of the LPM is given as follows:

$$L_m(x) = \sum_{j|x_j \leq x} (x - x_j)^m p_j \text{ for } a \leq x \leq b \qquad (5.27)$$

Lodwick [67] demonstrates that it is possible to use convex stochastic dominance directly on the original feasible set. This gives a set of N linear programs to be solved to determine the efficient set rather than a comparison of $1/2N(N-1)$ cumulative distributions. However, he concludes that it is generally faster to determine the admissible set first and carry out the optimization on the reduced efficiency sets.

[64] notes cases where using CSD led to the reduction in efficiency sets obtained by using the regular SD algorithms. He observes that the reduction in the FSD efficiency set is only slight but more significant reductions in the SSD (from a third to three-fifths) efficient set are obtained.

5.5.4 Almost Stochastic Dominance

The Stochastic Dominance approach to decision making is very comprehensive because it takes into account all the possible utility functions within each class it is being applied to. However, because of this comprehensiveness, it sometimes doesn't perform very well in terms of decision selectivity i.e. it is still possible for the SD feasible set to contain a large number of decisions. Many times, this is because within the large class of utility functions that are under consideration (for example, the class of *all* monotonically increasing funtions, U_1) there are often some utility functions that even though are mathematically feasible (like the Leontief utility [70] functions that correspond to the worst-case approach to decision making), rarely, if ever, correspond to the utility functions of real decision makers.

To deal with this Leshno and Levy [62] and Lizyayev [65] introduced the concept of Almost Stochastic Dominance, a relaxation of traditional Stochastic Dominance

to deal with such problems. Leshno and Levy define Almost First- and Almost Second order Stochastic Dominance (henceforth referred to as LL-AFSD and LL-ASSD, together known as LL-ASD) as follows:

Definition LL-ASD Let X and Y be two random variables, and F and G denote the cumulative disitribution functions of X and Y respectively. For $0 < \epsilon < 0.5$, we define:

1. **AFSD**: F dominates G by ϵ-Almost FSD if and only if

$$\int_{S_1} [F(t) - G(t)]dt \leq \epsilon \, \|F - G\| \tag{5.28}$$

2. **ASSD**: F dominates G by ϵ-Almost SSD if and only if

$$\int_{S_2} [F(t) - G(t)]dt \leq \epsilon \, \|F - G\| \quad \text{and} \tag{5.29}$$

$$E_F(X) \geq E_G(Y) \tag{5.30}$$

where the sets S_1 and S_2 are defined as follows:

$$S_1(F, G) = \{t \in [0, 1] : G(t) < F(t)\} \tag{5.31}$$

$$S_2(F, G) = \{t \in S_1(F, G) : \int_0^1 G(x)dx < \int_0^1 F(x)dx\} \tag{5.32}$$

and we define $\|F - G\|$ as

$$\|F - G\| = \int_0^1 |F(t) - G(t)| \tag{5.33}$$

In other words, the values of t where we test for LL-AFSD (the set S_1) are those where the normal FSD criteria of F over G doesn't hold and the same holds for the test for LL-ASSD (the set S_2).

In analyzing the LL-ASD formulations, Lizyayev [65] found them computationally intractable for use in portfolio applications and proposes an alternative formulation of the ASD framework which he defines as ϵ-Almost Stochastic Dominance (ϵ-ASD). He define this for both First and Second order Stochastic Dominance as follows:

Definition of ϵ-ASD For two random variables X and Y with distributions F and G respectively we have

1. ϵ-AFSD: A random variable X ϵ-almost dominates a random variable Y by FSD if there exists a non-negative random variable, Z such that $E[Z] \leq \epsilon$ and $X + Z$ dominates Y by FSD

95

2. ϵ-ASSD: X ϵ-almost dominates a random variable Y by SSD if and only if

$$\forall t \in S : F^{(2)}(t) - G^{(2)}(t) \leq \epsilon \qquad (5.34)$$

where $F^{(2)}(t)$ is defined as

$$F^{(2)}(t) = \int_0^t F_X(x)dx \qquad (5.35)$$

ϵ-ASSD can be similarly defined like ϵ-ASD i.e. where we have X ϵ-almost dominate Y by SSD if and only if there exists a non-negative random variable Z such that $X + Z$ dominates Y by SSD. Lizyayev proves the equivalence of both definitions in [65]. For this work, we have chosen to use ϵ-ASD for our analysis as it is both more intuitive and computationally easier to implement than the LL-ASD framework.

The big challenge however with ASD (both ϵ-ASD and LL-ASD) is the proper means of selecting the parameter ϵ. Up to this point, the SD framework has benefited from being largely non-parametric making its application universal to all utility functions in a specific class. With the parameter introduced by the ASD approaches, it can become a subjective decisioin making approach - manipulatable like other subjective parameters.

One idea, inspired by [72] is to make the value of ϵ dependent on the total amount of error in the process. When SD rules are implemented, they are carried out using probability distributions either estimated (or directly implemented) from samples with measurement noise. If we can connect the parameter to an error tolerance in the overall result, then we can ensure the transparency of the selection of ϵ and eliminate personal bias from the analysis process. This is highlighted in Chapter 9 in the possibilities for future work.

5.6 Stochastic Dominance and Objective Functions

In chapter 2, we examined the relationship of the different objective functions with utility theory and saw where they were equivalent i.e. conditions where the recommendations of the objective function corresponds directly to those that will be recommended by expected utility.

Under stochastic dominance, we get an efficiency set of maximal decisions given the class of utility functions that the utility function of the decision maker is known to belong. We can similarly think of the objective functions earlier examined as generating efficient sets (in some cases, giving rise to a single decision) and as such we can discuss the equivalence of these objective functions with stochastic dominance. In general, we seek to determine the objective functions whose efficient sets are either

1. Equal/Identical to the SD efficient set

2. Subsets of the SD efficient set i.e. they are contained entirely in the SD efficient set

3. Supersets of the SD efficient set i.e. the SD set is contained in the efficient set of the decision rule

4. There is some intersection between the SD efficient set and the decision rule

5. There is no intersection between the SD efficient set and the decision rule

These relationships are illustrated in Figure 5-8.

In general, we want decision rules in cases 1 and 2 above because, given the equivalence of Stochastic Dominance and Expected Utility, we are guaranteed that the decision rule would never give suboptimal decisions for a decision maker of a particular type of utility function - even if we don't know what it is. In cases 3 and 4, it is possible that for some utility functions, the decision rule and stochastic dominance will agree, but there is also the possibility that they will disagree and so in some cases where we can determine the range where they will disagree, we can tell where such rules shouldn't be used.

Knowing these relationships ahead of time help in determining a guide to the correct use of these rules in decision making, especially since many of these rules can be implemented much more easily than Stochastic Dominance, more so as we deal with larger scale and more complex problems.

5.6.1 Stochastic Dominance and Single Objective Functions

Single objective functions typically yield a unique ranking for projects, except for the case where the objective gives equal scores to the best projects. In evaluating equivalence with SD for single objective functions, we want to determine if it is possible for a project with a higher score when comparing a pair of projects (indicating it will be picked by the decision rule) to be dominated by the project with a lower score. If that is the case, the rule will have been shown to be inconsistent with SD and as such should be a flag to the users of the rule

Expected value For the **Expected Value** objective function, because one of the conditions of First and Second order stochastic dominance is that $\mu_F > \mu_G$, when project F dominates project G and μ_i is the mean of project i, then the higher mean project (F in this case) will never be dominated by the lower mean project (G), thus, selecting the decision with maximum expected value yields a decision that remains in the SD efficient set. Moreover higher order dominance rules i.e. nth order rules, where $n > 2$ don't eliminate the requirement on the means [63]

Risk Area Ratio (RAR) As discussed in chapter 2, this metric was proposed by Aseeri and Bagajewicz [8] as a way to incorporate risk into decisions. To recall, the
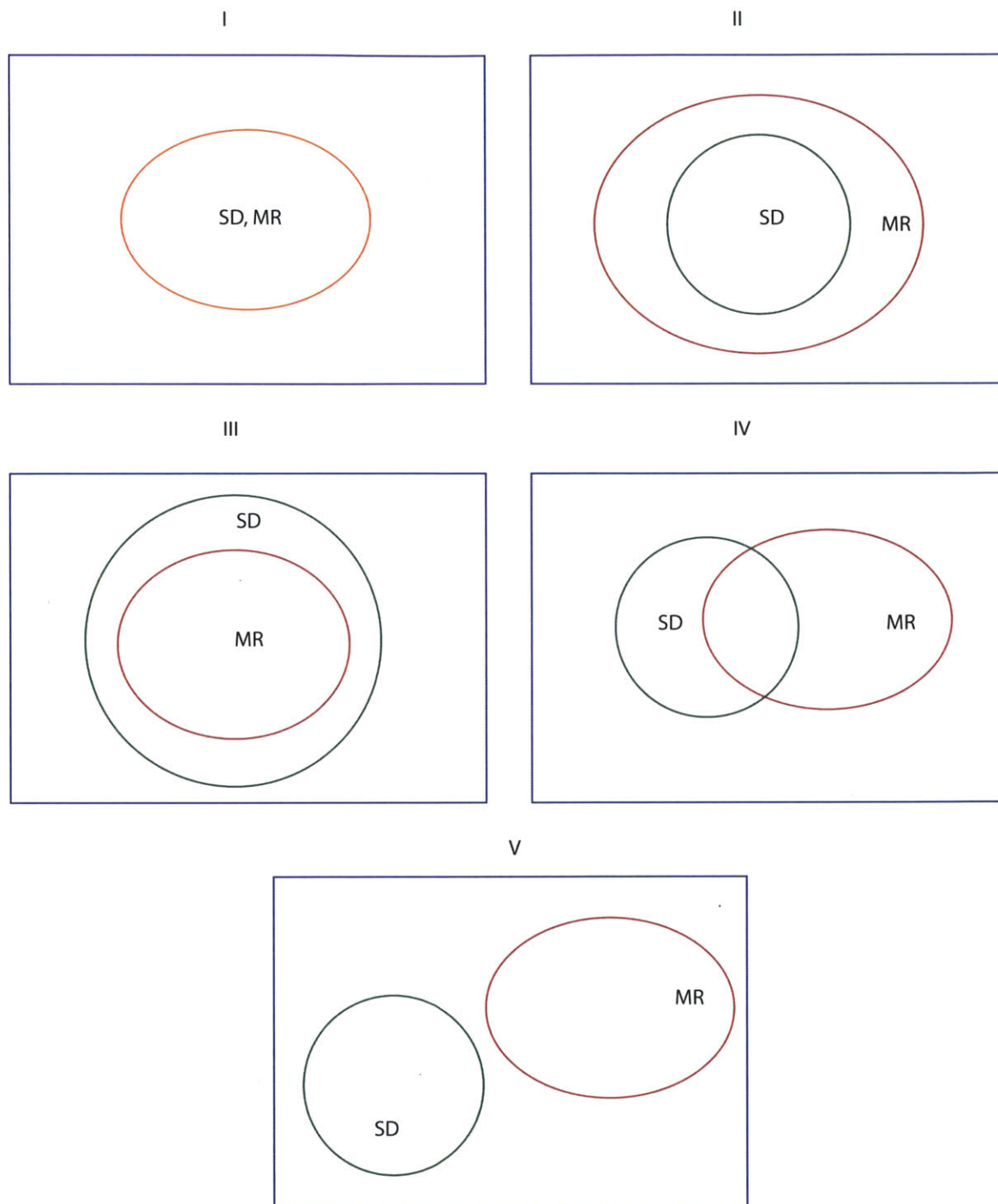
Figure 5-8: Illustration of the different relationships between the Stochastic Dominance efficient set (SD) and the mean-risk efficient sets (MR)
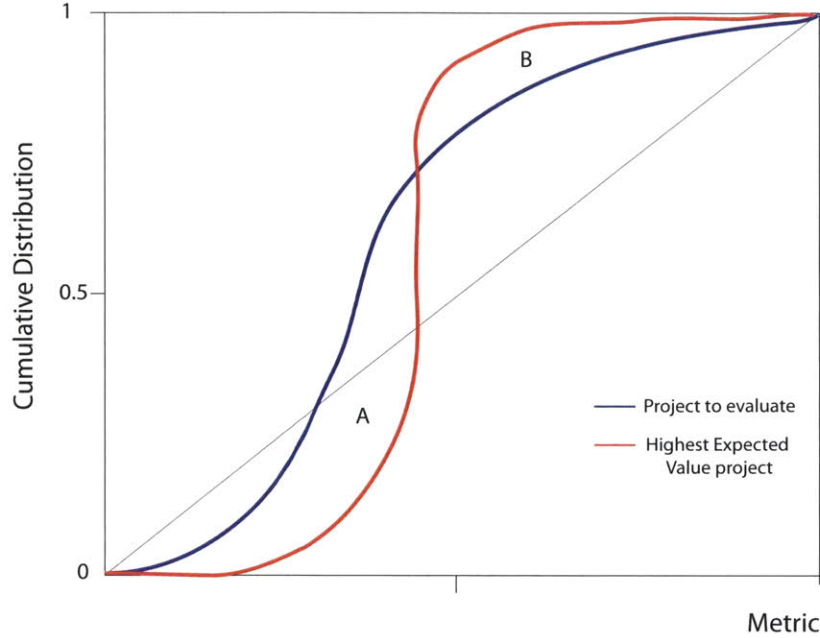
Figure 5-9: Illustration of Risk Area (region labelled A) and Opportunity Area (region labelled B)

defined the RAR as follows

$$RAR = \frac{Opportunity\,area}{Risk\,area} = \frac{\int_{-\infty}^{\infty} \psi^+}{\int_{-\infty}^{\infty} \psi^-} \tag{5.36}$$

$$\text{where } \psi^+ = \max\{F_X - F_Z, 0\}$$
$$\text{and } \psi^- = \max\{F_Z - F_X, 0\}$$

The RAR and the related figures are also displayed again below

From the definition above of the RAR, we can rewrite the first expression in **??** as

$$RAR = \frac{\int_z^b (F - \hat{F})}{\int_a^z (\hat{F} - F)} \tag{5.37}$$

where $\hat{F}$ is the cumulative distribution of the project with the highest mean, F is the distribution of an alternative project being compared against it, $[a, b]$ is the outcome range for the metric. To simplify our analysis, we assume that the two distributions only cross once at some point z, such that $a \leq z \leq b$. We also omitted the variable of integration - the argument of the cumulative distribution.

We can rewrite the above expression as

$$RAR = \frac{\int_a^b F - \int_a^z F - [\int_a^b \hat{F} - \int_a^z \hat{F}]}{\int_a^z (\hat{F} - F)} \tag{5.38}$$

99

With integration by parts we know that

$$\int_a^b F = xF|_b^a - \int_a^b x dF \tag{5.39}$$

Given that $F(a) = 0$ and $F(b) = 1$ and $\int_a^b x dF = \mu_F$ where μ_F is the mean of the distribution, then we have that the above expression becomes

$$
\begin{aligned}
RAR &= \frac{[b - \mu_F] - \int_a^z F - [b - \mu_{\hat{F}} - \int_a^z F]}{\int_a^z (\hat{F} - F)} \\
&= \frac{[b - \mu_F] - [b - \mu_{\hat{F}}]}{\int_a^z (\hat{F} - F)} + 1 \\
&= 1 + \frac{\mu_{\hat{F}} - \mu_F}{\int_a^z (\hat{F} - F)}
\end{aligned}
\tag{5.40}
$$

The final expression of the RAR in equation (5.40) in terms of the means of the distributions illustrates the observation that is made in [8] that the RAR will always be greater than 1. This is because the denominator second term on the right hand side is always positive (for a single point of intersection, z) and the numerator is also positive (by definition, $\mu_{\hat{F}} > \mu_F$), thus making the entire term positive. The RAR equals 1 only when the means of F and $\hat{F}$ are equal.

The decision rule for the RAR is simple: for any two projects, pick the one with the lower RAR. Thus, to test its validity as a rule, we will check if it is possible for an alternative that is inferior in terms of RAR to dominate, via stochastic dominance, another alternative that is superior.

We pick two projects, A and B, whose RARs are defined according to equation (5.40) and, without loss of generality, we assume that project A is preferred to project B according to the RAR metric. In that case, we will have $RAR_A < RAR_B$ or

$$RAR_B - RAR_A \geq 0 \tag{5.41}$$

Substituting the general expression in (5.40) we have

$$\frac{\mu_{\hat{F}} - \mu_{F_B}}{\int_a^{z_B} (\hat{F} - F_B)} - \frac{\mu_{\hat{F}} - \mu_{F_A}}{\int_a^{z_A} (\hat{F} - F_A)} \geq 0 \tag{5.42}$$

where z_A and z_B are the points in $[a\ ]$ where the cumulative distributions of A and B intersect the highest mean cumulative distribution respectively.

We expand this relation to obtain the relation

$$\frac{(\mu_{\hat{F}} - \mu_{F_B})[\int_a^{z_B} (\hat{F} - F_B)] - (\mu_{\hat{F}} - \mu_{F_A})[\int_a^{z_A} (\hat{F} - F_A)]}{(\int_a^{z_A} (\hat{F} - F_A))(\int_a^{z_B} (\hat{F} - F_B))} \geq 0 \tag{5.43}$$

Since we want the expression to be positive, we need the numerator to be (since

100

by construction, the denominator is positive). Hence we have

$$0 \leq (\mu_{\hat{F}} - \mu_{F_B})[\int_a^{z_B} (\hat{F} - F_B)] - (\mu_{\hat{F}} - \mu_{F_A})[\int_a^{z_A} (\hat{F} - F_A)] \tag{5.44}$$

$$= \frac{(\mu_{\hat{F}} - \mu_{F_B})}{(\mu_{\hat{F}} - \mu_{F_A})}[\int_a^{z_B} (\hat{F} - F_B)] - [\int_a^{z_A} (\hat{F} - F_A)] \tag{5.45}$$

Defining a constant, K, as

$$K = \frac{(\mu_{\hat{F}} - \mu_{F_B})}{(\mu_{\hat{F}} - \mu_{F_A})} \tag{5.46}$$

we can substitute into the expression in (5.45) to obtain

$$K \int_a^{z_B} (\hat{F} - F_B) - [\int_a^{z_A} (\hat{F} - F_A)] \geq 0$$

$$\text{or} \tag{5.47}$$

$$K[\int_a^{z_B} (\hat{F} - F_B)] \geq [\int_a^{z_A} (\hat{F} - F_A)]$$

The above inequality can be re-written and transformed to obtain the following relationship

$$K \int_a^{z_B} F_B - \int_a^{z_A} F_A \leq K \int_a^{z_B} \hat{F} - \int_a^{z_A} \hat{F} \tag{5.48}$$

If we compare the above equation to the second order Stochastic dominance requirement for B to dominate A i.e.

$$\int_a^z F_B(\zeta)d\zeta - \int_z^a F_A(\zeta)d\zeta \leq 0, \quad \forall z \tag{5.49}$$

we see that it is possible for B to dominate A by RAR when there is no SSD dominance between them, and, even worse, when A possibly dominates B because we can use the proof to construct an example where this is the case. This is one reason why stochastic dominance is important as a decision rule as it is useful for pruning metrics that turn out to be insufficient.

5.6.2   Stochastic Dominance and Mean-Risk Formulations

So far, we have seen two general approaches to incorporating risk in decision-making: the first is the mean-risk formulation, typified by the classical mean-variance objective proposed by Markowitz [68] and the other is via the use of utility functions, proxied by Stochastic Dominance. A good number of risk metrics that are used in conjunction with the mean/expected value have been generated since variance was first proposed and we discussed some of them in greater detail in Chapter 2.

Both approaches in general lead to an efficient frontier of choices and often do not give unique decisions as recommendations. Mean-risk formulations of the decision problem have the advantage of being intuitive (people can visualize the tradeoffs) and are computationally less burdensome to implement than SD. On the other hand, given its equivalence with utility theory, Stochastic Dominance is more general and it can be used for all types of decisions.

As discussed in the beginning of this section, it will be useful to compare the efficient frontiers of each objective with SD so that in the cases where they are found to be equivalent, we can reduce the computations required by the SD solutions by using the equivalent mean-risk formulation.

In a number of different papers [78] [79] [80] [81]. Ogryczak and Ruszczynski develop the partial equivalence of Second order Stochastic Dominance (SSD) and a number of mean-risk models. To do this, they first define the concept the concept of *mean-risk* dominance [78]. A project X is said to demonstrate mean-risk dominance over a project Y if the following relationship holds

$$X \geq_{\mu/r} Y \iff \mu_X \geq \mu_Y \text{ and } r_X \leq r_Y$$

and the above can be shown to lead to the following implication:

$$X \geq_{\mu/r} Y \Rightarrow \mu_X - \lambda r_X \geq \mu_Y - \lambda r_Y, \quad \forall \lambda > 0 \tag{5.50}$$

where μ_X is the mean and r_X the risk of project X (similarly for Y). We encountered this dominance in Chapter 2 when we first discussed the mean-risk approach to decision making. With mean-risk dominance, they then proceed to define the concept of *consistency* with Stochastic dominance. A mean-risk model is said to be consistent with stochastic dominance if the following relationship holds

$$X \geq_{SSD} Y \Longrightarrow X \geq_{\mu/r} Y \tag{5.51}$$

Since this mode of consistency is often too strong and restrictive (no risk model has been found to completely satisfy this relationship), the alternative was to relax the definition of consistency as follows : The mean-risk model is said to be consistent if there exists a positive constant, α such that for all X and Y

$$X \geq_{SSD} Y \Longrightarrow \mu_X \geq \mu_Y \text{ and } \mu_X - \alpha r_X \geq \mu_Y - \alpha r_Y \tag{5.52}$$

They showed that this relationship guarantees that

$$\mu_X - \lambda r_X > \mu_Y - \lambda r_Y \text{ for some } 0 < \lambda \leq \alpha \Longrightarrow Y_{SSD} X \tag{5.53}$$

From this relationship, they derive the following proposition: If the mean-risk model satisfies 5.52, then except for random variables with identical values for the mean and the risk, every random variable that is maximal by $\mu_Y - \lambda r_Y$ with $0 < \lambda < \alpha$ is efficient under SSD.

The above statement simply means that we can compress the mean-risk objective into a single one, $\mu_Y - \lambda r_Y$, and by maximizing this expression over the permissible

Risk measure	Expression	Consistent with SD?	Value of α
Absolute deviation [58]	$\int_{-\infty}^{\infty} \lvert \xi - \mu_X \rvert f_X(\xi)d\xi$	Yes	1
Absolute semi-deviation	$\frac{1}{2}\int_{-\infty}^{\infty} \lvert \xi - \mu_X \rvert f_X(\xi)d\xi$	Yes	1
Standard semi-deviation	$(\int_{-\infty}^{\mu_X} (\mu_X - \xi) f_X(\xi)d\xi)^{1/2}$	Yes	1
Variance (symmetric)	$\int_{-\infty}^{\infty} (x - \hat{x})^2 f_X(\xi)d\xi$	Yes	1
Variance (non-symmetric)	$\int_{-\infty}^{\infty} (x - \hat{x})^2 f_X(\xi)d\xi$	Depends	$1/h$

Table 5.6: Table of α-values of some mean-risk relationships with Stochastic Dominance

range of λ, we can generate SSD efficient projects. We should note that the consistency only holds for $0 < \lambda < \alpha$. Outside this range, only mean-risk dominance can be determined and it may or may not correspond to SSD.

In order to use this practically, it is important to know which risk metrics it holds for and what the value of α is. These are developed in [78] [79] [80] [81] and summarized below. Before we discuss them though, it will be useful to define the function $F^{(2)}$, defined also in [78] that plays a significant role in the subsequent analyses. This has the form

$$F^2(\eta) = \int_{\eta}^{-\infty} F_X(\xi)d\xi \tag{5.54}$$

As we see recall from the equation ??, we see that second order Stochastic dominace for a prospect, X, occurs over a prospect, Y, when

$$F_X^{(2)}(\eta) \leq F_Y^{(2)}(\eta) \quad \forall \eta \in D \tag{5.55}$$

The function, F^2 (which I will call from here on the second performance function), has a number of properties, principal among which are that it is continuous, convex, nonnegative and non-decreasing. Others are again highlighted in [78] and are summarized in Table 5.6 where $f_X(\xi)$ is the probability density function of the random variable X.

From the outset, variance has been known to be in general, inconsistent with Stochastic Dominance and can only be used as a rule for projects with equal means [35], [95]. With prospects that have unequal means, there is no known relationship with their variances (or standard deviations) that imply stochastic dominance. However, in light of this new definition of consistency, it has been shown that for *symmetric* random variables, variance is also consistent with SD with $\alpha = 1$. This relationship however doesn't hold for general, non symmetric variables. However, if we know a common upper bound, h for the random variables $(X - \mu_X)/\sigma_X$ and $(Y - \mu_Y)/\sigma_Y$, then the variance is in general consistent with $\alpha = 1/h$.

The relationships for the Value-at-Risk and Conditional Value-at-Risk metrics requre the concept of the dual function, $F'^{(-2)}$, of the cumulative distribution function and are examined in more detail in [80] and [81]. The connection between Stochastic Dominance and some of the risk metrics discussed is illustrated in Figure 5-10
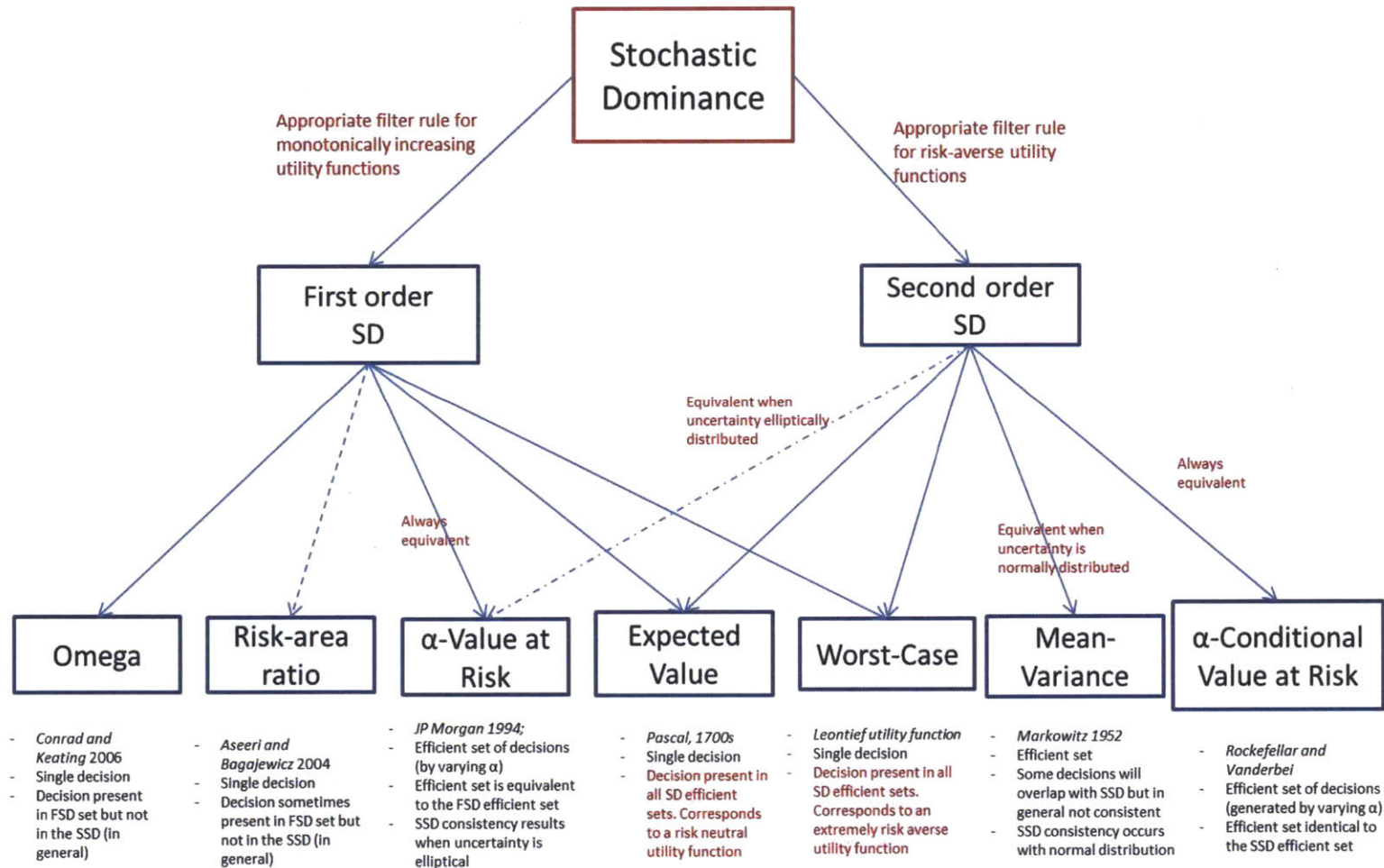
Figure 5-10: The relationship of Stochastic Dominance with other objective functions

5.7 Compuational Cost of Stochastic Dominance

Implementing Stochastic Dominance on project-selection problems, while possessing the advantage of including all preferences and being able to handle all the different forms of uncertainty also comes with added computational cost. In Table 5.7, we estimate the cost of four of the objectives: expected value, mean-variance, first and second order stochastic dominance in order to give a sense of the relative computational burdens imposed by each method. In the analysis, N represents the number of projects to be analyzed while m the number of samples that make up the outcome distribution of each project. Thus if there are 20 possible projects for investment and in the uncertainty analysis, we ran 1000 simulations for each project, then $N = 20$ and $m = 1000$. Our cost calculations also are done in terms of the number of 'arithmetic' operations used to implement each method. An arithmetic operation includes an addition, a square, a comparison, multiplication or division.

The analysis assumes that we have already generated the $N \times m$ matrix that contains the distributions for all the potential projects.The derivation of the expressions are in Appendix B.

Method	Number of arithmetic operations	Relative magnitude
Expected Value	$Nm + \frac{3N}{2}$	1
Mean Variance	$4 \times Nm + \frac{N(N-1)}{2}$	5
FSD	$Nm + N\log N + Nm\log m + \frac{Nm(N-1)}{2}$	1000
SSD	$Nm + N\log N + N(m\log m + \frac{m(m-1)}{2}) + \frac{Nm(N-1)}{2}$	2000

Table 5.7: Comparison of the computational cost of implementing Stochastic Dominance and other objectives

Figure 5-11 shows log-log plots for the expressions in the table.

The computational cost of implementing SD is perhaps one of the main reasons why, after its invention in the early 1970s, it didn't find its way into the mainstream like the other methods (e.g. mean-variance) did. However, computing power has been growing cheaper and cheaper over time as illustrated by Figure 5-12 from [22][2]. As a result, it is our expectation that the barrier to implementing the methodology on full scale problems will continue to fall and will allow it wider acceptance in industry.

The computational cost of the SD method is not the only barrier to consider in thinking about its implementation. In fact, it is often the case that the cost of running the simulation required to generate the matrix of distributions costs far more than the cost of running the SD algorithm. We illustrate this with the biomass-to-liquids case study that will be discussed in full in Chapter 8. In this example, we have roughly 1000 decisions (N) and we run 1000 simulations for each decision in order to obtain the matrix of distributions. In Table 5.8 we display the run times of the different components of the framework

---

[2]The speeds shown represent the speeds for the fastest supercomputers. While we don't expect this to be the typical speed available to the individual/corporate designer, it at least illustrates the trend in the growing amount of computing power available in the industry
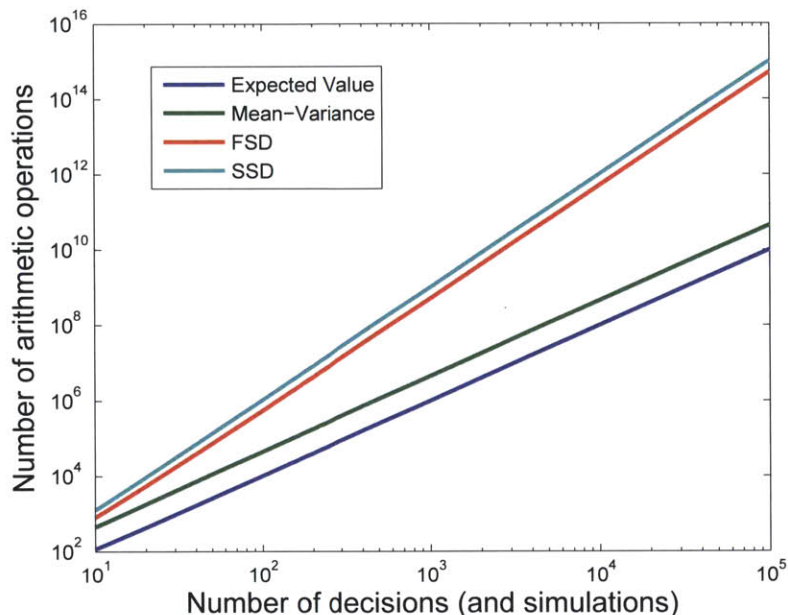
Figure 5-11: Plot of number of arithmetic operations with problem size for the different objectives

Component	Runtime (seconds)	Percentage of overall
Parameter simulation	0.02	0.01
Model run	137.90	93.18
SD analysis	8.01	5.41

Table 5.8: Breakdown of runtimes for different components of the biomass-to-liquids uncertainty analysis framework

To account for some vagaries in computer runtime, we ran the procedure about 3 times and averaged the results. As we can see, the time required for the model simulation vastly outweighs the amount of time required to run the stochastic dominance algorithm, accounting for 93% of the computing time. Improving the speed of model simulations - either by creating better screening models or implementing more efficient methods for uncertainty propagation via the use of alternatives like Polynomial Chaos Expansions (see [47], [38], [108]) may significantly improve the computational challenges encountered.

In this thesis, we didn't consider the use of dynamic uncertainty such as stochastic processes and random walks) in the simulation of the uncertain parameters. The inclusion of this will definitely add some extra burden on the cost, but, as can be seen from Table 5.8, it will still likely be dwarfed by the computational requirements of the model.
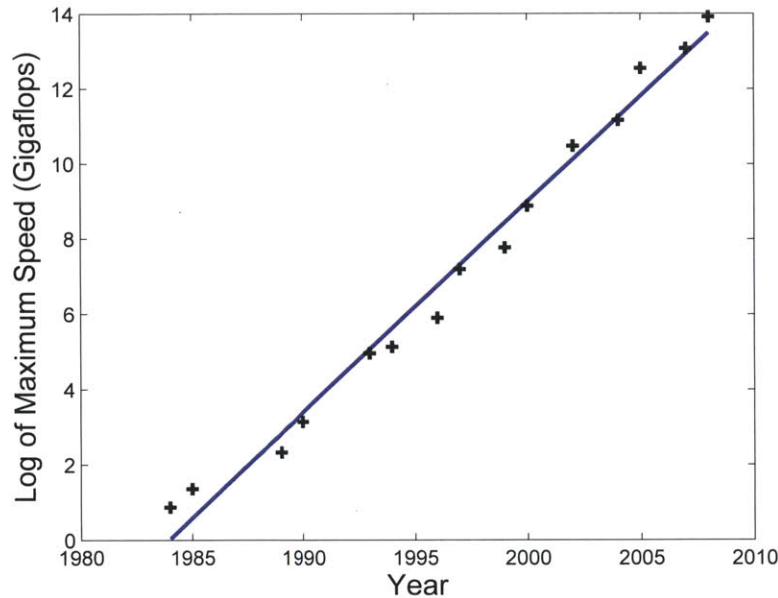
Figure 5-12: Speed of fastest supercomputers over different years. Data from [22]

5.8 Stochastic Dominance and Project Selection

So far, we have demonstrated the utility of Stochastic Dominance as an analysis and a screening tool but done much discussion on how we incorporate it into the process of design and project selection. The goal of this section is to demonstrate how we can combine the stage-gate approach to decision making with the reactor design problem. We illustrate the procedure outlined earlier in Chapter 2 in Figure 5-13.

Gates are the points where projects advance to another level of analysis (indicated by the vertical lines) while the portions between gates where the analysis of the different projects takes place are called stages. Our stage-gate process begins with three necessary inputs

1. Selection of a relevant metric (net present value, internal rate of return etc) on which the projects will be scored

2. Simple process models that convert the material and financial inputs into the relevant scoring metric.

3. A priori information/distributions for the relevant parameters

Our goal is to introduce Stochastic Dominance as a method for introducing preferences in the decision problem. First order stochastic dominance, FSD, serves as a basic screen and can be quickly used to eliminate inferior decisions in the set without any knowledge of the risk preferences of the decision maker. After pruning, the set is checked to see if a unique decision exists and if so, if it is acceptable. If an acceptable decision is found, the process ends. If the solution is unacceptable, the design process terminates. Time and resources are better spent in another design program.

Gate1
Does preliminary analysis suggest viability of some or all projects?

Stage 1

Build input - output models

Select a priori uncertainties

Propagate uncertainty and use outcome to do preliminary screening

Choose a priori Uncertainties

Start

No

Yes

Gate 1

Stage 2

Eliminate
Preliminary analysis indicates project not worth pursuing

Analyze
- Carry out studies and experients to narrow parametric uncertainties
- Proceed to incorporate uncertainties in the new model

Proceed to analyze project choices using outcome uncertainties

Stage 2 continued

Carry out First order Stochastic Dominance to screen projects

Unique Decision?

No

Is there information on risk class of decision maker?

Yes

Carry out Second order Stochastic Dominance to screen projects into respective classes

Yes

Is it acceptable?

No

Proceed to next gate

Yes

No

End

Implement

Terminate

End

Stage 2 continued

Analyze efficient set corresponding to risk class

Unique Decision?

No

Is there more information on risk class of decision maker?

Yes

Use knowledge to further restrict options

Unique Decision?

Yes

Is it acceptable?

No

Proceed to next gate

Is it acceptable?

Yes

No

Yes

No

Implement

Terminate

Implement

Terminate

Stage 3

Refine models

Refine parameters

Repeat analysis from prior stage

Modify selection metrics

Stage N + 1

Iterate process as needed until final project selected

Gate 2

108

Gate N

End

Figure 5-13: An illustration of the stage gate approach to design under uncertainty

With multiple possible projects, we can further prune the set of options using Second order Stochastic Dominance, if we know the class to which the decision maker's risk preferences belong (risk averse, risk neutral or risk seeking). The set of options that result from this will usually be smaller than the FSD set. If it is a unique decision, we can then decide to terminate or implement it. If not, we'll seek more information on the decision maker's preferences and use that to further prune the set. This could be in form of constraints (a certain tolerance for Value-at-Risk, variance etc) or could be via the Almost-Stochastic Dominance restrictions where the ϵ-parameter can be determined. In the case where no knowledge or preferences is present, we can either attempt to learn more about the preferences or advance the remaining projects to the next stage for further analysis.

Once all the knowledge of preferences has been incorporated and we have filtered out as much as we can with no unique process, the resulting projects advance through the next gate (gate 2). Once projects advance through to the next phase, more time and money is spent on refining the choices [116]. In this case, we could build better models, further refine the parameters in the models by carrying out more detailed experiments - say pilot plant studies - and possibly include other selection metrics to help refine the decision. After this is done, we begin an analysis that is similar in spirit to the ones described in the second stage and the process is iterated over as needed until the final design is selected.

The crucial addition in our design approach is the utilization of Stochastic Dominance as the proper screening metrics over alternative rules because these rules can sometimes allow inferior projects to pass through.

5.9   Summary and case-studies overview

This chapter introduced and reviewed the relevant concepts and current ideas in Stochastic Dominance. We examined the rules governing the different orders of Stochastic Dominance, and used examples to illustrate how they are applied.

The next few chapters look at the implementation of the method on three different case-studies: the design of a chemical reactor-separator system, the selection of a crop to grow for biomass production and ultimately the design of a biomass-to-liquids plant. In each of these case-studies we will be examining the performance of Stochastic Dominance first as a screening tool for reducing the size of the feasible set of options and comparing its performance with other objectives.

After generating the efficient set, we will then proceed to analyze the efficient set in two ways: First we will look at different approaches to reducing the size of the efficient set. We will use two different techniques to accomplish this - the use of Almost Stochastic Dominance (ASD) discussed and the repeat examination of the feasible set using Stochastic Dominance with another metric. This approach is illustrated in Figure 5-14.

The second way we analyze the efficient set is to determine the key uncertainties that have the biggest impact on the output uncertainty via a sensitivity analysis. This is important as it allows us to determine the variables that will give us the best
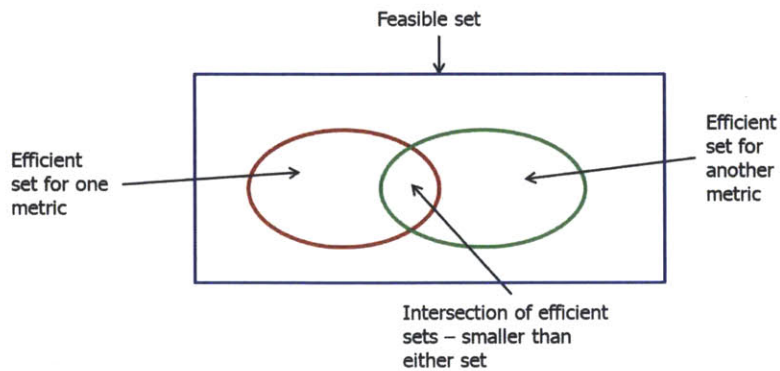
Figure 5-14: Using an alternative metric to prune the efficient set. The intersection of the two sets gives a reduced size set as, at most, the efficient set is as big as the smaller of the two different sets.
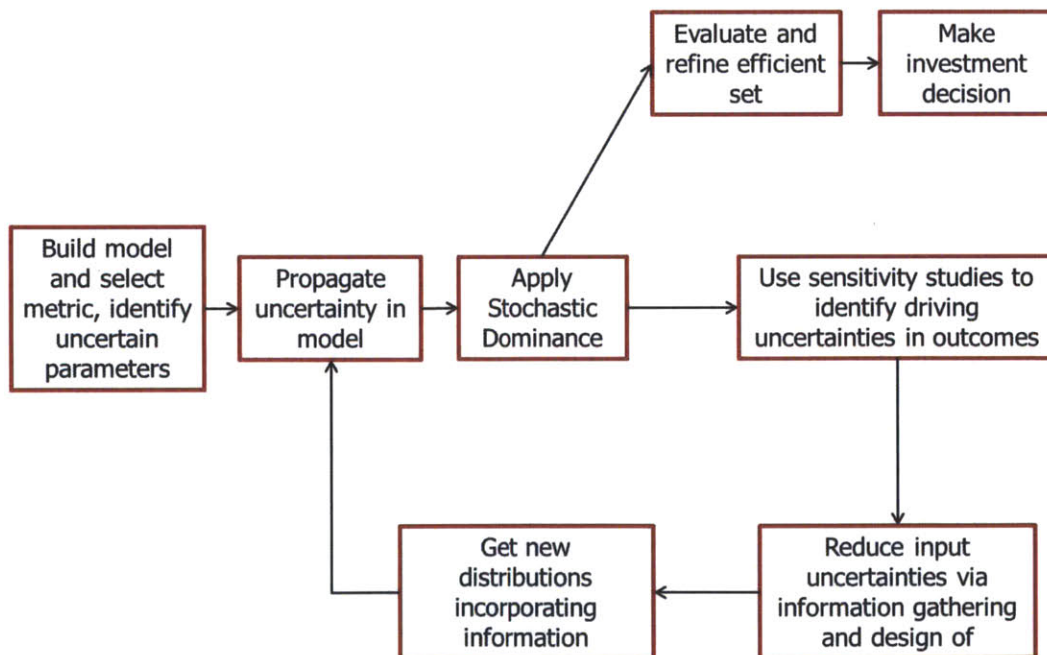
Figure 5-15: Using stochastic dominance with an implementation framework for analyzing the efficient sets

'bang for our buck' if we want to reduce the outcome variability. In addition, once more information is known, we can use it to generate a posterior distribution that can ultimately and the process used iteratively until we can eventually make a decision.

Chapter 6

Case Study 1: Design of a Reactor Separator

In this chapter, we demonstrate the use of the different objective functions analyzed so far on the design of a reactor-separator system for the production of a product C from reactants. We highlight the different decisions that result from the different metrics and show how we use the results of the analysis in determining the next steps in the design process.

6.1 Problem Description

The problem is centered around the decision to produce one of 2 products, C or D. C can be produced directly from a reactant A with simple first order kinetics i.e.

$$A \longrightarrow C$$
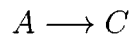
while the D is produced from the combination of two reactants A and B via second order kinetics i.e.

$$A + B \longrightarrow D$$
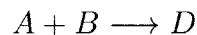
In addition to the choice of product, there are two other choices that together contribute to the decision to be made and these are

- Choice of reactor type: Continuous Stirred Tank Reactor (CSTR) or a Plug Flow Reactor (PFR)

- Reactor capacity/volume

These choices are illustrated in Figure 6-1. The thick red arrow shows the three choices that together constitute a decision i.e. the choice of product C, manufactured via a CSTR and a particular reactor volume choice.

The basic design structure for the project is a reactor-separator system illustrated (for the A to C system) in Figure A-1. The second case is illustrated in the Appendix A.
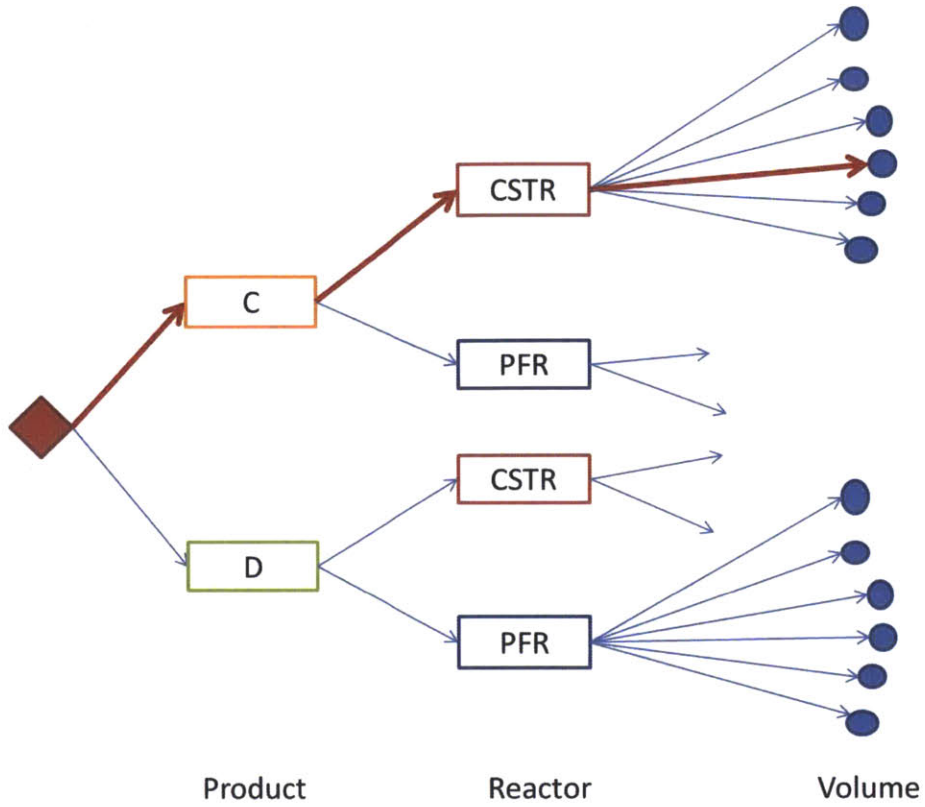
Figure 6-1: An illustration of the choices that constitute a project decision
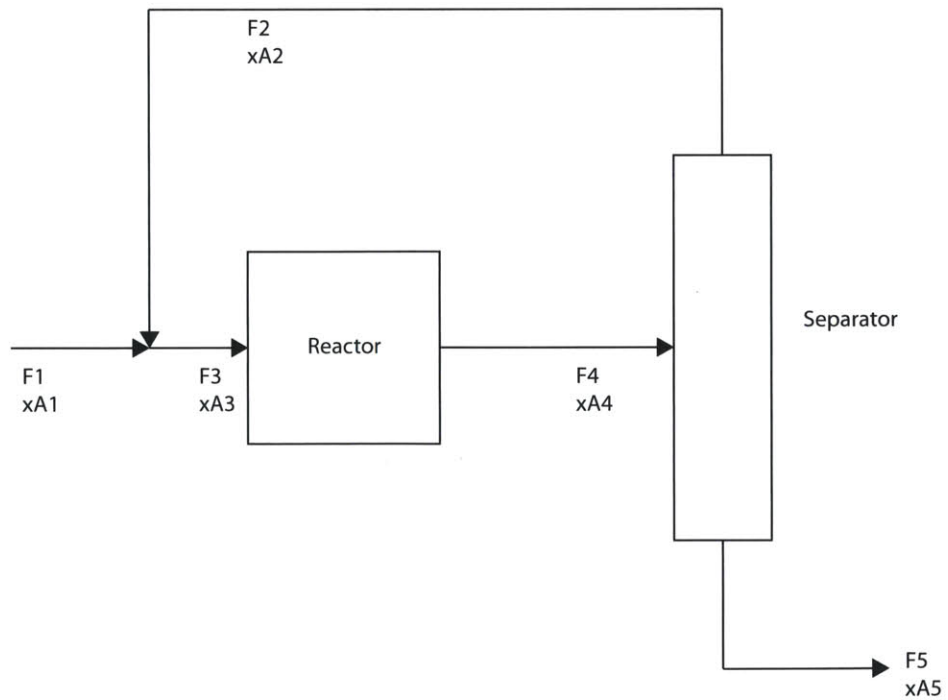
Figure 6-2: Model of reactor-separator system for a single reaction
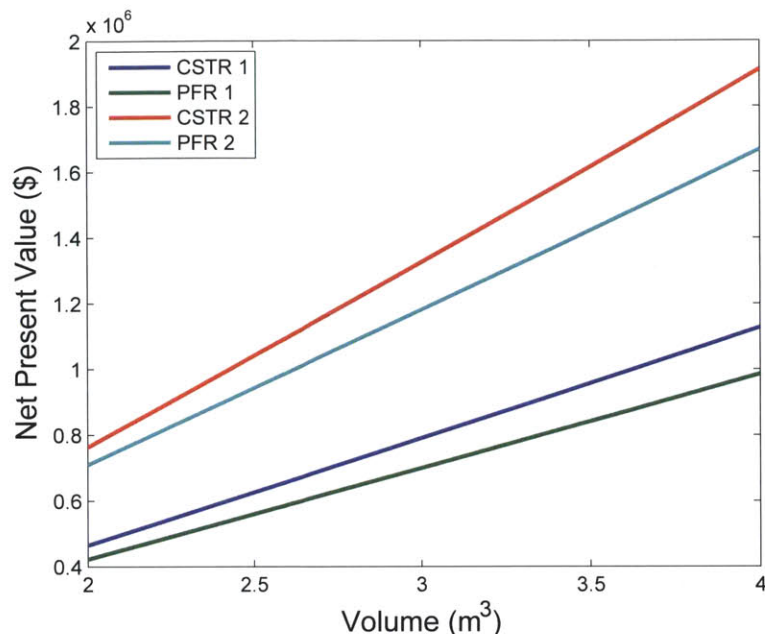
Figure 6-3: Net Present Values for different volumes and different process choices. 1 and 2 represent the manufacture of C and D respectively

The metric chosen for the evaluation of each project is the net present value (NPV). In order to analyze the decisions, we need to build technical and economic models that show how the inputs (raw materials A and B) are converted to the desired output (project NPV). These models are described in greater detail in Appendix A. To keep the problem simple, we made a few assumptions that are highlighted below:

First we assume that the choices are independent and mutually exclusive (the company ultimately wants to choose only one project for execution). We also assume that there are no cost constraints associated with any choice - all financial and demand constraints are included in the lower and upper bounds for the volume. In the technical models, we have focused just on mass balances and have omitted energy balances assuming costs can be adequately incorporated into the cost coefficients for the different stages. All of these were done in order to keep the system as simple as needed so that an adequate understanding of the Stochastic Dominance methodology is not lost in the complexity of the demonstrating examples.

6.2 Deterministic Analysis

With nominal values of the parameters for the problem we can evaluate each of the projects deterministically and compare their performance. Parameter values for the respective systems are given in Tables A.1 and A.2 in Appendix A. Figure 6-3 compares the Net Present values for different reactor volumes for the different decisions.

All the decisions show the NPV increasing as the volume increases largely because

115
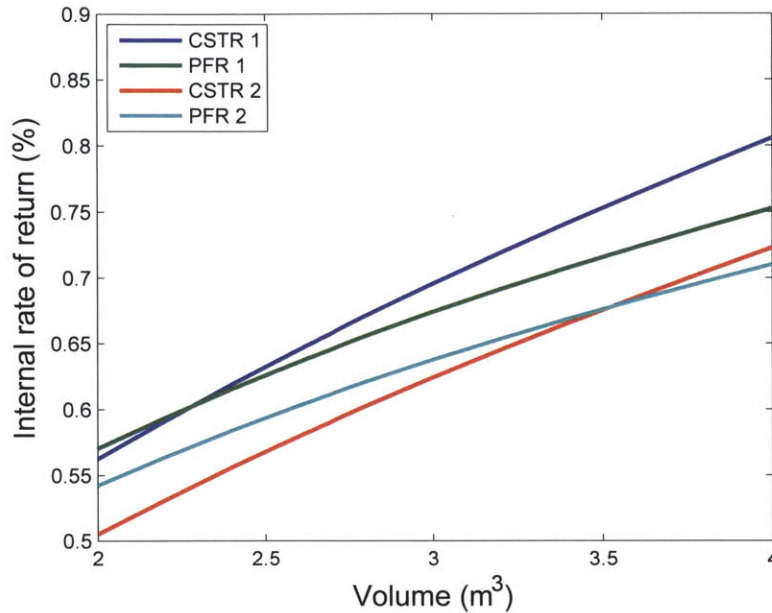
Figure 6-4: Internal rate of return for different volumes and different process choices. 1 and 2 represent the manufacture of C and D respectively

of the larger product output generated which translates to more sales. The CSTR and PFR for production of D (labelled as '2' in the plots) outperfom those for the production of C using NPV. And we can also see that the CSTR for the production of D outperforms that of the PFR for all the volumes considered. Thus, in the deterministic case, using the net present value as the metric of choice, the decision is simple - choose to produce D with a CSTR using a reactor volume of 4 m^3. However the story is a bit different when we use a different metric like the internal rate of return, IRR as shown in Figure 6-4. Here, the CSTR configuration for product C almost uniformly outperforms that of the others except for reactor volumes lower than 2.4m^3 where the PFR reactor for product C outperfomrs it. This example goes to illustrate the point made in Section 2.3 about how the choice of the metric can affect the selection of the product. In the one case (NPV), the choice would have unanimously been the CSTR for the production of D while in the other case, it would have been CSTR for product 2.

Capital costs can also form one consideration for the selection of projects. We illustrate the costs for the different projects in figure 6-5. In this case, we see that the project with the lowest capital cost is the PFR for the production of C. We have assumed that cost is not a consideration in this problem but if is and the company has budget constraints, then the CSTR and PFR configurations for the production of C now starts to look more attractive than those for the production of D.

In summary, while deterministic metrics can give concrete recommendations about choices to select, it is important to recognize that these recommendations are a function of the metrics themselves and proper care should be taken to ensure that correct metrics that represent the best approach for ranking the projects under consideration

116
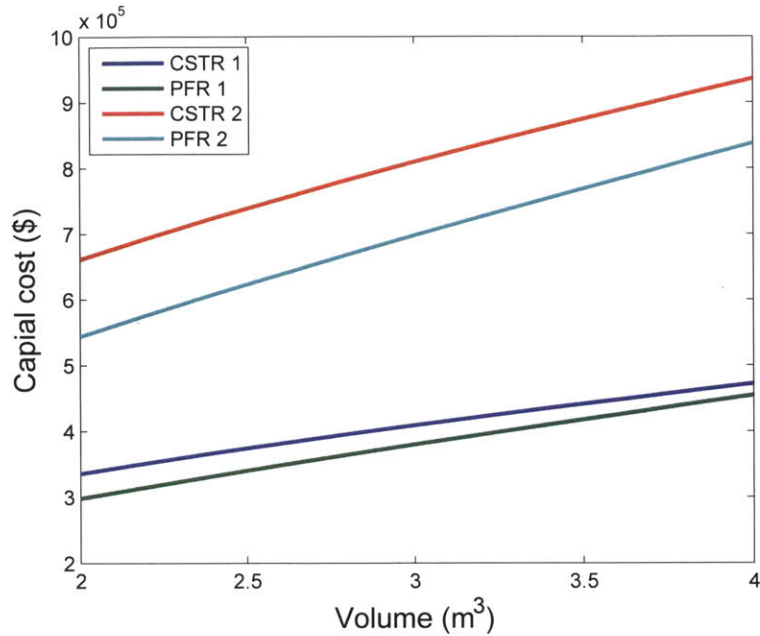
Figure 6-5: Payback period for different volumes and different process choices. 1 and 2 represent the manufacture of C and D respectively

are selected so that the best outcomes can be identified.

Thus, we see the importance of selecting the correct metrics when selecting among projects.

6.3   Model analysis with uncertainty

While the preceeding section makes use of deterministic approach to analyze the projects, we know that for real processes, there exists uncertainties in the parameter values. We introduce uncertainty into the models by selecting the parameters that we believed would most affect the output sensitivity, namely the reaction rate and the selling price of the product. We use lognormal distributions for the parameters of interest and illustrate their distributions in Figure 6-6
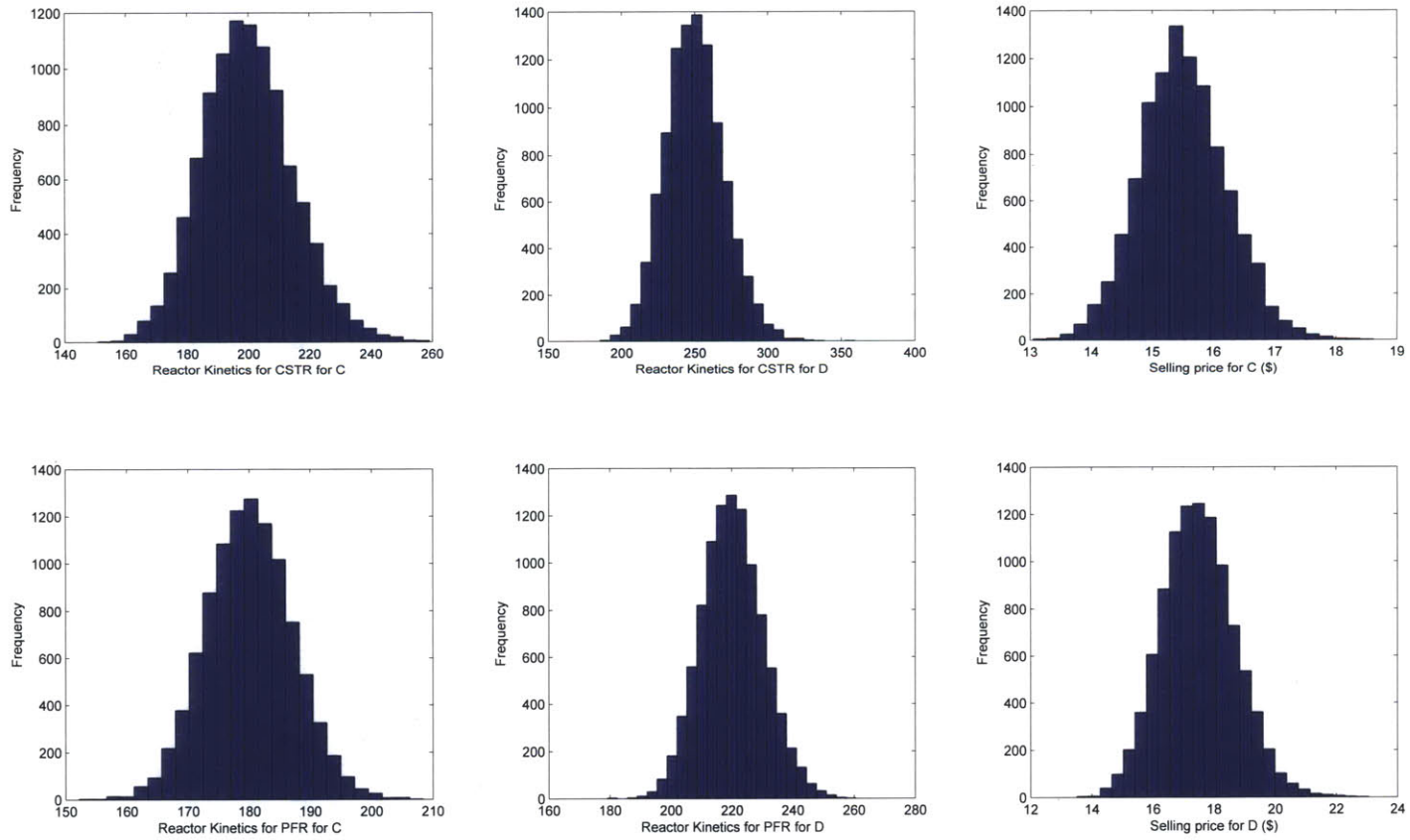
Figure 6-6: Distributions of the uncertain parameters for the CSTR and PFR and the selling prices of C and D
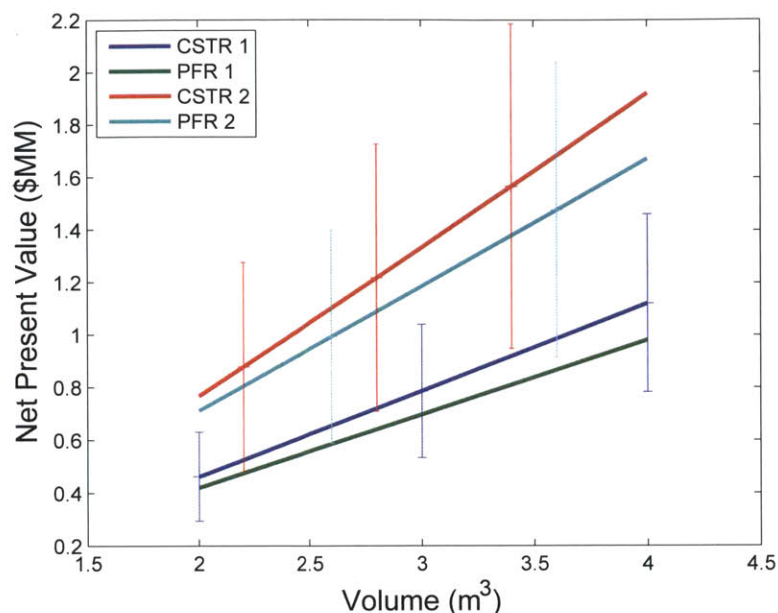
Figure 6-7: Net Present Values for different volumes and different process choices, with error bars to indicate the spread

When we include uncertainties in the distributions of the parameters, the decision results from the deterministic section are no longer as crisp as illustrated in Figure 6-7.

The figure shows the plots of the expected values of the decisions with a the standard deviations around a few points emphasized. The width of the bars is equivalent to two standard deviations - one above the particular point and one beneath. From the figure, we see that the presence of uncertainty now makes the decisions that were unreasonable now potentially viable. For example the production of D using a CSTR with a capacity of $2.2m^3$. This, deterministically, was shown earlier to be superior to all other options but we see that given the spread around its expected value, it is possible that under some circumstances, even the CSTR producing C (the blue curve) may generate a larger net present value.

We now turn to our developed framework to help with the decision.

6.3.1 Decisions from different uncertain objectives

Figure 6-8 below summarizes the different decisions made by the different metrics. Different shapes (also coloured differently) help to identify the different reactor-product configurations and the particular volume choice is shown in the appropriate shape above the volume scale.

There are two groups of outcomes - there are the single decision outcomes and the multiple decision outcomes. The single decision outcomes are the expected value, risk-area-ratio and the worst-case decisions. The expected value decision is to use a CSTR with capacity of $4m^3$ producing D and the other two objectives recommend

119

the production of C using a CSTR of volume $4m^3$.

For the multiple decision outcomes, there is a range of decisions that are recommended. Thus, we see that the First Order Stochastic Dominance set contains 13 decisions in total, comprising 8 possible volume choices for production of D by PFR, 3 choices of producing D by CSTR and one each of producing C by CSTR and PFR. The SSD solution contains 10 decisions in total - 6 for the production of D by PFR, 3 for its production using the CSTR and 1 for the production of C by CSTR. The diagram for the CSTR can be similarly interpreted.
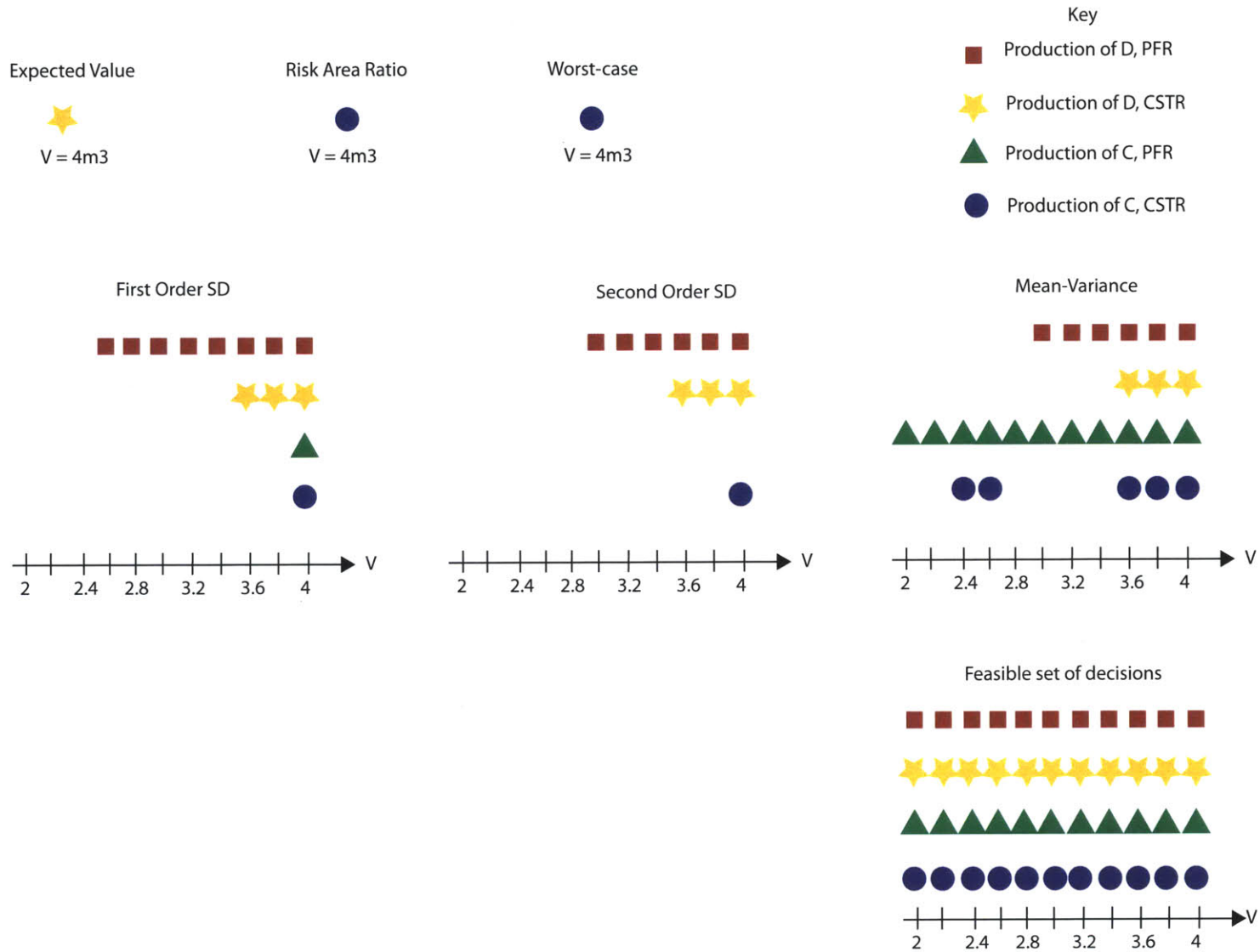
Figure 6-8: Decision results for different objectives for the Reactor-Separator design problem
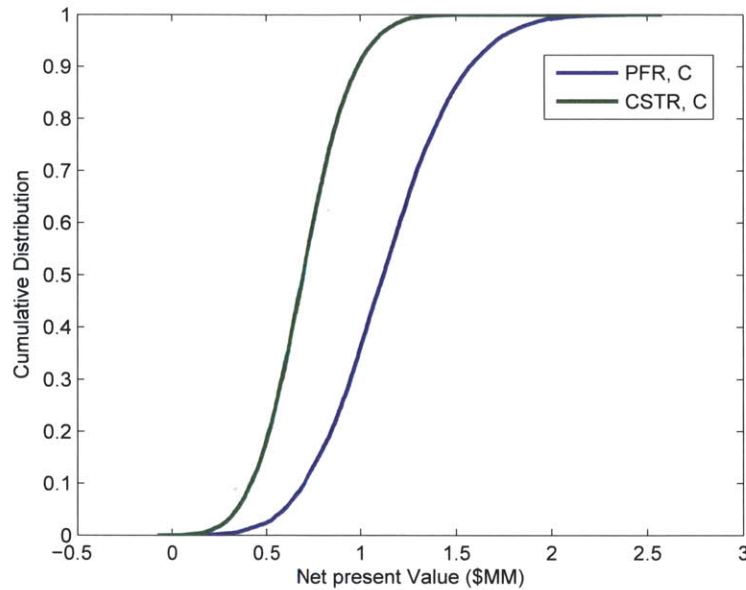
Figure 6-9: Comparison of a Mean-Variance Decision and a Stochastic Dominance Decision. The SD decision dominates the MV decision as its cumulative distribution is entirely to the right of the MV decision

From the figure, we see that FSD does a fair job of screening the initial number of designs - providing a reduction of about 67% in the size of the feasible set from 44 possible decisions to 13 decisions - without any need for the knowledge of the decision maker's risk preferences. Second order SD further reduces the FSD efficient set from 13 decisions to 10 decisions. This was done by assuming the decision makers were risk averse and demonstrates the nestedness of the SSD efficient set.

Comparing the stochastic dominance sets and the mean-variance set, we see that there are more than double the number of decisions present in the FSD set and almost three times the number present in the SSD set. Notable in the mean-variance set is the presence of the full list of decisions for the production of C by PFR. Many of these decisions are superseded by decisions in the Stochastic Dominance set and as a result can be misleading to decision-makers.

Take for example the decision to build a PFR with a volume of $3m^3$ to produce C. This decision is absent from the Stochastic Dominance set but is present in the MV set. The reason is because there is a design that supersedes this decision present in the feasible set and no decision-maker will choose this. We expand on this in Figure 6-9.

In the figure, we see that the decision to produce C with a PFR has a cumulative distribution given by the green curve. There is however another decision that dominates it - which is the decision to produce C with a CSTR and a volume of $4m^3$. This is even though both lie on the MV efficient set. As such, it is possible for the simple comparison of these two decisions on an MV basis to mislead decision makers. (it will be good to plot the probability distributions of these two decisions as well).

122

This decision is the basic flaw of the MV approach as discussed earlier. Because the variance (or standard deviation more accurately) penalizes both upside and downside variance, a low variance (like the PFR decision) penalizes upside as well. But we see that the larger variance of the CSTR decision contributes more to the upside. Even though there is a downside, it's not large enough to pull it into the lower NPV numbers categorizing the PFR decision.

6.4 Convergence Issues

One challenge in applying the Stochastic Dominance framework to this problem is the convergence of the size of the efficient set as the number of Monte-carlo samples of the uncertain parameters increase. Figure 6-10 demonstrates the variation in the size of the efficient set for different Monte-carlo runs.

In order to obtain the plots, we did the following; first we selected a number of sample sizes over which we were going to run the analysis[1] on the decisions. These ranged from 100 samples to 100,000 samples, three orders of magnitude. For each selected sample size, for example a sample size of 1000, we repeated the analysis a specific number of times (100 times in our case) and in each case, recorded the number of decisions in the efficient set and the decisions in each efficient set. These results are what the figure displays.

Looking closely at the figure, we observe a number of things. For the mean variance, there is a lot more fluctuation in the size of the efficient set with smaller samples [2] than for the larger one[3]. This is to be expected as will be demonstrated from standard Monte-carlo convergence analysis shortly.

For First order and Second order Stochastic Dominance, the story is a little different - at smaller sample sizes, there is a tendency towards a smaller size of the efficient set (close to 3) while at larger efficient set sizes there is a tendency toward larger efficient sets (10 or more). This is more clearly illustrated in Figure 6-11

As can be seen, while the size (and variation) of the mean-variance efficient set starts out large and eventually settles into a constant number with more samples (25 options), the number of decisions in the FSD and SSD efficient sets starts out much smaller and then grows, with a large fluctuation as well.

The main reason for this is the convergence of the tails of the distributions compared with the convergence of the central portions of the distributions - like the mean and the standard deviation. Figure 6-12 shows the convergence of the mean and the variance of the distributions of some decisions. As can be seen, most of them converge after 10,000 samples which explains the stability of the efficient set from 20,000 samples upwards

The convergence of the tails is quite different as is illustrated in the Figure 6-13 where we plot a number of percentiles for the same decisions as the mean and standard deviation.

---

[1]Note that 'analysis' here refers to determining the efficient sets of the different objectives
[2]the range is from 21 to 34 decisions for a sample size of 100
[3]for 100,000 samples, there are essentially consistently 25 decisions in the efficient set
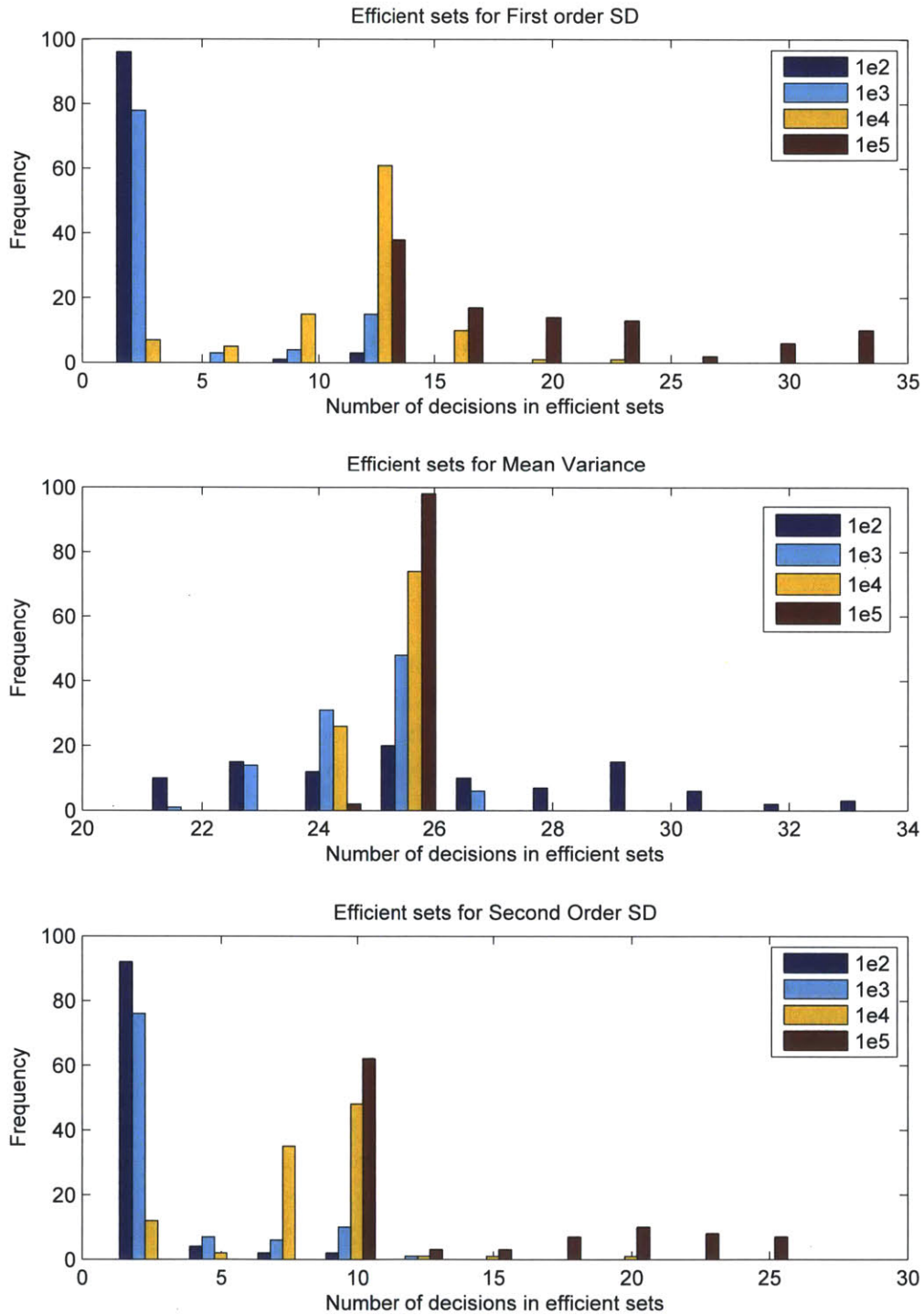
Figure 6-10: Histogram of the number of decisions in the efficient sets for different
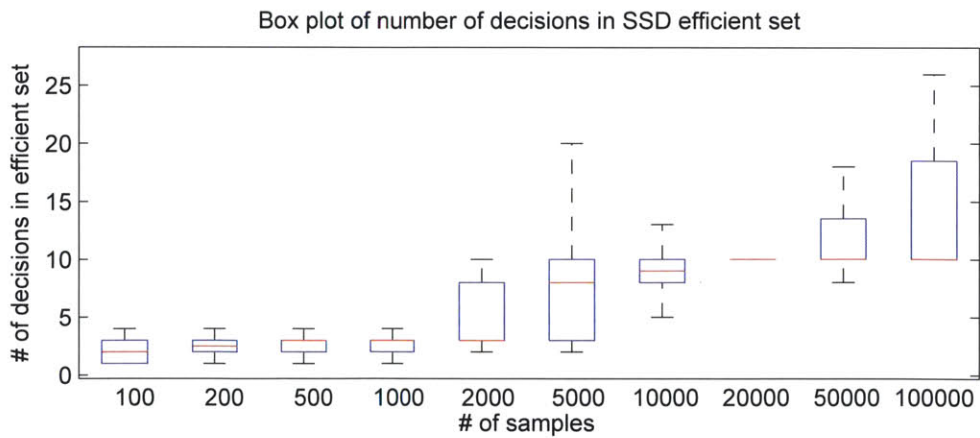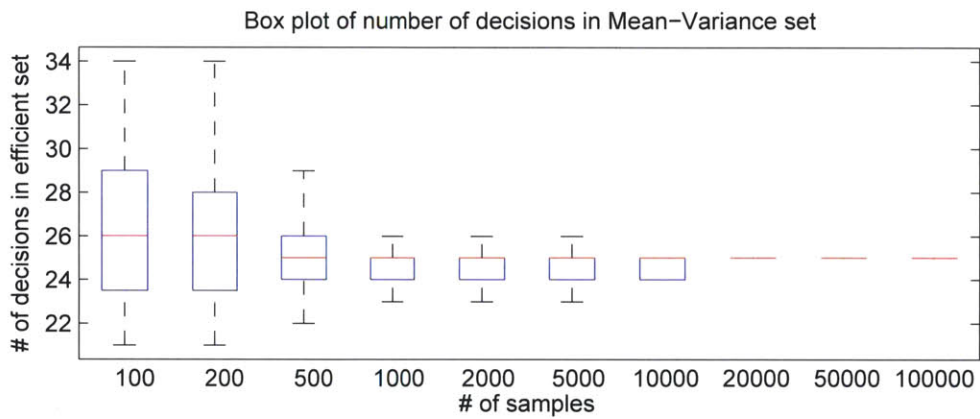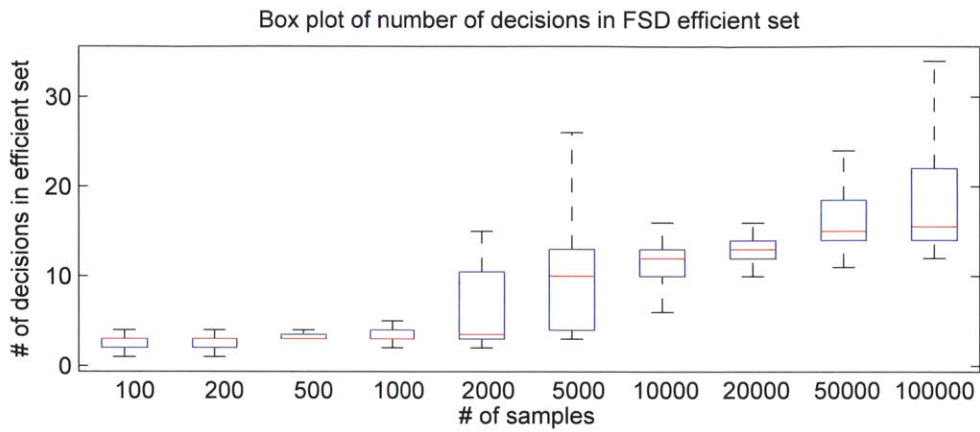
Figure 6-11: Box plots of the different efficient sets
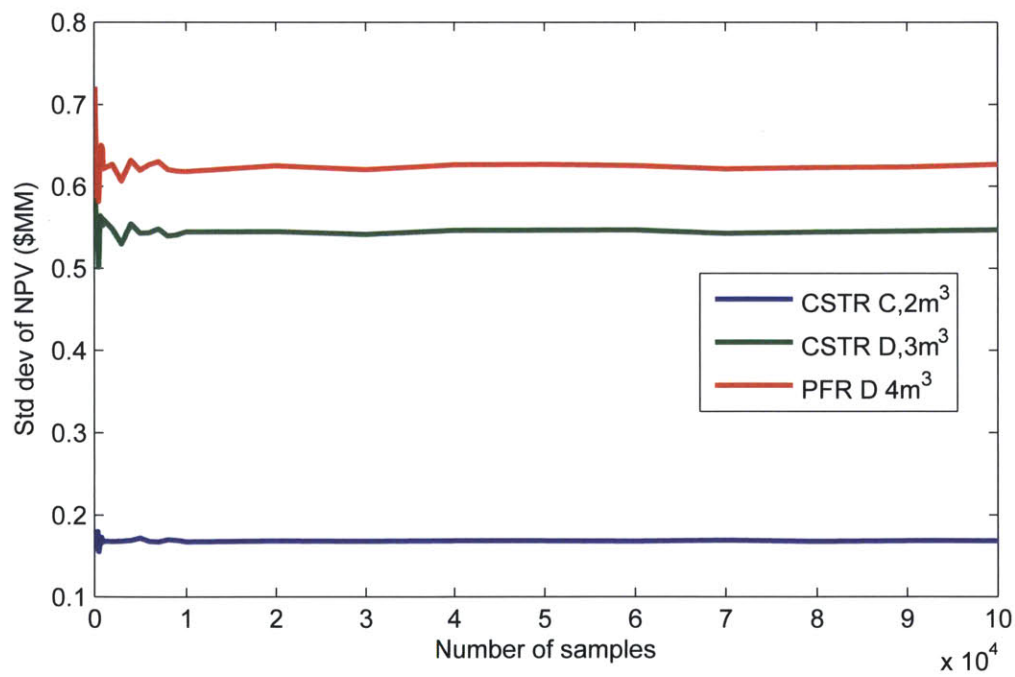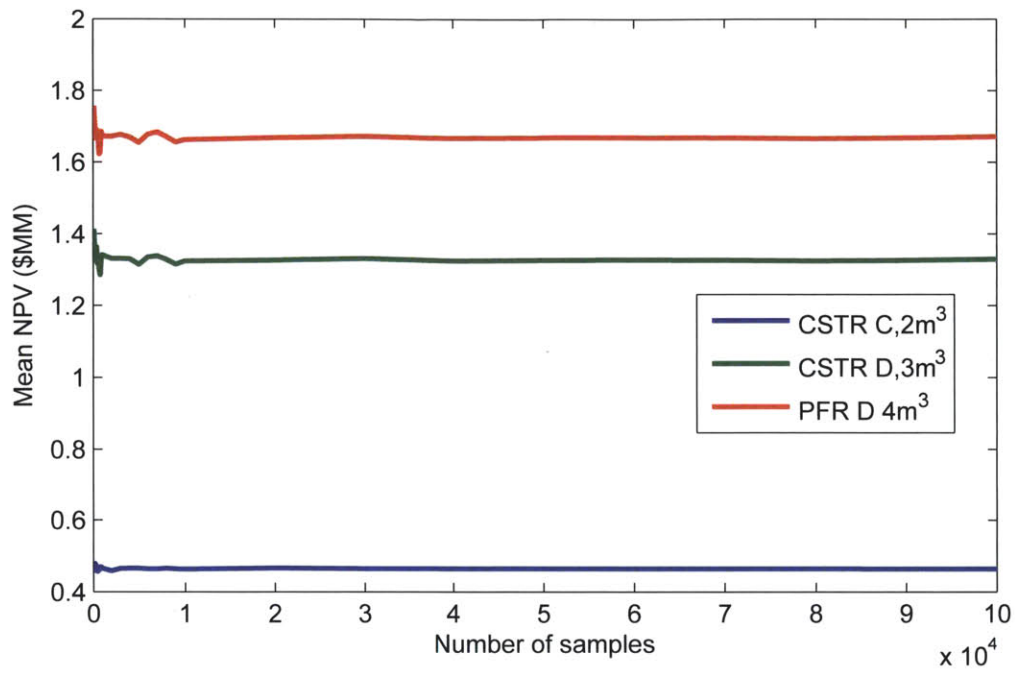
Figure 6-12: Convergence of Mean and Standard deviation for a few decision options
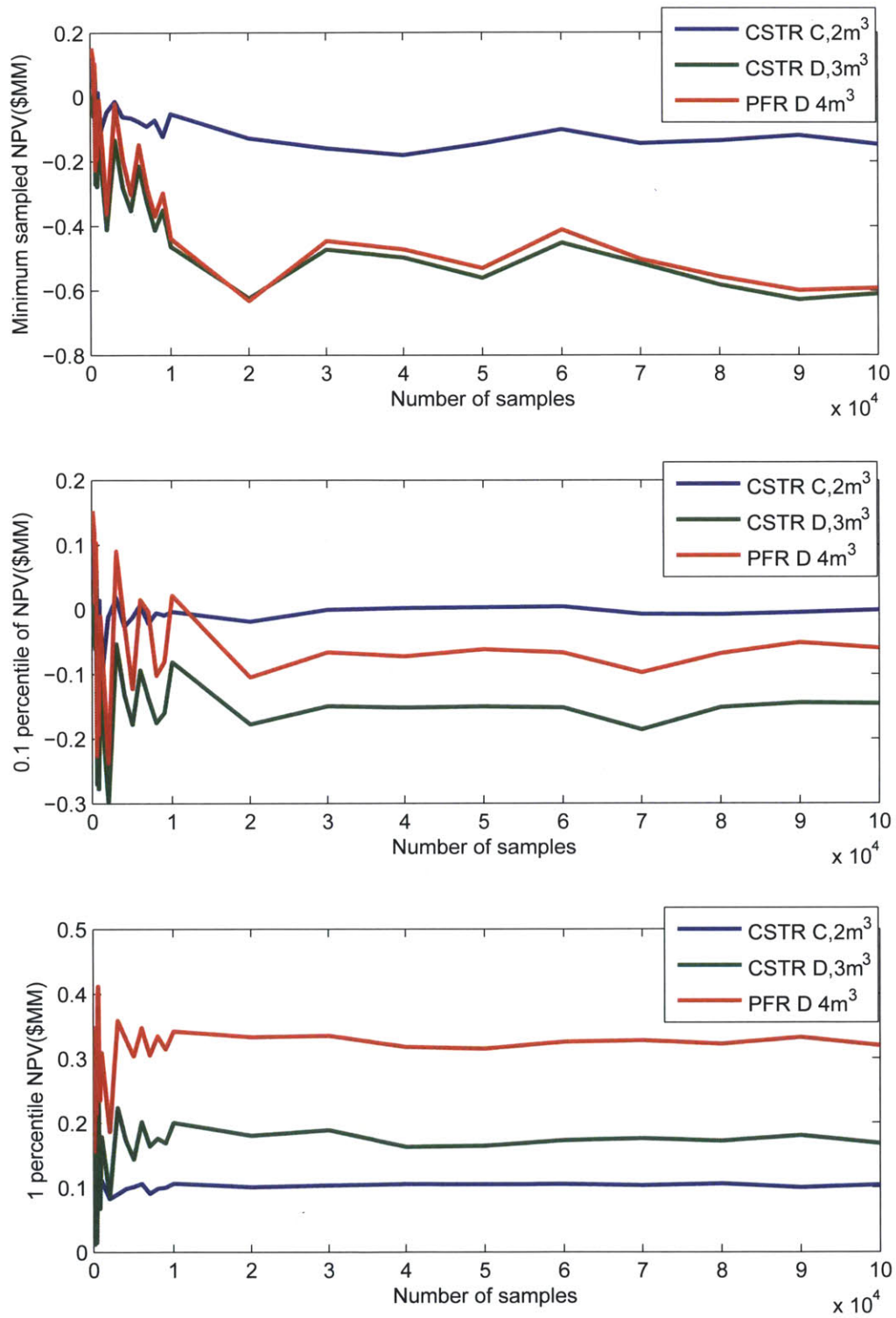
Figure 6-13: Convergence of different percentiles for a few decisio options

Reactor Type	Product	Volume (m^3)
CSTR	D	4.0
PFR	D	4.0
PFR	D	3.8
PFR	D	3,6
PFR	D	3.4
PFR	D	3.2
PFR	D	3.0
CSTR	C	4.0

Table 6.1: Decisions in the SSD efficient set using the NPV metric

In the figure, the minimum sampled NPV is a proxy for the zeroth percentile. The plots as such show the convergence of the zeroth, 0.1th and the 1st percentile. There is a visible increase in the convergence across all three but it is markedly different from that observed with the mean and the variance. This lack of convergence in the tails of the distribution partly accounts for the variation in the size of the efficient sets. First order Stochastic Dominance involves comparing values at each percentile value while Second order Stochastic Dominance involves comparing the integrals of values below each percentile value. With fluctuations in the sample values, it is possible for the percentile for a decision, A, to be larger than that of B, even though the 'true' value (as will be obtained in the limit of infinite samples - or sufficiently large samples) is lower.

6.5 From efficient sets to decisions: Pruning the efficient set

Even in the absence of fluctuations in the Stochastic Dominance efficient sets, it isn't common that the set will contain a single decision that can be implemented directly. In such a case, it becomes necessary to further prune the efficient set to a more acceptable number. As highlighted in the previous chapter, we demonstrate the use of other metrics and the use of the Almost Stochastic Dominance framework in achieving this goal.

The use of other metrics - Once second order Stochastic Dominance has been used to reduce the efficient set to a manageable size, it is possible to use other metrics to further pare down the set to arrive at a design decision. For example, as we discussed earlier, the Net Present Value is but one metric that can be used to measure profitability. We can run SSD analysis on the efficient set using another metric like the internal rate of return to see what reduction there is.

The SSD efficient set in this case contains 8 decisions (from an initial set of 44) and this was obtained using 10,000 samples of the uncertain variables. These decisions are represented in the Table 6.1.

Reactor Type	Product	Volume (m^3)
CSTR	D	4.0
PFR	D	4.0
CSTR	C	4.0

Table 6.2: Decisions in the FSD efficient set using the IRR metric on the NPV, SSD efficient set

Reactor Type	Product	Volume (m^3)
CSTR	C	4.0

Table 6.3: Decisions in the SSD efficient set using the IRR metric on the NPV, SSD efficient set

We now analyze these decisions using Stochastic dominance on the uncertain internal rates of return. The decisions left in the FSD and SSD sets are presented in the tables below. With this metric, we were able to make a decision using SSD on the 8 decisions that were left. Using FSD gave us 3 decisions, with SSD finally reducing that decision to 1.

An interesting observation in this case is that the SSD set resulting from the application of the IRR metric on the initial set (all 44 decisions) is the same as applying it on the smaller efficient set that resulted from applying Stochastic Dominance using the NPV metric. This is not surprising as the two metrics are not entirely independent. This will not always be the case though - especially when the metrics are quite different. In those cases, it is possible that the order in which the different metrics are applied leads to different decisions. This is one of the dilemmas of lexicographic decision making discussed in [70] and the approach should be used with caution. When tradeoff coefficients between the different metrics are known, it is best to use the combined weighted average metric rather than applying the metrics in sequence.

Almost Stochastic Dominance The rigour of Stochastic Dominance means that even in cases where one distribution virtually dominates another, except in a region of very low probability, both decisions will be present in the efficient set. As an example, consider two decisions in the efficient set whose cumulative distributions are represented in the Figure 6-14

The distributions represent the choices of producing D with a CSTR (blue) and a PFR (green) both with a reactor volume of $4m^3$, and both of which are in the Second order Stochastic Doiminance efficient Set (using the NPV metric). However, it is initially unclear from Figure 6-14 why the PFR decision is in the efficient set given that the CSTR's distribution seems to lie entirely to the right of the PFR decision. The complete picture is seen more clearly near the low probabilities as we demonstrate by the zoomed in image in Figure 6-15.

Here we see that near the tails, the (empirical) distributions begin to cross, leading to no clear dominance in the two. decisions. However, if the scale of the y-axis is examined, we see that this crossing over occurs around a cumulative probability less
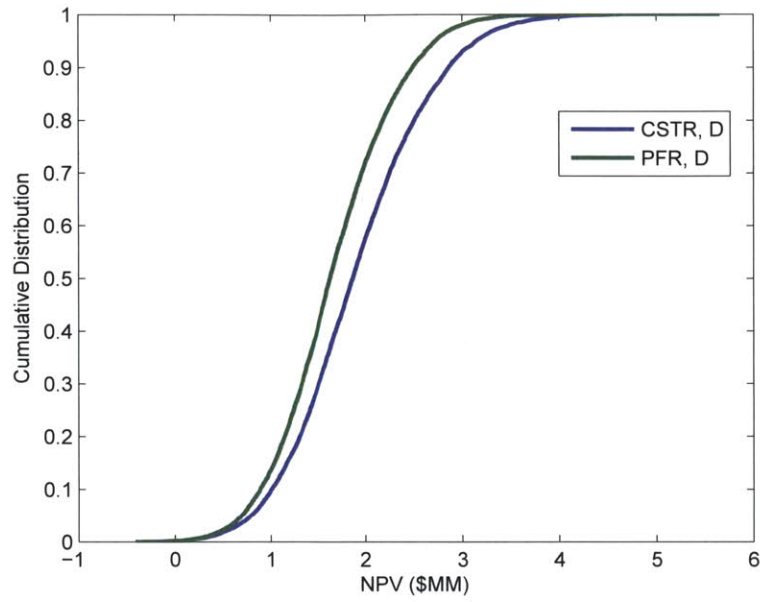
Figure 6-14: Cumulative distributions of the different decisions to produce D with CSTR and a PFR
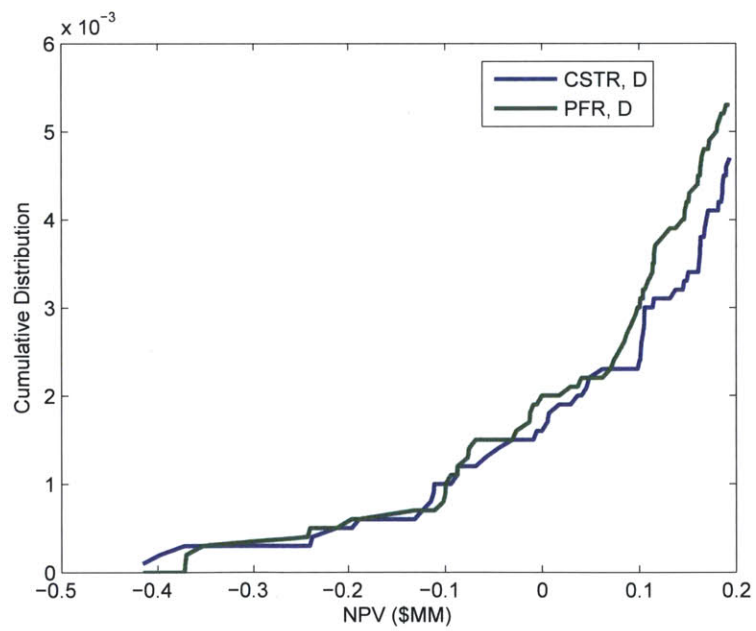
Figure 6-15: Tail cumulative distributions of the different decisions to produce D with CSTR and a PFR
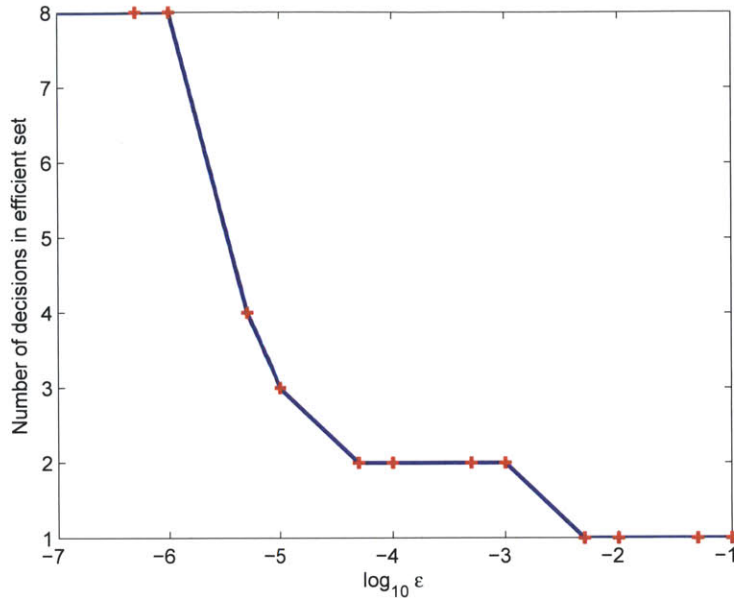
Figure 6-16: Decrease in the size of the efficient set with the ASD parameter, ϵ

than 0.01, the first percentile of the distribution. This region represents a set of NPVs that have a probability of less than 1% of occuring. While there are utility functions for which the PFR decisions will be preferred to the CSTR decision, these utility functions will lead to the acceptance of gambles that most people will find untenable such as the example given in [62].

As explained in the previous chapter, Almost Stochastic Dominance, helps to eliminate some of these decisions, however it depends on the value of a particular parameter, ϵ. The lower bound for this parameter is zero, at which point it is equivalent to the regular Stochastic Dominance. Ruszczynski and Lizyayev [66] give a physical meaning for the parameter when used in the ϵ-ASSD sense as 'the smallest value of the mean return of a random variable that needs to be added to a random variable, X in order for it to dominate another, Y'. We screen the decisions in the SSD efficient set using different values of ϵ and observing the reduction of the efficient set. These are presented in Figure 6-16.

The figure is a semi-log plot where the logarithm of ϵ (in base 10) is plotted against the number of decisions in the efficient set. The number of decisions begins with the total number of decisions in the SSD efficient set, 8, and then steadily drops until $\epsilon = 5 \times 10^{-3}$ when only a single decision is left in the efficient set. This decision is the decision to produce D with a CSTR of volume $4m^3$. The very small value of epsilon shows that most decision makers (5×10^{-3}) shows that most risk-averse decision makers will be easily satisfied with this decision. In this case however, and in almost every case where ASSD is used to reduce the number of decisions to a single outcome, the final decision left in the efficient set will always be the decision with the largest expected value, thus corresponding with the decision that results from expected value maximization.

Note that this decision is different from the decision that resulted earlier when we used another metric to pare down the SSD set. In general, this will be the case and as such, it is recommended that the reduction of efficient sets should be done in tandem with the decision maker as the preferred approach to reducing the size of the efficient set will inform the analyst on the decision maker's particular preferences, resulting in the recommendation of a decision that will be easily accepted.

Use of hierarchical models It is also possible to further refine the efficient set by a refinement of the mathematical models used to run the analysis. Kasas et al. [53] illustrate with deterministic metrics how using models of increasing accuracy can help improve the trade-offs made in decision making, in effect reducing the size of the efficient set. We didn't investigate this approach here, but for completeness mention its use.

6.6 Conclusions

This chapter examined the design of a reactor-separator in the absence and presence of uncertainty. The key design variables were the choice of reactor, volume and the particular product to be made as this affected the reaction kinetics. The uncertain parameters that were analyzed were the reaction rate and the product selling price as a preliminary sensitivity analysis had shown both parameters to be the most important.

The best design in the deterministic case turned out to be a CSTR geared towards the production of D (implying second order kinetics) with a volume of $4m^3$. This design also turned out to be the best expected value solution as well in the presence of uncertainty. There were 10 decisions in the SSD set although the number was found to vary as we changed sample size since there was a lot of the crossing over in the tails of the cumulative distributions. This issue was identified using Almost Second order Stochastic Dominance to analyze the efficient set as the size of the set dropped very quickly with minor changes in ϵ, with the expected value decision remaining as the sole decision while ϵ was still quite small (order of 10^{-3}).

The use of an alternative metric was also able to reduce the efficient set down from 10 decisions to three, with the expected value decision under the NPV scenario still remaining in the set. Thus, it seems that in this particular case, the use of the expected value solution will be the best decision to pick.

This study illuminated some of the issues that can arise with the use of Stochastic Dominance for project screening - particularly the fluctuations that can arise in the tails of distributions and how they can affect the size of the efficient set - an issue that isn't solved by the use of more samples. The ASSD approach thus comes in handy in eliminating decisions present in the efficient set because of tail fluctuations and should be used whenever such situations arise.

Chapter 7

Case Study 2: Growing energy crops

This second case-study demonstrates the use of Stochastic Dominance on the selection of the crop mix and land area for a farmer interested in growing renewable energy crops for biofuel production. We first develop the model and solve it deterministically before including uncertainty. This problem is the first part of the larger problem of designing a plant that converts biomass to liquid fuels and potentially other products. As analyzed in the first chapter, there are three main decisions to be made.

1. The choice of raw materials/feed to be used

2. The processing pathway

3. The product mix

In general there will be some interaction between the three design stages as it is easy to see how the choice of any one of the stages will affect the rest. We address some of this in the next case-study where we analyze the system-wide design of the biomass to liquids plant. In this chapter however, we will focus on the initial problem of selecting the biomass feedstock for the plant assuming that, to first order, the decision can be separated from the rest of the plant design. This is not a poor initial assumption because the generation of products from biomass often requires that biomass be first converted to syngas via the gasification process. Gasification presents a natural decoupling point for the feed selection problem from the design - since the rest of the plant design can also be analyzed on the basis of syngas generated from the plant. The interaction linking the feed to the rest of the plant via the syngas lies in the amount of clean-up that has to be done for the syngas based on the type of feed. But syngas clean-up is a small fraction of the overall cost of the plant that, again, to first order, can be neglected and considered in a later, more detailed analysis.

7.1 Problem description and model development

First generation biofuels were largely made from corn and sugarcane but a number of reasons have led to the need for alternative sources for the manufacture of biofuels.

The first is the food-fuel competition which led to an increase in prices for food, especially corn, as supply became scarce due to the alternative demand source. The second is the debate over the fossil fuel neutrality of corn-ethanol with respect to fossil fuel as some argue that it is consumes more fossil fuel in its creation than it replaces while other studies show it to be mildly positive, issues which are explored in mroe detail by Johnson [50].

These two reasons have led to the search for new crops and carbon sources from which biofuels can be made that have a greater energy ratio and don't compete with food supply. Of the many alternatives that have been proposed, four of them have been prominent - switchgrass, miscanthus, willow (an example of a short rotation woody crop) and corn stover - and we will focus on them in our analysis.

The decision problem is that of a farmer deciding to be a supplier of one (or more) of these energy crops. His decision is constrained by his initial overall capital (how much he has or can raise), and, given the set up costs (all done on a per unit area of land cultivated), this affects the size of land he can purchase and subsequently will determine the net present value of his returns.

In order to determine which crop he should grow, we first build a simple structural model that relates the main factors and parameters to the outcome we care about - the net present value. Our model couples all costs incurred in the growing process but uniquely separates out tranport costs which have been found to be a very significant factor in farm/crop growth [17].

First, we assume that the revenue from a crop can be given by the relationship

$$Z = p \times Y \times A \tag{7.1}$$

where Z is the return from the crop, p is the selling price of the crop, Y is the crop yield per unit farm area and A is the total farm area harvested. The figure below shows the relationships in detail

For the elemental area (in blue) we have

$$dA = r dr d\phi \tag{7.2}$$

and as a result we get

$$dZ = pY r dr d\phi \tag{7.3}$$

which upon integrating yields

$$Z = \int_0^{2\pi} \int_0^R dZ$$
$$= Yp \int_0^{2\pi} \int_0^R r dr d\phi$$
$$Z = pY \pi R^2 \tag{7.4}$$

Figure 7-1: Schematic digaram for farm model

where R is the radius of the farmland. The total cost of growing crops is similar to the above and is given by:

$$
\begin{aligned}
Growthcost &= cost \times mass \\
&= c \times Y \times \pi R^2
\end{aligned}
\tag{7.5}
$$

Transportations costs on the other hand are given by

$$
T = c_t \times Y \times A \times D
\tag{7.6}
$$

c_t is the transportation cost per unit distance per unit mass of crop and D is the total distance travelled. For the elemental area shown, we have that

$$
dT = c_t \times Y \times r dr d\phi \times 2r
\tag{7.7}
$$

Which upon integrating yields

$$
\begin{aligned}
T &= \int_0^{2\pi} \int_0^R dT \\
&= Y c_t \int_0^{2\pi} \int_0^R 2r^2 dr d\phi \\
&= \frac{4\pi c_t Y R^3}{3}
\end{aligned}
\tag{7.8}
$$

Capital cost, we model as a function of the area of the farm (essentially land costs

and preparations) and this is given by

$$Capitalcost = costperarea \times totalarea$$
$$= D \times \pi R^2 \tag{7.9}$$

where D is the establishment cost of the crop.

Furthermore, we assume that there is a total budget, B, available for the initial capital investment. Given the budget, there is a constraint on the amount of farmland that can be purchased i.e.

$$D\pi R^2 \leq B$$
$$0 \leq R \leq \left(\frac{B}{D\pi}\right)^{\frac{1}{2}} \tag{7.10}$$

This budget constraint translates to a maximum radius, R_{max}, which can be calculated for each crop once the initial investment cost, D_i, is known.

The investment/establishment cost is a one time cost (a capital expense) whereas revenue and crop production cost terms (as well as the transportation costs) are recurring annual items. In order to combine the equations, we either need to annualize the capital cost of the farm or sum the discounted revenues and costs of the farm over the lifetime of the crop. The crops all have a varying lifetimes so this as to be accounted for as well.

Miscanthus and willow (our chosen short rotation woody crop) both have a lifetime of about 20 years while switchgrass has a lifetime of about 10years. We gave the corn a lifetime of 10 years as well. We thus levelled the returns of all crops to the lifetime of 20 years - but included a factor that accounted for reinvestment costs for corn stover and switchgrass after year 10. We assume a constant average yield over the lifetime of the farm for each crop which allowed us to use a single factor to account for the net present value of the crop returns. We discount at a risk free return of 5% per year.

Thus, with a lifetime of 20 years, the discount factor is

$$f_d = \sum_{k=1}^{20} \frac{1}{(1+r)^k} \tag{7.11}$$

For the capital cost correction, we assume that whatever the amount was invested in the beginning for the crops with a lifetime of 10 years will be spent again to account sustain the current farm size. Thus, we added a capital cost discounted over 10 years to account for this. Thus switchgrass and corn stover have capital cost multipliers given by

$$f_{inv} = 1 + \left(\frac{1}{1+r}\right)^{10} \tag{7.12}$$

Since the lifetime of miscanthus and willow are both 20 years, their investment factor is 1 (since there would be no reinvestment done). With these, we can now

write the mathematical model of the farmer's problem as

$$\max_{R} \left\{ f_d((p-c)Y\pi R^2 - \frac{4\pi c_t Y R^3}{3}) - D\pi R^2 (f_{inv}) \right\}$$

$$st. \quad 0 \le R \le \left(\frac{B}{D\pi}\right)^{\frac{1}{2}}$$

(7.13)

Equation (7.13) helps the farmer determine the optimal size of the farm for a particular crop. Since R is the only variable, we can use the techniques of calculus to obtain the optimum decision and the corresponding net present value. We will temporarily assume the solution lies at an interior point and thus neglect the complications that arise with including boundary restrictions.

At the optimum we have

$$\frac{dNPV}{dR} = 0 = 2f_d(p-c)Y\pi R - 4 * f_d c_t Y\pi R^2 - 2D\pi R f_{inv} \qquad (7.14)$$

$$0 = 2(p-c)f_d Y - 4f_d c_t Y R - 2D f_v \qquad (7.15)$$

which gives the optimal deterministic radius as

$$R = \frac{Y(p-c)f_d - Df_v}{2f_d c_t Y} \qquad (7.16)$$

Since R cannot be negative, for there to be any investment at all in the farm, we must have, from equation (7.16) that

$$\frac{Yf_d(p-c)}{f_v D} \ge 1 \qquad (7.17)$$

which, in economic terms, translates to the fact that the total discounted returns from the farm must exceed the initial capital investement.

$$\frac{d^2 NPV}{dR^2} = -4c_t f_d Y \le 0 \qquad (7.18)$$

as required by the necessary conditions for a maximum. For a farmer that faces a number of potential crops to grow from which he has to chose one (and only one), the decision problem can be written as

$$\max_{R,i} \left\{ (p_i - c_i)Y_i\pi R^2 - \frac{4\pi c_t Y_i R^3}{3} + Br - D\pi R^2(1+r) \right\}$$

$$st. \quad 0 \le R \le \left(\frac{B}{D\pi}\right)^{\frac{1}{2}}$$

(7.19)

where i is the index of each crop and n the total number of crop options the farmer

has.

7.2 Crop selection under certainty

We solve this problem using data given in Table 7.1 for the four energy crops introduced at the beginning of the section- switchgrass, miscanthus, willow and corn stover. We show the plot for miscanthus in Figure 7-2 and that for all four together in Figure 7-3. For all the crops, we used a transportation cost of about \$0.25/ton/km. The establishment costs for the different crops were obtained from a survey of literature.

Prices are, where it was possible to determine, how much can currently be obtained from the sale of the crop while costs were estimated to be the total non-establishment and transportation costs incurred in growing the crop.

Corn stover establishment costs were determined from the establishment costs of corn and dividing by two since the corn crop yields about 50% grain and 50% stover. Since corn averages about 3-5 tons per acre(about 7.5-15 tons per hectare) , a stover harvest rate of about 30% of the standing mass is used since the stover usually acts as a replenishment for the soil after harvest. Selling price was estimated as an average of the ranges given in and the cost of growing estimated the same way.

Of all the crops though, establishment costs for Miscanthus varied the most - ranging from \$1000 to about \$2500. We used the average \$1750 to run the deterministic analysis.

Crop	Price (\$/ton)	Cost (\$/ton)	Yield (tons/ha)	Establishment (\$/ha)
Switchgrass	65	60	15	450
Miscanthus	65	55	25	1750
Willow	65	55	14	2000
Corn Stover	50	67	5	350

Table 7.1: Data for the farmer problem

Parameter	Value	Unit
Budget, B	75×10^6	\$
Risk free rate, r	5	%
Transportation cost, c_t	0.25	\$/ton.km

Table 7.2: General farm data

As we start out from smaller radii and move out to large ones, we see a rise in profitability because the marginal revenue obtained from adding each additional bit of land more than compensates for the increasing marginal transportation cost until the optimum distance where both are equal. Beyond this distance, the extra cost incurred on transporting the crops exceeds the marginal revenue and this gives rise to the decreasing profitability which continues to fall until the maximum distance (determined by the budget constraint is reached.
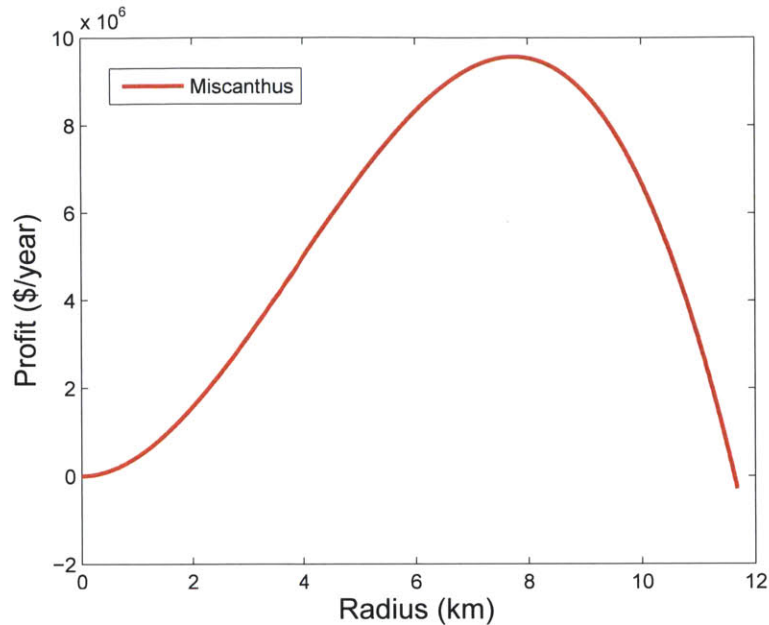
Figure 7-2: Farm profit as a function of radius for Miscanthus

As can be seen from Figure 7-3 the peaks are all different as well as the range of the farm radii available to farmer. The table below gives the results for all the crops together with their respective annualized profits. from the table, it can be seen that miscanthus provides the largest profitability and the most amount of biomass generated for the plant - almost 500,000 tons per year. This is followed by corn stover (in terms of profitability) and then by switchgrass and finally by willow. Given the data that we currently have, there is no radius for which the willow crop is profitable and as such it will be dropped out of consideration.

Thus, for the deterministic problem, with the four choices of the crops and a net present value criterion for determining profitability, the farmer will choose to grow switchgrass and will do so on a farm with radius 7.8km and will expect to provide about 500,000 tons of biomass per annum to the farm.

Crop	Max Radius (km)	**Opt Radius** (km)	Profit (MM)	Biomass (ktons)
Switchgrass	23.03	**2.17**	20.56	748
Miscanthus	11.68	**7.76**	1.80	679
Willow	10.92	**0**	0	0
Corn Stover	26.12	**13.94**	16.05	440

Table 7.3: Optimum farm radii and annual profit for the different crops
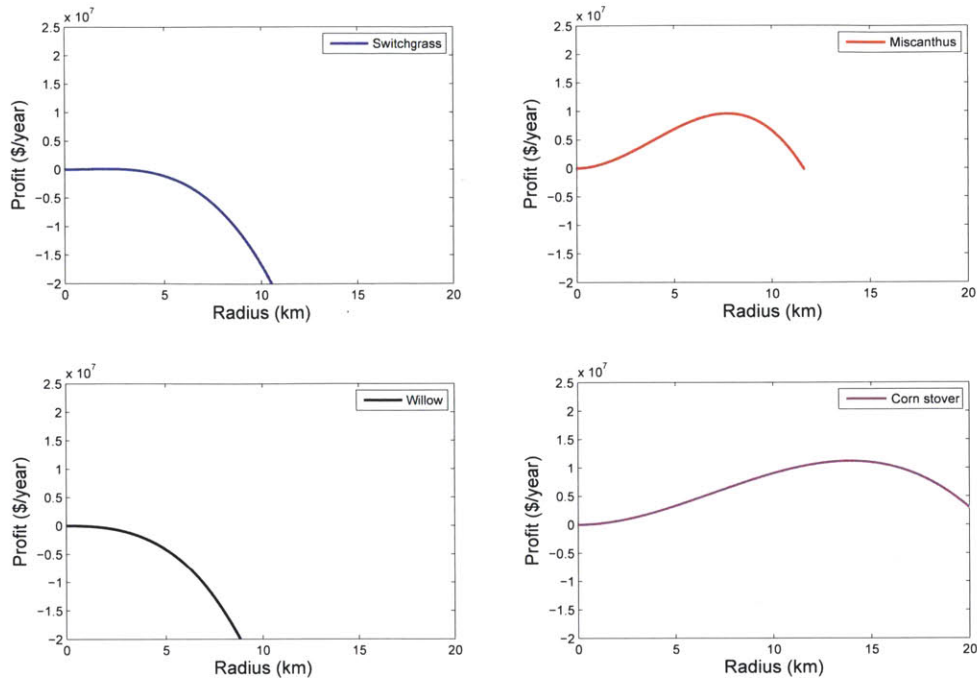
Figure 7-3: Profit as a function of radius for all crops

7.3 Crop selection under uncertainty

The analysis presented above is for the case where all the parameters defining the problem are known with certainty - an idealized scenario. In reality, these parameters are known to vary and the real farmer will have to make decisions regarding the crop to grow in the presence of uncertainty. The same base structural model is used for the uncertain case - but the parameters are now allowed to vary. Most of the variation was chosen to be uniform as the literature for a lot of these parameters is quite uncertain.

Information about the the parameters vary quite a bit in literature given that the growth of many of these crops is still in the research phase and the locations of farms are spread out geographically - leading to soil, weather and economic variations. To simplify the analysis, we chose five of the variables to be uncertain - the growth cost and selling prices of the different crops, the transportation costs, the yields and the establishment costs. All the distributions are constructed from sources in literature. We also assumed independence of the different uncertain variables. Furthermore, we assumed that the parameters of the problem don't change with time i.e. the distributions are static.

For all the variables except the yields, we assumed simple uniform distributions with the spread given in Tables 7.4, 7.5, 7.6 and 7.7. For the yields, we used uniform distributions for corn-stover and willow. A triangular distribution was constructed for switchgrass yields and a modified uniform distribution was used for the yields of miscanthus. The histograms from 10,000 samples of the miscanthus and switchgrass
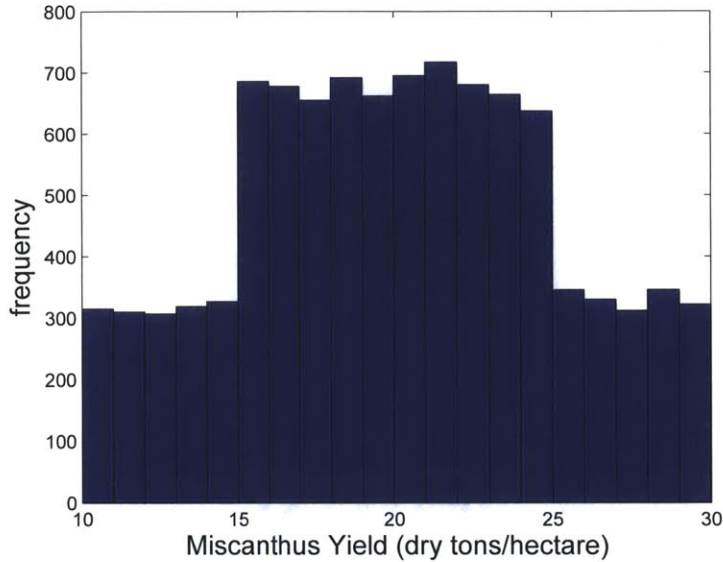
Figure 7-4: The histogram of yields for miscanthus

Parameter	Distribution	Data	Units
Selling Price	Uniform	[65, 75]	$/ton
Growth cost	Uniform	[55,65]	$/ton
Establishment cost	Uniform	[350, 550]	$/ton
Transportation cost	Uniform	[0.1, 0.5]	$/ton.km
Yield	Triangular	[6, 15, 25]	ton/hectare

Table 7.4: Parameters for Switchgrass

yields are displayed in Figures 7-4 and 7-5.

Each decision consists of a crop to grow and the farm radius. Since the farm radius is a continuous variable, we discretize it first in order to use the Stochastic Dominance algorithm. Like we did with the deterministic case, we determine a maximum farm radius based on the budget and the capital cost per unit area of investment. The farm radii for each crop is then obtained by selecting 21 evenly spaced points between the maximum and minimum farm radii (both included). For each choice of a crop and a farm radius, the uncertain parameters are then propagated through the model to get the distribution of the net present value for the decision. These distributions are then compared using the various metrics presented in Chapter 2 and Chapter 5. We present the results of the decision analysis in the Table 7.8.

Because the First order Stochastic Dominance efficient set contains the decisions in the second order stochastic dominance set, we include in a separate table the decisions present in the FSD set that are not in the SSD set. In addition, we include the results from the use of single objectives in Table 7.10.

One thing we see is that the SSD approach does a decent job of cutting down the number of decisions in the feasible set. While First order Stochastic Dominance

141
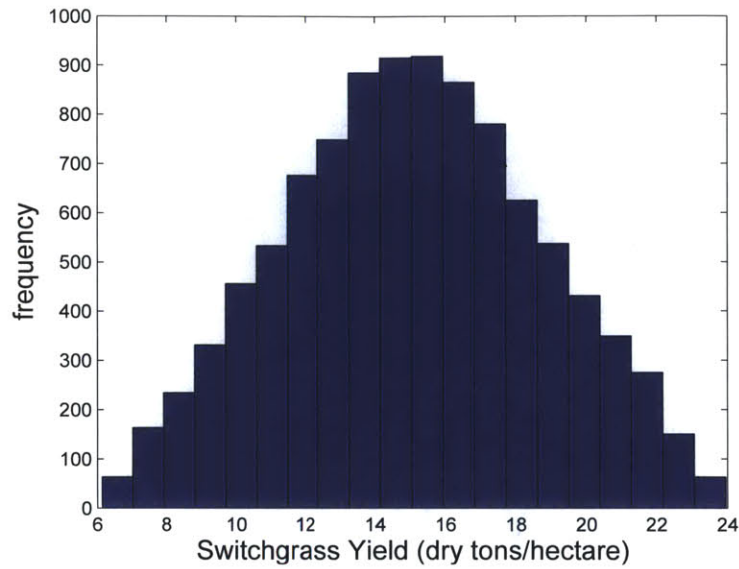
Figure 7-5: The histogram of yields for switchgrass

Parameter	Distribution	Data	Units
Selling Price	Uniform	[60, 70]	$/ton
Growth cost	Uniform	[50,60]	$/ton
Establishment cost	Uniform	[1000, 2500]	$/ton
Transportation cost	Uniform	[0.1, 0.5]	$/ton.km
Yield	Modified uniform	[6, 15, 25]	ton/hectare

Table 7.5: Parameters for Miscanthus

Parameter	Distribution	Data	Units
Selling Price	Uniform	[60, 70]	$/ton
Growth cost	Uniform	[50,60]	$/ton
Establishment cost	Uniform	[1500 2500]	$/ton
Transportation cost	Uniform	[0.1, 0.5]	$/ton.km
Yield	Triangular	[6, 15, 25]	ton/hectare

Table 7.6: Parameters for Willow

Parameter	Distribution	Data	Units
Selling Price	Uniform	[60, 70]	$/ton
Growth cost	Uniform	[43,60]	$/ton
Establishment cost	Uniform	[350, 550]	$/ton
Transportation cost	Uniform	[0.1, 0.6]	$/ton.km
Yield	Triangular	[6, 15, 25]	ton/hectare

Table 7.7: Parameters for Corn Stover

Crop	Radius (km)	Crop	Radius (km)
Miscanthus	0.58	Miscanthus	7.00
Miscanthus	1.17	Miscanthus	7.59
Miscanthus	1.75	Miscanthus	8.18
Miscanthus	2.34	Willow	0.55
Miscanthus	2.92	Corn Stover	1.31
Miscanthus	3.50	Corn Stover	2.61
Miscanthus	4.09	Corn Stover	3.92
Miscanthus	4.67	Corn Stover	5.22
Miscanthus	5.26	Corn Stover	6.53
Miscanthus	5.84	Corn Stover	7.84
Miscanthus	6.42	No Crops	-

Table 7.8: Second order Stochastic Dominance efficient set

Crop	Radius (km)	Crop	Radius (km)
Switchgrass	1.15	Corn Stover	9.14
Switchgrass	23.03	Corn Stover	11.75
Miscanthus	8.76	Corn Stover	20.89
Miscanthus	9.34	Corn Stover	22.20
Miscanthus	9.93	Corn Stover	23.51
Miscanthus	10.51	Corn Stover	24.81
Miscanthus	11.10	Corn Stover	26.12
Miscanthus	11.68	Corn Stover	10.45
Willow	1.09	-	-

Table 7.9: Additional decisions present in the FSD efficient set absent from the SSD set

Objective	Crop	Radius (km)
Worst-case	No crops	-
Expected value	Miscanthus	8.18
Risk-Area-ration	Miscanthus	7.59

Table 7.10: Decisions from the single objectives

cuts down the number of decisions to about half the original (44 instead of 84 in the feasible set), the SSD approach further eliminates half of that, reducing the FSD set to about half its original size. The mean-variance efficient set is virtually identical to the SSD set for this example [1].

The predominant decisions in the efficient sets are those of miscanthus and corn-stover for different farm radii. This is not surprising since the deterministic approach shows that the other two crops - switchgrass and willow - are either profitable at very small radii (switchgrass), or not at all (willow) with. Switchgrass interestingly is effectively out-competed by miscanthus. This result can be partially attributed to the prices and costs we assumed.

If we, for example assume that a farmer is able to charge the average premium of $10 over growth costs then we have both the deterministic and uncertain plots change to give the results in Figure 7-6 and Table 7.11
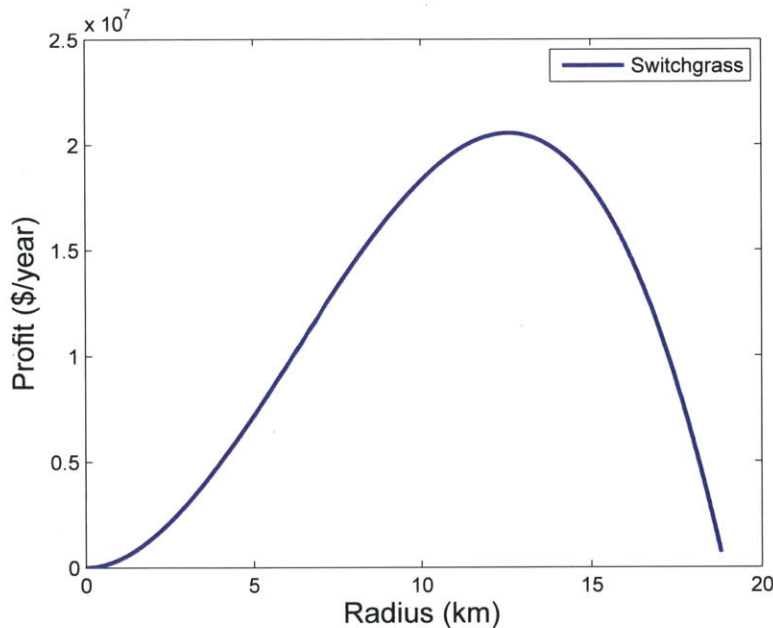
Figure 7-6: Farm profit as a function of radius for Switchgrass

We see that the SSD set changes quite dramatically with the modification of the price of switchgrass. This is not surprising as sensitivity results show that the prices of the crops have a strong effect on the profitabiity and this will be discussed in more detail in the next section.

Also, save for one decision (crop and radius), the mean-variance set is identical to the SSD set - an outcome that is different from the results of the previous chapter. One reason for this could be that the distributions were fairly symmetrical about the mean and so penalizing spread (variance/standard deviation) is the similar to

---

[1]Virtually, because the only decisions missing are the options to not plant any crops as well as the decision to grow the Willow crop with a farm radius of 0.55km.

Switchgrass	Miscanthus	Willow	Corn Stover
11.52	0.58	0.55	2.67
10.36	-	-	-
9.21	-	-	-
8.06	-	-	-
6.91	-	-	-
5.76	-	-	-
4.61	-	-	-
3.45	-	-	-
2.30	-	-	-
1.15	-	-	-
No crops	-	-	-

Table 7.11: Second order Stochastic Dominace efficient set with the assumption of a change in the price of switchgrass

Crop	Radius (km)
Miscanthus	4.09
Miscanthus	4.67
Miscanthus	5.26
Miscanthus	5.84
Miscanthus	6.42
Miscanthus	7.00
Miscanthus	7.59
Miscanthus	8.18
Corn Stover	2.61
Corn Stover	5.22

Table 7.12: Decisions left in the final ASD reduced set

penalizing the downside risk. When such is the case, as we saw in Chapter 5, we know that both methods can be virtually identical.

7.4 From efficient sets to decisions

Like the previous chapter, given the size of the efficient set (20 decisions) and the significant uncertainty still inherent in it, we explore possible actions that can proceed from the preceeding analysis.

Almost Stochastic Dominance As we did in the previous chapter, we used the ASD framework to try to reduce the size of the efficient set. Figure 7-7 shows the result

As we can see, the size of the reduced efficient set is still relatively large compared to the original set, with only about 55% reduction achieved (from 22 to 10 decisions).
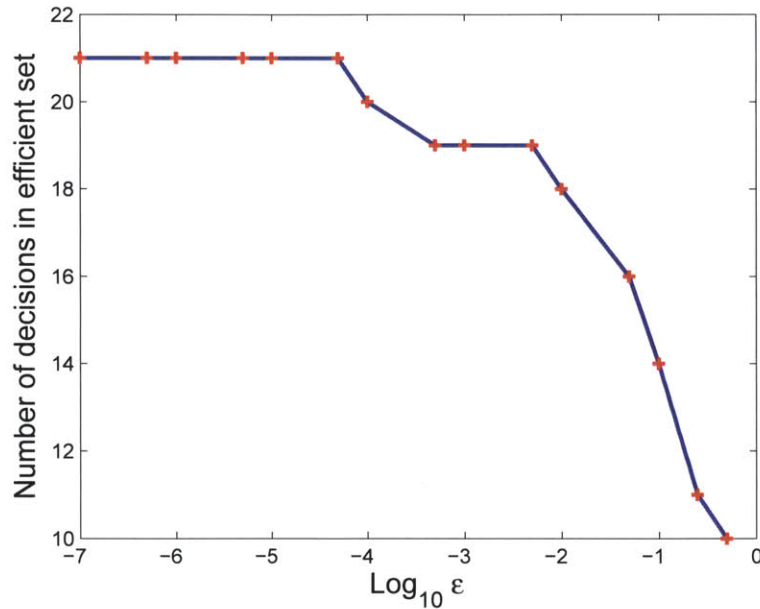
Figure 7-7: Change in the size of the efficient set with using ASD with different values of ϵ

Crop	Radius (km)
Miscanthus	0.58
Corn Stover	1.31
Willow	0.54
No Crops	-

Table 7.13: SSD efficient set for the profitability index metric

The decisons that comprise the final set are given in Table 7.12. As we see, Miscanthus still dominated the crops present in the set, accounting for 8 out of the 10 decisions in there.

Alternative metrics   Net present value, as we saw, is but one metric that can be used to decide between alternatives. We illustrate the use of the Profitability index as a second alternative for reducing the size of the efficient set. The profitability index is defined as

$$PI = \frac{Profit}{Capcost} \tag{7.20}$$

To do this we run the uncertain analysis using the different metric and pick the decisions that lie at the intersection of the two SSD efficient sets. The table below shows the efficient sets that result from the application of the SSD procedure to the profitability index metric is given below

The challenge with this metric (and many ratio metrics) is that they tend to favour small - medium investments (cite reference). From the analysis and a comparison with

Crop	Price	Cost	Yield	Transportation cost	Establishment cost
Switchgrass	0.36	0.36	0.15	0.20	0.02
Miscanthus	0.28	0.28	0.15	0.03	0.24
Willow	0.34	0.34	0.05	0.03	0.21
Corn stover	0.20	0.58	0.06	0.15	0.04

Table 7.14: Averaged total sensitivities for the various parameters determined for each crop

Parameter	Switchgrass	Miscanthus	Willow	Corn stover
Selling price (old)	[60, 70]	[60, 70]	[60, 70]	[60, 70]
Selling price (new)	[65, 70]	[60, 65]	[65, 70]	[60, 65]
Growth cost (old)	[55, 65]	[50, 60]	[50, 60]	[43, 60]
Growth cost (new)	[60, 65]	[55, 60]	[50, 55]	[50, 60]

Table 7.15: Table of parameters assumed for 'new information' entering the efficient set. Units for data are in $/ton

the original SSD efficient set in Table 7.8 we see that not much reduction in the size of the efficient set is obtained, unlike the previous case study. Thus, we can't always rely on using alternative metrics to reduce the efficient sets.

Experiment design and uncertainty reduction   As introduced in the Chapter 3, one of the big uses of uncertainty propagation and analysis is the determination of the uncertainty drivers of a problem with an eye to reducing them - in order to aid decision making. We carried out such an analysis on the uncertainties in the inputs and the results are given in Table 7.14, a table that shows the average sensitivities for the different crops and for the different parameters. This averages are taken over the possible ranges of the radius that was allowed by the budget.

In general, the crop selling price (amount the farmer expects to receive from the sale of the crop) and growth cost are the parameters to which all the crops are most sensitive to while the remaining three parameters vary in order of importance depending on the crop. For miscanthus and willow, the establishment cost dominates the yield and the transportation costs for importance while for corn stover and switchgrass, there is a greater sensitivity to transport costs than establishment cost and yield.

So from Table 7.14 we see that the selling prices and the growth costs are, on average, the biggest drivers of the uncertainty of the two crops - miscanthus and corn stover - that form the bulk of the decisions in the efficient set. Knowing this means that it may be helpful to explore ways to reduce the uncertainties in these decisions. We explore this a bit by assuming that we are able to reduce the uncertainties in some of those parameters to distributions resembling. The new parameters are given in Table 7.15 with the resulting efficient set given in table ??

The use of new information results in a significant reduction of the SSD, elimi-

Crop	Radius
Miscanthus	2.34
Miscanthus	1.17
Willow	0.55
Corn Stover	6.53
Corn Stover	5.22
Corn Stover	3.92
Corn Stover	1,31
No crops	-

Table 7.16: Resulting SSD set upon new information reducing the size of the efficient set

nating a number of decisions in the original set. While it doesn't reduce the efficient set to a singleton decision (easiest to implement), it advances the analysis by further reducing the number of decisions left in the efficient set.

7.5 Conclusions

This case study focused on the problem, from a farmer's perspective, of the best renewable energy crop to grow and on what farm size to grow it on to yield maximum value to the farmer. We analyzed the problem both deterministically and in the presence of key uncertainties in some of the parameters affecting the crop returns - in particular the establishment costs, the yield, selling price, cost price and the transportation cost.

In the deterministic case, it was found that the best decision was to grow Miscanthus on a farm radius of about 7.8km. This was despite the rather high establishment costs of miscanthus as the combination of lower growth costs and its high yield and long lifetime contributed to overcoming this barrier. In the presence of uncertainty, Miscanthus still retained its dominance as the crop to grow as 14 of the 22 decisions in the SSD efficient set were all Miscanthus choices. The farm size did vary quite a bit within the Miscanthus set, ranging from a low of 0.6 km to a high of 8.2 km. Larger farms produced better NPVs in the best times but did worse in poorer times and it is in the selection of this variable that the particulars of the decision maker's risk profile may be needed.

The SSD efficient set and the mean-variance set were nearly identical, illustrating that in this case, the advantage to using SSD wasn't nearly as big as that in the previous study.

Sensitivity analyses showed that the price and the yields were the biggest influences on the uncertainty in the output and changing the price of switchgrass made a big difference in the composition of the SSD efficient set. In modelling the effect of new information entering the set via the reduction of the variances of the inputs, we were able to illustrate a reduction in the size of the efficient set.

Variable	Sources
Yield	[61] [83] [85] [56] [74] [115] [119]
Establishment cost	[32] [6]
Selling price and growth cost	[83] [1]

Table 7.17: Data sources for the Switchgrass

Variable	Sources
Yield	[118] [56] [73]
Establishment cost	[6] [26]
Selling price and growth cost	[2] [118][56]

Table 7.18: Data sources for Miscanthus

Almost stochastic dominance was not as efficient in reducing the size of the efficient set as the previous case study as, after the analyis, we still had about half the original decisions present in the set. The use of an alternative metric to weed out the decisions helped to reduce the efficient set down from the initial set of although it favoured the smaller farm sizes and will prove to be unsatisfactory when we consider the next chapter focused on the larger scale problem of the production of biofuels from biomass. There, the advantages of economies of scale in production output encourage the growth of bigger farms to supply adequate biomass input - and it may be advantageous for the operators of such a facility to offer higher sale prices to farmers to encourage them consider larger farms.

The analysis thus indicates that Miscanthus is the best crop for the farmer to grow. More information that reduces the uncertainty about the prices the farmer for the crop from buyers, in addition to the growth cost of the crop will be useful in determining the eventual size of the farm.

7.6 Notes on sources

The data for that was used for the simulations was obtained from a range of sources. There was quite a bit of variation in the numbers across the literature so in the end we often had to select either what seemed most reasonable after consulting a number of sources or take an average of a set of the numbers we had. We took this approach since the goal of this study was a demonstration of the SD analysis methodology on a simple yet realistic case. Rigorous recommendations will require better models and more data. Tables 7.17, 7.18, 7.19 and 7.20 presents some of the sources from which the data for the crops were estimated.

Sources for transportation costs (assumed same across the crops) include [1], [56] and [6] and costs in different currencies (like the Euro) were converted using an average that prevailed around the time of the data publication. Uncertainties were modelled by adding appropriate spreads around the best numbers we came up with from the sources, taking the range observed across the literature into account.

Variable	Sources
Yield	[5] [44] [120] [124] [73] [104]
Establishment cost	[44] [59] [73] [54]
Selling price and growth cost	[5] [44] [120] [109]

Table 7.19: Data sources for Willow

Variable	Sources
Yield	[60] [110]
Establishment cost	[40]
Selling price and growth cost	[48] [60] [110]

Table 7.20: Data sources for the Corn stover

Chapter 8

Case Study 3: Pathway selection for the production of biofuels

Increasing concern about the rate of use of fossil fuels to power today's global economy as well as the security issues that go hand in hand with the increase in fossil fuel use has led a lot of countries to invest heavily in the development of renewable, greener sources of energy in the near future to replace the ubiquitous fossil fuels.

A major aspect of that effort has been to try to utilize biomass - a resource which is widely available. The first of these efforts was the conversion of biomass to ethanol, using corn as the biomass of choice. This was initially hailed as successful and promising for the future of energy generation. However, in retrospect, corn ethanol has suffered from very serious drawbacks [50] like

1. Its alteration of the economics of food prices as more farmers were switching to growing corn for conversion to ethanol since they benefitted from government subsidies and thus led to a shortage of corn for food and driving up the price

2. Its near zero net energy output as many studies showed that the manufacture of bio-ethanol from corn took up as much energy as was used in the conversion process, making the entire process redundant

As a result, more efforts in the recent past have now been focused on the conversion of non-food biomass like (primarily cellulose) into alternative fuels. Two main research streams diverge at this point in trying to accomplish this goal. The first is the use of microbes and the manipulation of biochemical pathways within these organisms to directly produce industrial chemicals. Examples of such include (list them). Many of these approaches are still in the laboratory phase although current research also focuses on how to scale these up to industrial production levels.

The second stream is the use of thermochemical means to convert the cellulose to crops. Most of this has been adapting technology developed in the late 20th century for the transformation of coal into sythesis gas (gasification) and subsequently into synthetic fuels (Fischer-Tropsch conversion). This approach to biomass utilization is much closer to industrial scale application and a number of pilot plants have been (or are being built) to test the economic and long-term feasibility of the technology. It is
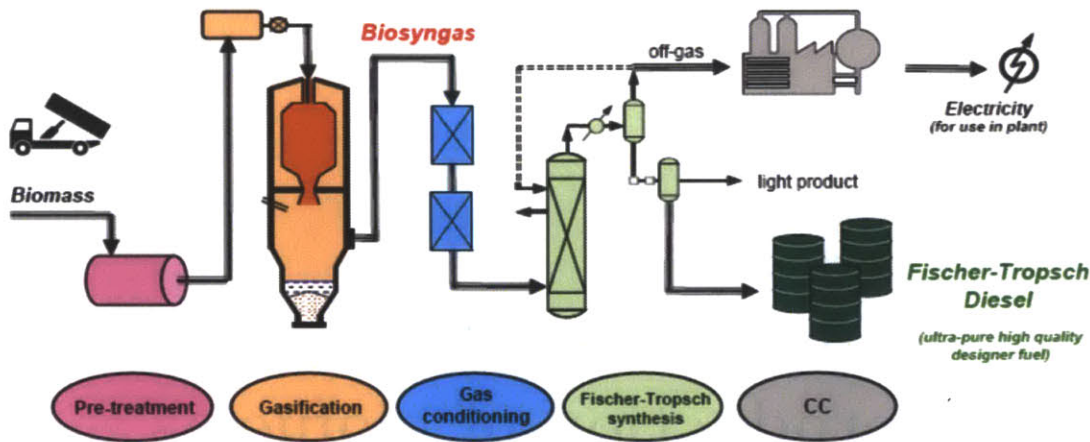
Figure 8-1: Schematic that illustrates the different steps that form the essence of the biomass to liquids process. From [17].

this path that we will be examining in greater detail in this chapter. First we discuss the overall structure of the biomass process and then develop simple mathematical models that can be used for the analysis and pathway selection in the presence of uncertainty. The discussion is that follows is largely a summary from the references [102] and [19] and readers interested in a more depth are invited to consult them for more details.

8.1   The Biomass to Biofuel conversion process

The conversion of biomass to biofuels takes place via the following main steps

1. Feed processing and pretreatment

2. Gasification

3. Gas clean-up and conditioning

4. Fischer-Tropsch synthesis

5. Product upgrading/reforming

These are illustrated in Figure 8-1 and we discuss each in some more detail in the subsections below

8.1.1   Feed processing and pre-treatment

Harvested biomass crop is not often in the best form appropriate for its use in the gasification stage. Depending on the source of the biomass, a number of modifications

152

to the form and or content needs to be done before it can be used. Size reduction and or compaction form (e.g. chopping, baling) form one primary pretreatment step as they allows for better handling and processing of the biomass during combustion or gasification. Drying is also another important pretreatment step as too much moisture in the biomass reduces the heating value of the fuel. And, depending on the distance between the source of the biomass and the processing facility, some of the steps are carried out on the farms. The side benefit of this is that it can result in lower transportation costs and thus allow for biomass to be sourced from distances much further away from the farm.

8.1.2 Gasification

Gasification is the conversion of solid biomass to combustible gases in the presence of an amount of oxygen insufficient for complete combustion. The goal of gasification in the BTL process is to convert the solid biomass into a gaseous fuel mix known as synthesis gas - a mix of carbon monoxide and hydrogen - which can later on be either completely burned to generate electricity or converted to liquid fuels via Fischer-Tropsch synthesis.

Gasification takes place in three stages. The first stage is the pyrolysis stage where the biomass is heated in the absence of oxygen, a process useful for driving off the volatile components present in the fuel like water vapour, carbon monoxide and carbon dioxide. This stage occurs at lower temperatures (225 - 500 degrees Celsius [19]).

After the pyrolysis stage comes the gasification stage that is marked by 4 key reactions, all of which occur between the char and the vapours released during the pyrolysis. The first is the oxidation of carbon (char) that remains after the pyrolysis stage to carbon dioxide i.e

$$C + O_2 \longrightarrow CO_2 \tag{8.1}$$

Next comes the Boudard reaction where the carbon dioxide is reduced by more char (carbon) to give carbon monoxide:

$$C + CO_2 \longrightarrow 2CO \tag{8.2}$$

Two other important reactions also take place in this stage. One of the more important ones is the water-gas reaction where water vapour released in the pyrolysis stage further reacts with some of the char and is reduced to hydrogen i.e.

$$C + H_2O \longrightarrow CO + H_2 \tag{8.3}$$

And finally comes the methanation reaction where some of the hydrogen produced also reacts with the char to produce methane

$$C + 2H_2 \longrightarrow CH_4 \tag{8.4}$$

Some reactions also take place in the gas-phase between the gases produced by

the gas-solid reactions. The most important of these are the water-gas shift reaction where carbon monoxide is further oxidized by water vapour to produce carbon dioxide and yet more hydrogen gas:

$$CO + H_2O \longrightarrow CO_2 + H_2 \qquad (8.5)$$

And finally the gas-phase equivalent of methanation where carbon monoxide reacts with some hydrogen to produce more methane gas as well as water vapour

$$CO + 3H_2 \longrightarrow CH_4 + H_2O \qquad (8.6)$$

These are the principal reactions that describe the gasification process. Energy for the reactions is supplied by the exothermic reactions - carbon oxidation and methanation reactions - although most of it is supplied by the incomplete oxidation of solid carbon to carbon monoxide. The fnal composition of the synthesis gas exiting the gasifier as well as the temperature of the gasifier is dependent on the amount of oxygen (or air) fed into the gasifier and the composition of the initial biomass. Composition is represented by the oxygen-carbon (O/C) and hydrogen-carbon (H/C) ratios of the fuel and the amount of oxygen fed into the system represented by the **equivalence ratio**.

To define the equivalence ratio, we first note that an overall stoichiometric equation for the combustion of biomass given as

$$CH_{1.4}O_{0.6} + 1.05O_2 \longrightarrow CO_2 + 0.7H_2O \qquad (8.7)$$
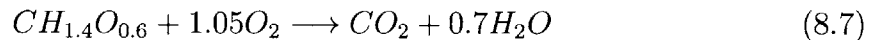
where $CH_{1.4}O_{0.6}$ is a rough approximation of the composition of biomass [19] . Thus roughly about 32.16 kg of oxygen gas is needed for the complete combustion of 23 kilograms of biomass. The equivalence ratio for gasification is defined as the ratio of the amount of oxygen fed into the process to that required for complete combustion. Prins et al. [87] note that (molar) equivalence ratio can range from about 0.24 for cellulose to about 0.5 for graphite, highlighting the dependence of equivalence ratio on composition. However, they also note that most fuels level off at an equivalence ratio of about 0.29 at lower gasification temperatures (less than 1000 degrees Celsius) and about 0.32-0.33 at the higher temperatures (higher than 1200 degrees Celsius).

It should also be noted that besides gasification with oxygen or air, it is possible to carry out gasification with steam where the water vapour is the carrier of the oxygen required for the oxidation of carbon. In our analysis however, we will focus on gasification with pure oxygen. Using oxygen instead of air for gasification reduces the size of process equipment required for gasification (less volume flowing through the process) and it can also achieve higher conversion efficiencies (cite reference)

Gasifiers are predominantly classified according to the processes by which the fuel is transported/conveyed through the equipment. This classification gives rise to three types of gasifiers

1. Fixed bed

2. Fluidized bed

154

3. Entrained flow

In fixed bed gasifiers, hot gas is passed through bed of solid biomass. There are two main types - the updraft and the downdraft fixed bed gasifiers. In both cases, feed is introduced into the system from the top of the gasifier and gasification occurs as the biomass falls through it. In updraft gasifiers, the oxidizer (air or oxygen) is fed at the bottom in counterflow to the movement of the biomass and the syngas produced exits near the top of the gasifier. Downdraft gasifiers on the other hand have the flow of the oxidizer co-current with the flow of biomass and the product gas exits the gasifier near the bottom.

In fluidizied bed gasifiers, gas moves through the bed of solid biomass at such a velocity that the mixture begins to behave like a fluid (hence the name). Solid, crushed fuel is mixed with inert material and continuously introduced into the gasification chamber. The large velocities needed for fluidization means that air (as opposed to oxygen) is often the oxidizer of choice, because of the large volumes needed to sustain such velocities. As with fixed-bed gasifiers, there are also two types of fluidized bed gasifiers - bubbling fluidized bed (BFB) and circulating fluidized bed (CFB) gasifiers. In bubbling fluidized bed gasifiers, the velocity of the gas stream is such that it is just enough to keep the bed agitated but in the same position. In circulating fluidized beds though, the gas velocities are high enough to physically transport the bed particles which have to then be returned to the position using cyclones.

Entrained flow gasifiers are the final type of gasifiers and here, the solid fuel is injected as very fine particles into the oxidizing stream and the reaction occurs as this stream flows through the gasifier at very high temperatures.

There are advantages and disadvantages to using each of the different types of gasifiers. Fixed bed gasifiers are cheap and simple to construct but tend to produce syngas with a lot of tar and particulates - leading to high clean-up costs.

Fluidized bed gasifiers give high reaction and heat transfer rates and uniform conditions across the bed. These lead to higher conversion of the biomass with low tar production. However, fluidized beds tend to also be operated at lower gasification temperatures than the other two types of gasifiers to prevent the ash content in the fuel from reaching fusing temperature which would cause the bed to defluidize. These lower temperatures lead to some loss in efficiency and biomass conversion that could have taken place if the temperatures were high enough for cause tar cracking.

Entrained flow gasifiers would give the highest conversion of biomass to syngas and very little residual tar since the temperatures are high enough to cause tar-cracking. However, the amount of feed processing required for the potential.

8.1.3 Syngas clean up and conditioning

Regardless of the source of biomass and the choice of gasification, impurities usually end up in the syngas product from the gasifier and these have to be removed before the gas can be synthesized into fuels in the downstream Fisher-Tropsch processes as the impurities poison catalysts and can significantly reduce their performance and lifetime. Permanent gaseous impurities like oxides and hydrides of sulfur and nitrogen

are removed using solvent-based methods where vapourized impurites like tars and alkaline compounds are removed by first cooling the product gas to temperatures that allow the vapours to condense (and or solidify) and then subsequently filtering the mixture in order to get rid of the condensates alongside the particulate impurities like ash.

Syngas conditioning takes place after clean-up and here the goal is to optimize the $CO - H_2$ ratio in the syngas for the particular FT application. This is primarily achieved via the water-gas shift reaction where steam is used to oxidize some of the carbon-monoxide to carbon dioxide, with the subsequent production of more hydrogen gas.

8.1.4   Fischer-Tropsch Synthesis

Fischer-Tropsch synthesis is the name given to a wide variety of reaction mechanisms that aim to convert synthesis gas to liquid hydrocarbons of various forms. This reaction was originially discovered by Fisher and Tropsch in the 1920s but since that time, many similar reactions with different catalysts have been observed. The generic reaction is given as

$$nCO + 2nH_2 \longrightarrow (CH_2)_n + nH_2O \tag{8.8}$$
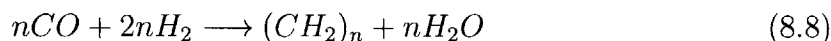
The final product distribution is given by the Anderson-Schultz-Flory distribution as

$$W_n = n(1 - \alpha)^2 \alpha^{n-1} \tag{8.9}$$

where n is the number of carbon atoms in a product, W_n is the weight fraction of a product of chain length n in the final product mix and α is the growth probability factor. Theoretically, only methane ($n = 1$) and waxes ($n > 19$) can be produced with near 100% selectivity - corresponding to very low and very large growth probability factors respectively - with the intermediate chain-length products like gasoline and diesel having only a peak selectivity of about 40 - 50 percent maximum. The distribution is illustrated in Figure 8-2. [43] state that the α is dependent on the catalyst, temperatures, pressures and the applied Fischer-Tropsch technology (currently developed by different companies) but other factors like the feed gas composition and the presence of promoters can affect the distribution too [19].

In general though, the choice of catalyst for the FT reactor determines a number of other significant process parameters. The use of cobalt-based catalysts gives rise to a higher selectivity for the heavier end products hence producing more diesel and wax while iron gives products on the lighter end, with a preference towards the production of gasoline. Cobalt catalysts also lend themselves to lower reaction temperatures while iron catalysts can utilize a greater range of reaction temperatures. Cobalt also requires a narrow and high range of the hydrogen-carbon monoxide ratio while iron can tolerate a wider range since it can also catalyze the water-gas shift reaction unlike cobalt.
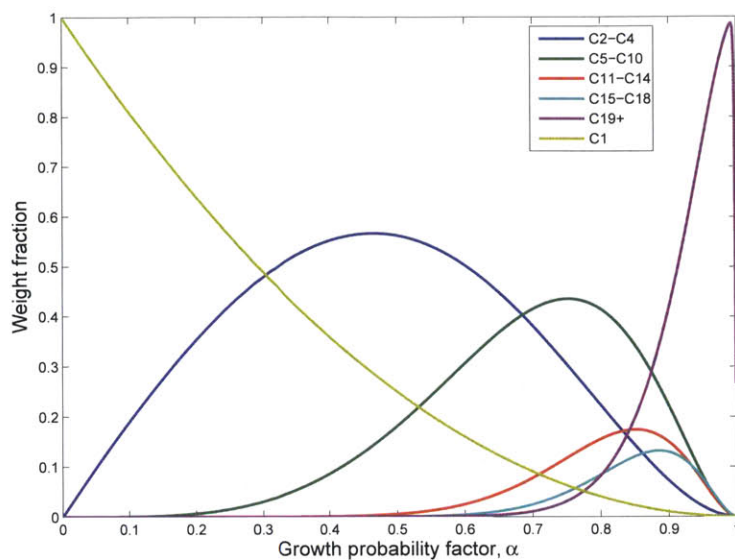
Figure 8-2: Weight distribution of the different chain lengths from FT synthesis

8.1.5 Product upgrading

The production of transportation fuels via FT does not take place with complete selectivity. There is often the generation of solid waxes and even the necessary products generated are often impure and need to be treated before they can be sold. For the production of diesel fuels, the waxes and other large-carbon chain compounds are hydrocracked to give the final product [111].

8.1.6 Electricity and Chemical Products

While the above processes have outlined the general process for transforming biomass to liquid (transportation) fuels, it should also be highlighted that there is the potential of generating other products in the biomass transformation. Electricity can be generated by the combustion of some syngas or utilization of heat generated during gasification. Other chemicals like Methanol, dimethylether and even hydrogen can be generated as alternative product options beside the main-stream transportation fuels. In this study, we will only consider the production of the transportation fuels like diesel and gasoline and the possible generation of electricity as the principal desired outcomes of the BtL plant we will be analyzing.

8.2 Mathematical Model for Biomass to Liquids Process

In order to select a production pathway among all the alternatives, it is necessary to build a mathematical model that adequately represents the physical transformations

157

and economic value added. Before we do so, it will be useful to highlight some of the key decisions that need to be made

1. The amount of biomass feed to be processed (tons per day)

2. The choice of gasifier for the gasification process - either an entrained flow or circulating fluidized bed gasifier

3. The choice of reforming the methane gas generated in the gasifier for production of FT liquids or using it to generate electricity

4. The fraction of the syngas generated that should be diverted for the production of electricity

5. The choice of catalyst for the FT process as well as the product mix.

6. The selectivity of the FT reaction

7. The choice of a once-through or a twice through reaction process

The schematic diagram for the decision framework presented is illustrated in Figure 8-3. Of the variables, three of them are continuous in nature: amount of feed processed, fraction of syngas used for electricity generation and FT selectivity, while the others are binary choices. To use the model with our decision analysis framework, it is important that we discretize it in order to have unique decisions. The range of choices for these variables, as well as how they affect the eventual outcome are discussed in detail in the subsequent sections.

The overall model consists of technical and economic components. The technical model tracks the physical (and chemical) transformation that converts the input biomass into output products. The economic model tracks the capital and operational costs of setting up the plant as well as the revenues to received from the sale of plant products. We will develop both the technical and economic models separately.

Both models follow that developed by Hamelinck et al. [43] with appropriate revisions made where necessary in order to adapt it to our uses. Since Hamelinck et al. borrow from the model developed by Tijmensen et al. [111], some of the structure in that model is also evident in ours as well and we highlight as much where that is the case. We refer interested readers to both papers for more details on the models as we discuss only parts of it, highlighting the areas where modifications were made.

8.2.1 Technical model for the Biomass to Liquids Plant

Feed preparation and conversion to syngas As mentioned, we will be using the results of the previous chapter, assuming that we have a source of biomass supply for the feed. We assume that we receive wet unprocessed biomass and as such estimate that pre-treating the biomass will result in some loss (moisture and some carbon content). Thus the amount of usable biomass that will eventually make it to the gasifier is determined by the equation

$$x_2 = k_1 \times (1 - mf) \times x_1 \tag{8.10}$$
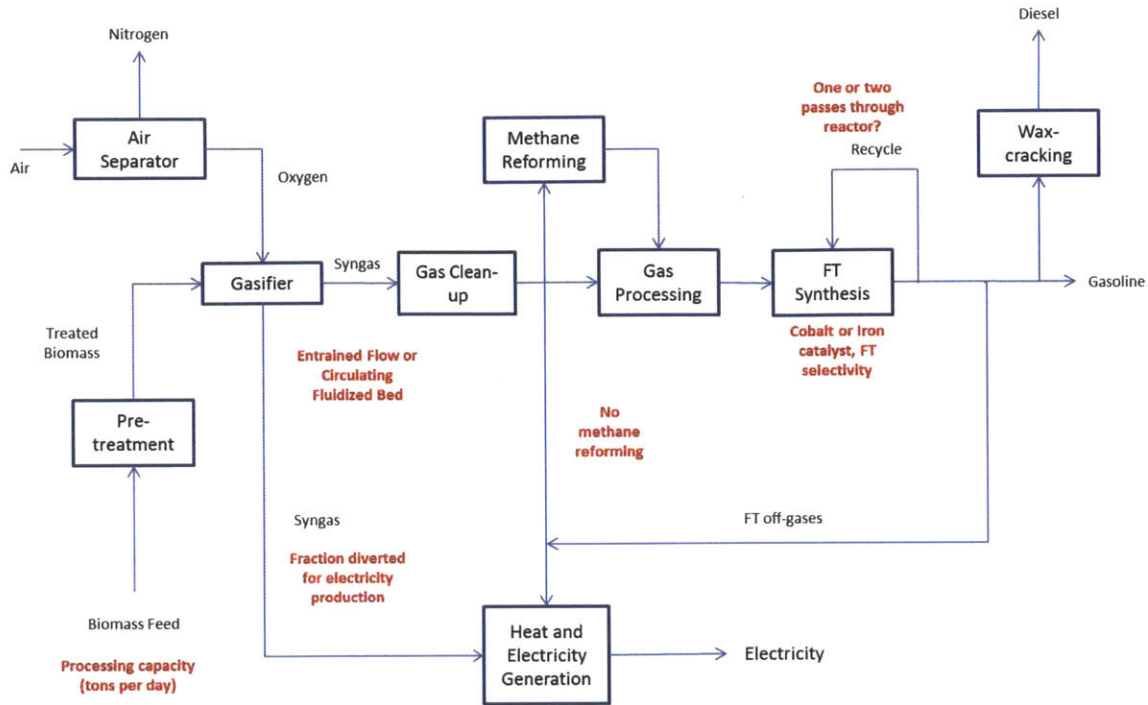
Figure 8-3: Schematic strucure for BTL model. Text in bold red font represent the decisions to be made

where x_1 represents the amount of biomass feed coming into the plant (tonnes/day), x_2 represents the prepared biomass feed that goes into the gasifier (tonnes/day), mf is the fraction of moisture present in the feed, and k_1 represents the efficiency of the overall process, assumed to be about 95%. The value of k_1 depends on the nature of the biomass assumed to be coming into the plant and how much mass is lost during processing. The amount of moisture present in the biomas feed varies and this variation will be considered when we analyze the model in the presence of uncertainty. We assume a mean value of about 30% in line with [111].

Gasification   Gasification can either take place using air or oxygen. In our process, we assume a choice of oxygen for gasification and this is supplied using an air separation unit. The capacity of the oxygen plant required is a function of the amount of biomass that the facility processes and the equivalence ratio required for the biomass gasification

$$X_{O_2} = k_{eq} \times x_2 \tag{8.11}$$

where X_{O_2} is amount of oxygen supplied and k_{eq} is the equivalence ratio. The equivalence ratio in general varies between 0.26 and 0.3 and so we assumed a mean value of 0.28. The products of gasification were generated via a material balance from the overall gasification equation

$$CH_{1.4}O_{0.6} + yO_2 \longrightarrow aCO + bH_2 + cCO_2 + dH_2O + eCH_4 \qquad (8.12)$$
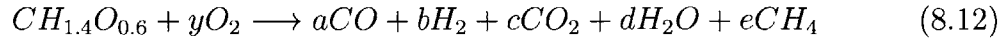
This equation is a sum of the principal components of all the gasification equations highlighted in the process description section. y represents the equivalence ratio of the plant and the letters a - e represent the components of the respective species in the equation i.e. carbon monoxide through methane. From an atom balance, three independent equations can be written for the variables. The carbon balance yields

$$a + c + e = 1 \qquad (8.13)$$

and the hydrogen balance yields

$$b + d + 2e = 0.7 \qquad (8.14)$$

and finally the oxygen balance gives

$$a + 2c + d = 0.6 + 2y \qquad (8.15)$$

Since there are five variables and only 3 balances, we need two other equations to uniquely determine the composition of the exiting gas. These are dependent on the nature of the gasifiers. Entrained fluid gasifiers give higher quality gas than the fluidized bed gasifier and this is reflected in the percentage of carbon monoxide and hydrogen components of the syngas produced. We set the percentage produced by the entrained fluid gasifier as 0.85 and that produced by the circulating fluidized bed gasifier at around 0.65 [1]. This gives a fourth equation for the set of variables above as

$$a + b = r_1 \times (a + b + c + d + e) \qquad (8.16)$$

where r_1 is the ratio. And finally, since the goal of gasification is the production of CO and not the production of the other components, we set the final quality equation ratio of CO in the syngas to that of the other carbon bearing components i.e.

$$\frac{a}{c + e} = r_2 \qquad (8.17)$$

We assume that r_2 is about 3.5 for CFB and about 4 for entrained fluid gasifiers (check and make sure your ratios are correct. These allow us to solve the gasification equation to get the composition of the exiting syngas that can then be used for the production of FT liquids.

Gas cleanup and processing Hamelinck et al. include equipment for gas clean-up and processing and the equipment needed in those require technical calculations. The basis for the calculations are all given in the basis of the volume of gas exiting the gasifier. This volume is calculated using the ideal gas law as an approximation and then

---

[1]These numbers are within the range suggested by [4] - a vendor publication

accounting for contractions in non-ideality. For a given pressure and temperature, the volume of an ideal gas is given by

$$V = f \times NRT/P \tag{8.18}$$

where T represents temperature, P, pressure, V, volume, N, total number of moles, V, the volume of the gas and f the factor that accounts for non-ideality of the gas[2]. Appropriate conversion units are used to convert the volume. Entrained flow gasifiers are cleaner than fluidized bed gasifiers so we modify the costs and equipment appropriately to account for this difference. We omit cyclones and particle filters in the cost for the entrained flow as well as the tar cracker but leave all these in the fluidizied bed gasifier.

The exiting flowrates of all the gases were obtained by multiplying the composition obtained from the gasifier with the inlet molar flow rate into the gasifier. Molar flow rate was calculated using the average biomass composition. An effectiveness factor was incorporated to allow for the fact that not all the biomass is converted to syngas - that some is lost in the conversion to syngas via tar and other particulates. A conversion factor of 0.95 is used for entrained flow and 0.9 is used for the CFB gasifier.

After cleaning up the gas, we also process it to have it in the form needed for FT synthesis. Again, Hamelinck et al. [43] provide a breakdown of the equipment needed to accomplish this and we modify these as necessary depending on the particular choice of the FT process.The ratio of hydrogen to carbon monoxide produced after gasification is often smaller than is required for the synthesis of FT liquids. The reaction requires about 2 moles of hydrogen to 1 mole of carbon monoxide but the output of the gasifier is usually less than 1:1. To accomplish the desired ratio, some water gas shifting and/or steam/autothermal reformation of methane to syngas components. The water gas shift reaction is already given in the gasification section in equation 8.5. This equation produces 1 mole of hydrogen gas for every mole of carbon monoxide reacted. The steam reformation of methane is given as [90]:

$$CH_4 + H_2O \longrightarrow CO + 3H_2 \tag{8.19}$$

$$H_2O + CO \longrightarrow CO_2 + H2 \tag{8.20}$$
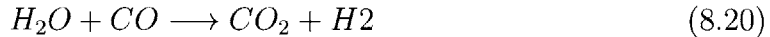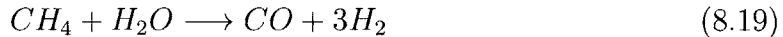
where the second equation is simply the water-gas shift reaction. Halabi et al[42] give the overall reaction for autothermal reforming as

$$CH_4 + xO_2 + (2 - 2x)H_2O \longrightarrow CO_2 + (4 - 2x)H_2O \tag{8.21}$$

where x in the above equation represents the oxygen to carbon molar ratio.

The big difference, as can be seen from the two reactions presented above is that steam reforming doesn't require oxygen while autothermal reforming does. Autothermal reforming also produces a lower amount of hydrogen than the steam reforming

---

[2]We ignore the factor by setting its value to 1. For more advanced models, more appropriate equations of state can be used to estimate the value of f under different conditions

Variable	Definition	Variable	Definition
b	Ratio of H_2 to CO in FT reactor	z	Amount of CO shifted
X_1	Amount of CO produced in gasifier	Y_1	CO entering FT reactor
X_2	Amount of H_2 produced in gasifier	Y_2	H_2 entering FT reactor
X_5	Amount of CH_4 produced in gasifier	Y_3	Total CO_2 produced

Table 8.1: Variable definitions for the FT balance

(can be seen from the equations) but is likely cheaper to implement [43] and will be used in this study as was used in our reference.

To thus calculate the amount and components of the syngas flowing into the reactor, we need to include the shifting and reforming. Shifting produces 1 mole of hydrogen gas for every mole of carbon monoxide consumed while we will assume that autothermal reforming produces about 2.6 moles of hydrogen to 1 mole of CH_4[3]. We assume that this number covers all of inefficiencies in the conversion process. To determine the carbon content of the entering feed, we have to do a material balance using the equations. We define the variables used in the following analysis in Table 8.1

Given those variables, the amount of CO that needs to be shifted to H_2 is given as

$$z = \frac{bX_1 - X_2 - 2.5X_5}{1+b} \tag{8.22}$$

The resulting amount of carbon monoxide and hydrogen entering the FT reaction phase as a result of the shift and reformation becomes

$$Y_1 = X_1 - z \tag{8.23}$$
$$Y_2 = X_2 + z + 2.5X_5 \tag{8.24}$$

the final amount of the carbon dioxide generated in the gasification and processing stage is

$$Y_3 = X_1 + X_5 + z \tag{8.25}$$

And with these equations, we can determine the appropriate downstream variables.

FT processes take place over a range of pressures - and the particular pressure desired determines the cost of the compression. We choose a pressure of around 40 bars as the representative pressure which is the average between the range of 20 and the pressure of 60 bars found in literature [43]. We assume a mean temperature of about 550 Kelvin which is near the average of the temperatures required for the different catalysts (Iron has higher temperatures of 300-350 degrees Celsius and cobalt between 200 and 250 C [107]). We assume to first order that the compressive work is

---

[3]This was one of the yields obtained in simulations reported by [42]

that obtained by an isothermal compression of the gas to the desired pressure which gives that the work done is obtained by

$$W = NRT \times \ln(P_2/P_1) \tag{8.26}$$

where the final pressure, P_2 was eventually chosen to be 40 bars and the initial pressure, P_1 was the gasification pressure of 20bars.

FT production and upgrading The amount of FT products generated were calculated based on the Anderson-Schultz-Flory distribution parameter, α. This parameter features in our model as a decision variable based on the choice of catalyst. Swanson [107] gives the distribution of α as

$$\alpha = [0.2332 \times \frac{y_{CO}}{y_{CO} + y_{H_2}} + 0.6330] \times [1 - 0.0039(T - 533)] \tag{8.27}$$

As can be seen, the value of α is dependent both on composition and Temperature. FT with a cobalt catalyst requires a hydrogen-carbon monoxide ratio of about 2:1 while FT with an iron catalyst can do with a wider range of 0.7 - 1.7. Because we can tune the value of α by adjusting a number of other things like pressure, temperature and composition, we make it a decision variable and not a variable. However, given the stricter composition requirements for the cobalt catalyst and the more variable requirements with iron, the possible ranges of the value of alpha varies depending on the catalyst choice and this is reflected in the model. Cobalt gets a smaller range of 0.8 - 0.9 while Fe gets a larger range of 0.6 - 0.9.

Cobalt as a catalyst is more expensive than iron however it has a longer half life. To account for the difference in catalyst cost balanced for different lifetimes, we assume that the FT reactor cost for iron is decreased by a factor of 4 (the reference paper uses cobalt as the basis catalyst). We chose this because it we lacked data to estimate the actual ratio[4] and it gave a reasonable order of magnitude for the range of costs.

To size the reactor, attempts where made to use the reaction rates developed in [43] in order to determine total catalyst weights and subsequently catalyst volumes. However, data was only presented for the cobalt case and the presented data didn't seem to match up with the observed volumes and overall rates they obtained after our calculations. Attempts to normalize the calculations based on their results were unsuccessful when we proceeded to try extending to the case of iron as catalyst. In the end, we resulted to modelling the capacity after the data given by Tijmensen et al. [111] - a model which produced stable results that were within the range of costs anticipated. This model determines the cost and size of the reactor based on the energy content of the FT liquids produced.

Once the selectivity parameter has been selected, the amounts of the different fractions are obtained from the ASF distribution. The gas fraction is taken as the components with 1-4 carbon atoms, the gasoline fraction that with 5-8 atoms, the

---

[4]Such data include the lifetimes of the catalysts, rate of poisoning and the relative amounts required as a function of the reactor volume

diesel fraction that with 9-18 atoms and the wax fraction taken as everything with more than 20 carbon atoms [5]. Once the respective fractions have been calculated, the mass flow rates of each are calculated as

$$F_i = mr \times W_i \times M_{FT} \tag{8.28}$$

where M_{FT} is the mass flowrate of the feed into the reactor. Note that from the FT reaction, the total hydrocarbons produced from the syngas feed is a just a fraction of the mass of the feed as the rest of the mass is water. This ratio is the variable mr in equation 8.28 and can be determined to be approximately 14/32 from the overall equation.

Note that separation costs are not factored into this model separately although we include cracking costs. All the wax is sent to the cracker for upgrading and the products from cracking are modelled principally as diesel while the gasoline fraction is sold at the price of gasoline.

Power Generation   Included in the model is the possibility of generating electricity along side liquid fuels by combustion of fuel gases generated within the process. These gases can come from one of three places

- Methane generated as a by product in the gasification process

- Actual diversion of some of the syngas generated in the gasifier for electricity generation

- Combustion of the gaseous component of the FT products

Power generation from all of these sources is calculated using the following equation

$$P_j = m_j \times HHV_j \times \eta \tag{8.29}$$

where j denotes the source (one of the three possibilities), m_j the flowrate of the gas, HHV_j the higher heating value of the gas and η the efficiency of converting thermal energy to electrical energy. The choice to divert a fraction of the syngas generated is a decision variable included in the model. An efficiency factor of 0.8 and 0.7 is included in the calculation of the energy content of the methane from the gasification process (where it is not reformed to hydrogen) and the off-gases from the FT process to account for impurities and other inefficiencies.

8.2.2   Economic model for the Biomass to Liquids plant

The economic model for this plant comprises both the capital costs and the running costs. The net present value is used as the profitability criterion and eventually balances both components. We first develop the capital costs and then the operating costs follow

---

[5]http://www.zero.no/transport/biodrivstoff/fischer-tropsch-reactor-fed-by-syngas (accessed January 10, 2013). Note that these are rough proportions and only used to give first order estimates

Capital costs Hamelinck et al. [43] develop a fairly detailed economic model with parameters for the biomass to liquids process and we utilize the model and the data they give in developing our economic model. They use the following general structure for developing the capital cost of all the equipment in the process

$$C_2 = C_1 \times (\frac{S_2}{S_1})^a \qquad (8.30)$$

where C_2 is the capital cost of the equipment, C_1 is the capital cost of a reference size of the equipment, S_2 is the capacity of the equipment to be costed and S_1 is the capacity of the reference, S_2 the capacity of the equipment to be cost and a the exponent. Where insufficient data exists to determine the value of the exponent, the value of 0.6 is usually taken as a good starting point [105] but Hamelinck et al. also provide data for the exponents as well. We present the data from their paper but include modifications where their equipment options do not cover all the possibilities. Some of the variables include maximum sizes which we didn't take into consideration.

Operating costs The biggest elements of the operating costs of the plant are the costs of biomass which we calculate as

$$Feedcost = Feedmass \times Feedprice \qquad (8.31)$$

We assume a feed price of about \$70/tonne although as capacity increases, we include a factor that allows the costs to increase since biomass may have to be sourced from much further out, adding to the transportation costs. In addition to the feed cost, we assume that other overheads, administrative and plant maintenance costs can be approximated at about 10% of the capital cost of the plant [111].

Revenues The principal sources of revenue for the plant come from the sale of FT liquids and revenue from the sale of electricity. Revenues from the sale of FT liquids are obtained by calculating the number of gallons of gasoline and diesel produced per year multiplied by the price of each product i.e

$$Revenues = \sum gallons_{fuel} \times price_{fuel} \qquad (8.32)$$

The revenues for electricity are obtained by determined the number of megawatt-hours of electricity the plant produces per year and then multiplying by the price of electricity per megawatt-hour which, in the deterministic case we assumed to be \$70 per megawatt-hour. We assumed the plant to be running for 330 days in a year.

With all these numbers we can then determine the net present value of the plant as follows

$$NPV_{BTL} = f_d \times (Revenues - Operatingcosts) - Capitalcost \qquad (8.33)$$

where the factor f_d captures the discounting that occurs in determining net present value and was defined in Chapter 7.

8.3 Model Validation

To check that our implementation of the published models were correct, we compared our outputs against the numbers that the papers obtained - while trying to make sure we were able to replicate the model structure assumed by the papers when producing the numbers. The results obtained by Tijmensen et al. [111] are given in Table **??**. The cost break down is for a plant with a plant of capacity 400 MW HHV input. Under certain assumptions[6], we determine that this correlates to a plant processing about 2500 metric tonnes of wet biomass per day. We use information from their process to determine the appropriate process structure for comparison. Syngas quality suggests the use of an entrained flow gasifier and they explicitly mention all concepts requiring the use of a cobalt catalyst - further emphasized by the need for a shift reactor. There is also no reforming of the syngas and we asume a once-through process becaue of the co-production of electricity. Though they state a once-through process, they use an efficiency that corresponds to our twice through efficiency and so we base the model on that. Their model neglects the cost of upgrading diesel so in the comparison, we do the same.

On an overall basis, our model compares quite well as our model yields an overall cost of $328 million while theirs yields a cost of about $340 million which is only a difference of about 4%. Furthermore, We compared the outputs of the model (total capital cost, production etc) with some another set of numbers from the literature. The principal reference was the paper by Boerrigter [17][7].

To check the numbers, first we determine the specific Total Capital Investment of the plant using the following equation he gives:

$$TCI_s = 52000 \times (\frac{34000}{scaleX})^y \tag{8.34}$$

where y is an appropriately chosen exponent between from 0.3 to 0.5 depending on the scale of the plant[8]. This equation gives the specific total capital investment (TCI_s) of the plant i.e. the capital cost per barrel of FT liquid it produces per day. To then get the TCI, we simply multiply the specific TCI by the barrels of FT liquids produced per day.

Since the basis of our plant capacity is the amount of biomass processed per day, we need to convert his capacity basis to ours. From the paper, the FT fuel output seems to be based on an overall (energetic) conversion of biomass to FT liquids of 56%. At this efficiency, a plant that processes about 2500 tons of wet biomass per

---

[6]To back calculate the amount of biomass the plant processes per day (the basis of our plant), we assume a moisture and non-biomass content of about 30% and a higher heating value for biomass of about 20 MJ/kg

[7]This paper seemed apropos because it doesn't reference the paper by Hamelinck et al. thus allowing us to reasonably conclude that the papers are independent An examination of the references of both papers also showed independence - no mutual papers cited - although an extensive comparison of collaborator names within the papers wasn't carried out

[8]He notes that at scales less than 20,000 barrels per day, the exponent should be closer to 0.4 and even closer to 0.5 when plant production drops to 5000 barrels of FT liquids per day. We will use a value of 0.4 for the rest of the analysis

Parameter	Unit	Value
Cost of biomass	\$ per wet tonne	70
Cost of electricity	\$ per megawatt	70
Price of diesel	\$ per gallon	4
Price of Gasoline	\$ per gallon	3.5
Moisture content of wood	kg/kg	0.3
Efficiency of electricity generation	MJ/MJ	0.55
Plant (economic) lifetime	Years	15
Discount rate	%	10

Table 8.2: Nominal values of the parameters used in running the BTL model

day produces 2100 barrels per day of FT liquids[9].

Using his cost correlation in Equation 8.34, the capital cost of a plant with this capacity is about \$320 million dollars. Our model, assuming a CFB flow gasifier with an iron catalyzed FT process and a biosyngas-FT conversion of 80% (double pass type efficiency)and an α of 0.8 gives a plant with a capital cost of \$317 million dollars, a value that is about 1% smaller than the Boerrigter estimate - a remarkably close agreement for the difference in approach for arriving at the calculations. While there is the possibility that there is a mutual model that both models are derived from, at the very least it gives us the confidence that our model was implemented correctly and we can proceed with confidence of agreement with literature.

8.4 Pathway selection in the deterministic case

With the model that we have built, we can now attempt to optimize for the process pathway based on the NPV metric. To run the deterministic case - i.e. the case where all the parameters are known with certainty - we need a set of base values with which to determine the outputs. These values are highlighted in Table 8.2. We also give a list of the variables and the range of values that they assume in Table 8.3

A process design consists of a value selected for each of the variables presented in Table 8.3. The continuous variables are discretized in order to be able to do this. Using the parameters, we run the entire range of decisions through the model and obtain the NPV for each decision. We can then sort the decisions based on the NPVs to determine the decision that gives the maximum NPV. The histogram in Figure 8-4 shows the range of net present values obtained using all the decisions[10].

The average NPV characterizing this range of decisions is about -\$240 million

---

[9]This assumes that the FT liquids are largely diesel with an energetic content of about 36 MJ per liter. Assuming a conversion of 1 barrel to 160 liters, an energy content of wood (dry basis) of 20 MJ per kilogram and a moisture content of 30%, we arrive at the final figure

[10]In this section we will be using a number of metrics that are often associated with distributions to describe the outcome of the deterministic decisions. This is not because the outcomes are random, but rather that the range of decision options is so large - there are about 21000 possible designs - that the best discussions are obtained in some form of aggregation

Variable	Type	Range
Biomass feed	Continuous	500 - 5000 tons per day
Gasifier choice	Binary	Entrained Flow/Fluidized bed
FT catalyst choice	Binary	Cobalt/Iron
Inclusion of reformer	Binary	Yes or no
FT selectivity	Binary, Continuous	0.6-0.8 (iron); 0.8 - 0.9 (cobalt)
One or two pass efficiency	Binary	1 or 2
Syngas diverted for electricity	Fraction	0 - 0.9

Table 8.3: Decisions in the BTL process and their range of values
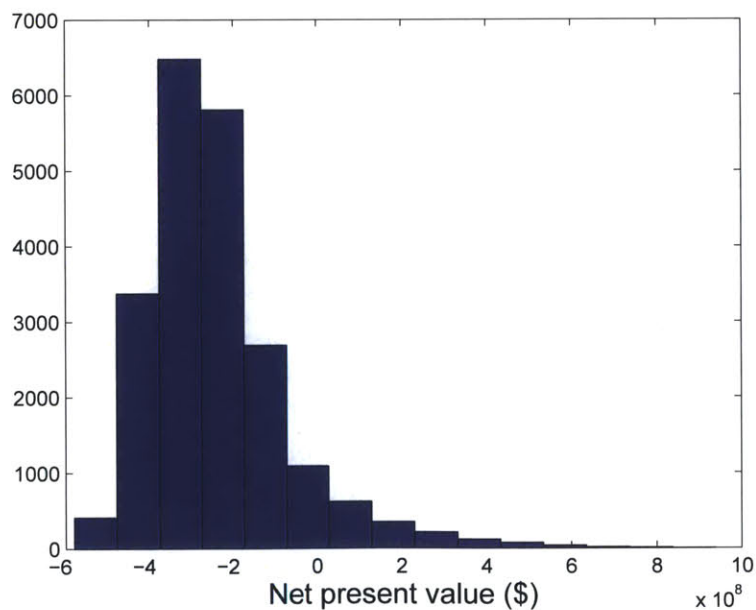
Figure 8-4: Histogram of the net present values for all the designs

Variable	Best NPV decision	Worst NPV decision
Biomass processed	5000	5000
Gasifier type	Entrained Flow	Fluidized bed
Syngas fraction for electricity	0	1
FT Catalyst	Cobalt	-
FT selectivity	0.0.88 (cobalt maximum)	-
Autothermal reformer	Yes	-
One or 2 passes through FT	2 passes	-
Net present value	$940 million	-$580 million

Table 8.4: Decision variable for the best and worst NPV outcomes for the BTL process

Variable	$X_i - X$	Variable	$X_i - X$
EF Gasifier	$44 million	FB Gasifier	-$44 million
Co catalyst	$70 million	Fe catalyst	- $70 million
Reformer	$ 1.9 million	No reformer	- $1.9 million
1 pass	-$10 million	2 passes	$10 million

Table 8.5: Effects of the different binary variables

while the median NPV is about -$270 million - highlighting the skewness in the range that can be seen in Figure 8-4. Overall, the realizable net present values range from -$580 million to a positive of $940 million. Both extremes occur at the maximum processing capacity of 5000 tons per day. Table 8.4 below gives the decision variables that yield both best and worst designs. The flowsheets corresponding to these two designs are illustrated in Figure 8-5 and Figure 8-6.

From the table and the diagram, we see that the critical variables that affect the process are the type of gasifier and the choice to produce electricity as all the other variables take on the same values for both decisions. We however won't just rely on the extreme decisions for the effect of the decisions as we analyze the effect of each variable subsequently. We take the average of each variable held at its value as a measure of its overall effect on the process. The averages are reported in Table 8.5 and the continuous variables are illustrated in Figures 8-7 to Figure 8-10

From the table and the figures, a number of things are clear. First we see that capacity alone is not sufficient to guarantee great value although the extremes do better on average than the middle. This, at first seems counter-intuitive because incresing economies of scale ought to mean that in general, as the capacity increases, the designs should get better. However, close examination reveals that with a poor structural design, say the worst case structure in Table 8.4, larger capacity just amplifies the bad outcome in the same way it amplifies a good outcome. Thus the effect of capacity is dependent on the process structure and isn't necessarily always positive.

The gasifier choice has a significant positive effect on the net present value because it leads to a higher conversion of the carbon in the biomass to the useful syngas that can either be used for electricity generation or for fuel production.
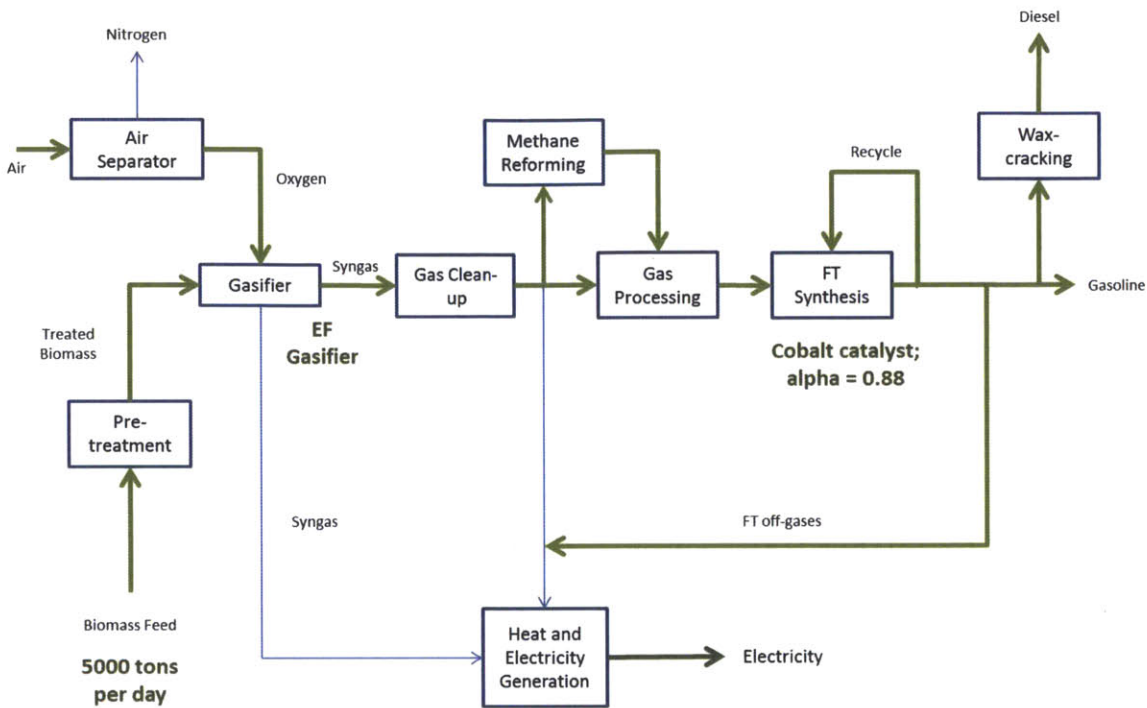
Figure 8-5: BTL flowsheet for design with the maximum (deterministic) net present value
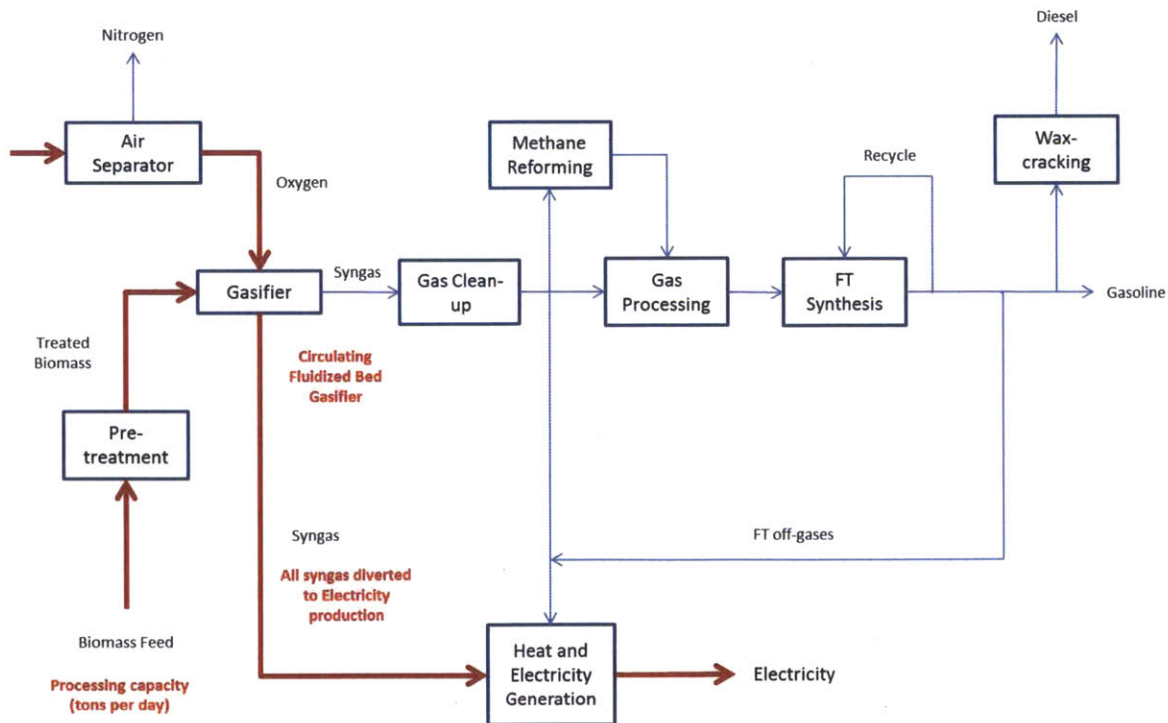
Figure 8-6: BTL flowsheet for design with the minimum (deterministic) net present value
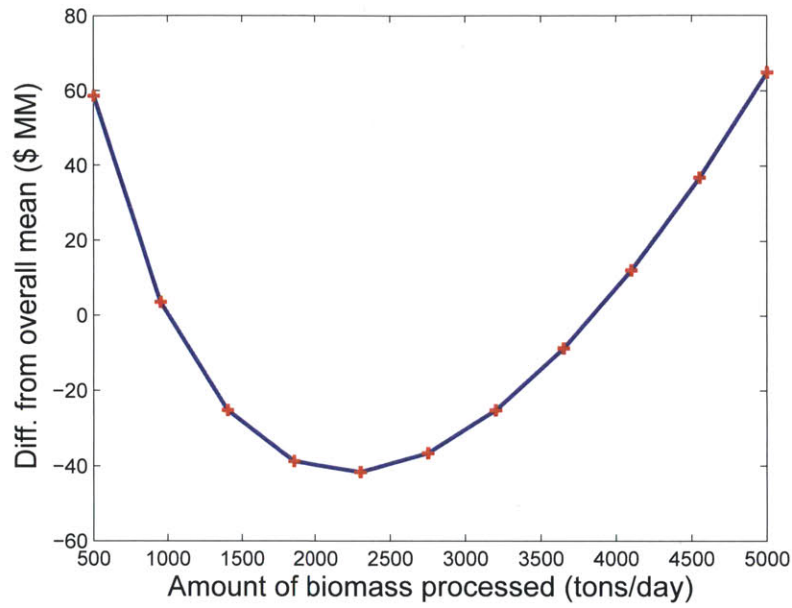
Figure 8-7: Improvement over the overall mean as a function of inlet processed biomass
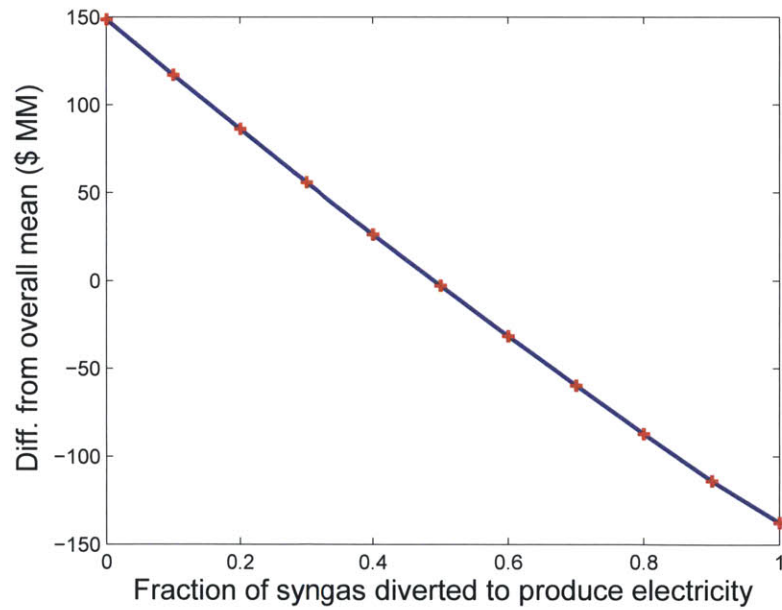
Figure 8-8: Improvement over the overall mean as a function of the fraction of syngas diverted to electricity production
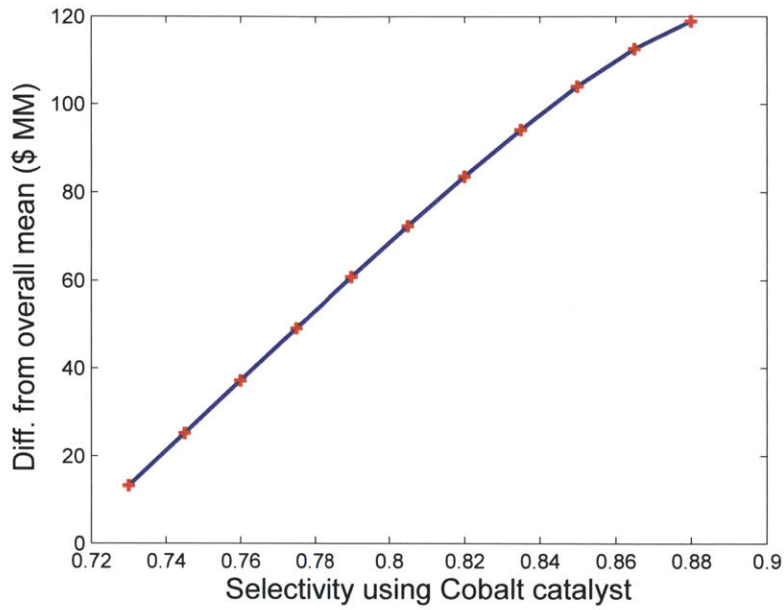
Figure 8-9: Improvement over the overall mean as a function of the FT selectivity parameter for cobalt
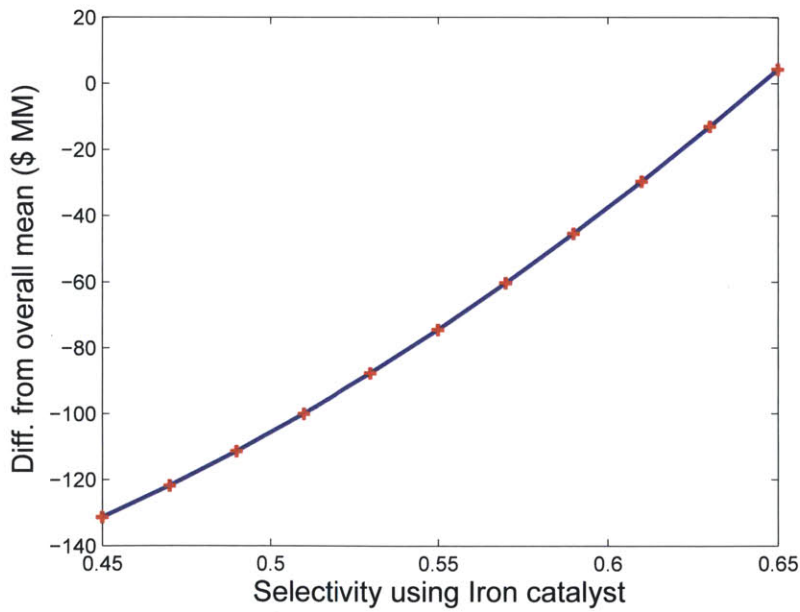
Figure 8-10: Improvement over the overall mean as a function of the FT selectivity parameter for cobalt

172

Figure 8-8 shows a very interesting trend - diverting syngas to the production of electricity instead of liquid gives a linear relationship between the deviation from the overall mean and the fraction diverted. The average improvement in net present value in diverting no syngas and diverting all the syngas produced to electricity generation is about \$300 million - and this change occurs in a linear increment as the fraction of syngas diverted is increased.

The linear relationship can be understood by considering the fact that the overall conversion efficiency (energy input to energy output) to either fuel or syngas stays about constant. As a result, the energy in a unit of processed biomass can either be converted to a certain fraction of electrical energy or to another fraction of fossil fuel energy. However, the prices of these units of energy are different. For the deterministic model, we assumed a price of \$70 per MWh of electricity, which is about \$20 per MJ of electricity[11]. Assuming, for simplicity, that all of our fossil fuel energy was converted to diesel fuel which we priced at \$4 per gallon, this translates to a selling price of about \$30 per MJ of fossil fuel energy[12].

Thus every megajoule of syngas energy diverted to the production of electricity instead of liquid fuels (assuming relatively equal overall energetic efficiencies) gives \$10 less in returns and it is this average constancy that is reflected in the slope. The price differential in the outputs also reflects why in Table 8.5 any variable that leads to an increase in the amount of fuel produced over electricity - introduction of a reformer, recycling unconverted syngas through the reactor a second time to increase overall efficiency - leads to an improvement of the NPV.

This also explains Figures 8-9 and 8-10 that show the increase in value that comes with higher FT selectivities. At higher selectivities, more of the syngas going through the FT reactor is converted to liquid fuel versus gaseous products (see Figure 8-2), resulting in less electricity produced (since the gaseous products are burned to produce electricity). However, the selectivity curves are not as linear as the electricity curves because the tradeoff between gaseous products and liquid fuels is not exactly a linear function of the selectivity as the Anderson-Flory-Schultz distribution shows.

The relationship of the choice of the gasifier is not readily intuitive. Entrained Flow gasification gives higher overall conversion but also has a higher cost of feed preparation. The positive relationship between the choice of this gasifier and the average performance of a design with this choice relative to the overall average illustrates that the cost of pre-treatment on average is, over the lifetime of the plant, less than the average value of the extra returns obtained from increased overall efficiency.

And finally, the positive relationship of the cobalt catalyst is essentially tied to the fact that it gives higher FT selectivities than the iron catalyst - and the returns from this higher selectivity in terms of more liquid fuel produced as opposed to more electricity generated outweighs the extra cost of the catalyst as factored in by the model.

Given these relationships, we can now understand the variable choice that pro-

---

[11]1 MWh = 3.6 MJ

[12]Energy content of diesel is about 36.4 MJ per liter and there are 3.78 liters of diesel in a gallon of fuel

Variable	Nominal value	Range	Unit
Biomass price	70	35 - 140	$ per wet tonne
Electricity price	70	30 - 100	$ per MWh
Gasoline price	3.5	1.75 - 6	$ per gallon
Diesel price	4	2 - 6.5	$ per gallon
Moisture content	0.3	0.2 - 0.4	-
Electricity conversion efficiency	0.55	0.5 - 0.6	-
Overall cost coefficient	1	0.7 - 1.3	-
Discount rate	10	7 - 13	-
EF Gasifier syngas conversion	0.8	0.75 - 0.85	-
EF Gasifier carbon conversion	4	3.5 - 4.5	-
CFB Gasifier syngas conversion	0.65	0.55 - 0.75	-
CFB Gasifier carbon conversion			-
FT 1 pass efficiency	0.6	0.5 - 0.7	-
FT 2 pass efficiency	0.8	0.7 - 0.9	-

Table 8.6: List of parameters in the BTL model that will be varied for the uncertainty analysis

duces the best NPV. As to structure, we want to maximimize the amount of liquid fuel produced over electricity - leading to a choice of an entrained flow gasifier, a cobalt catalyst, no syngas diversion for electricity production, inclusion of autothermal reformer, using the highest possible selectivity available and recycle through FT reactor for higher conversion efficiency. Once these are done, we can then take advantage of economies of scale by building a structure to maximize throughput/capacity.

8.5 Pathway selection under uncertainty

The analysis of the design procedure under uncertainty is similar to those done in the previous case-studies. First we select a list of key parameters that we will be varying in the design. These are listed in Table 8.6 together with the range of uncertainties. For the analysis, we will be assuming that all the distributions are uniformly distributed with their distributions often centered around the nominal parameters used in the deterministic case.

Once the uncertain variables are implemented, we run the analysis framework to get the decisions present in the stochastic dominance efficient sets in addition to those in the mean-variance efficient set. Table ?? shows the decisions present in the single objectives while Table 8.8 shows the number of decisions in the Stochastic Dominance and Mean-Variance efficient set.

A few things to note about the distributions. First, except for the gasoline and diesel prices, all the uncertain parameter were determined to be independent. This is a simple assumption - but one that can easily be changed in the model as more information on the correlations among the parameters are known. Secondly, one uncertain variable was used to simulate the diesel and gasoline prices on the assumption

Variable	Expected Value	Worst-Case
Biomass processed	5000	500
Gasifier type	Entrained Flow	Entrained Flow
Syngas fraction for electricity	0	0
FT Catalyst	Cobalt	Cobalt
FT selectivity	0.73	0.88
Autothermal reformer	Yes	Yes
One or 2 passes through FT	2 passes	2 passes

Table 8.7

Objective	Number of Decisions
FSD	11
SSD	11
Mean-Variance	122

Table 8.8: Number of decisions in the Stochastic Dominance and Mean-Variance efficient set

that these prices in general track each other closely in the world market.

In addition, the efficient sets generated in Table 8.8 were obtained using a sample size of 1000 simulations for the uncertain variables. This is because this was the largest thousand number of simulations that the MATLAB environment we used[13] was able to handle before running out of memory. Our compromise was to run this simulation (and a few others with sample sizes smaller) and examine the breakdown of the decisions in the efficient sets to see if any variables were uniformly excluded. Once identified, these were removed from the set of choices and bigger sample sizes were run with the remaining decisions so that better representation of the decisions in the efficient set could be obtained.

Table 8.8 displays the sizes of the efficient sets when all the decision variables were made to vary. The FSD and SSD sets are identical and contain the same 11 decisions. Essentially these correspond to the discretized capacity dimension - with all other variables remaining the same, with the essential structure given in Table 8.9.

The break down of the decisions that comprise the mean-variance efficient set are given in Table 8.10 and in Figure 8-11.

Figure 8-12 displays the cumulative distributions for the decisions - indexed by their biomass capacities. For greater contrast the cumulative distributions for the largest capacity and the smallest capacity are also plotted and shown in Figure 8-13 These, by the way, also correspond to the expected value decison and the worst case decision respectively.

As mentioned earlier, we also ran the analysis, eliminating the electricity variable as well as the choice of using one or two passes for the FT reactor in order to in-

---

[13]Computer was a DELL precision T7400 model with a Intel Xeon processors, Windows XP operating system and 4GB of RAM.

175
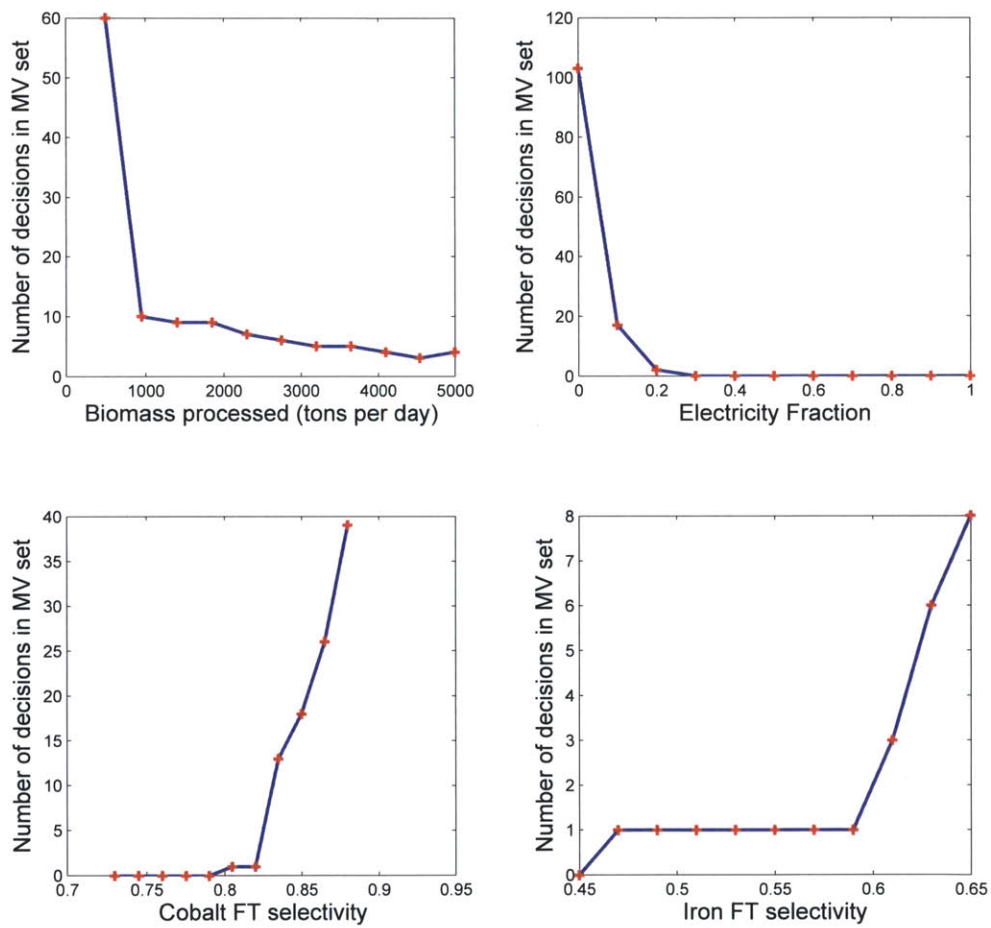
Figure 8-11: Breakdown of continuous variables in the Mean-variance efficient set

Variable	Value
Gasifier type	Entrained Flow
Syngas fraction for electricity	0
FT Catalyst	Cobalt
FT selectivity	0.88 (cobalt minimum)
Autothermal reformer	Yes
One or 2 passes through FT	2 passes

Table 8.9: Structure of BTL process represented in the FSD and SSD set, where all the variables were varied

Variable	Number in set	Variable	Number in set
EF Gasifier	88	Autothermal reformer	88
CFB Gasifier	34	No reformer	34
Co catalyst	98	One pass through	29
Fe catalyst	24	Two pass through	93

Table 8.10: Breakdown of process structure variables in the Mean-variance efficient set
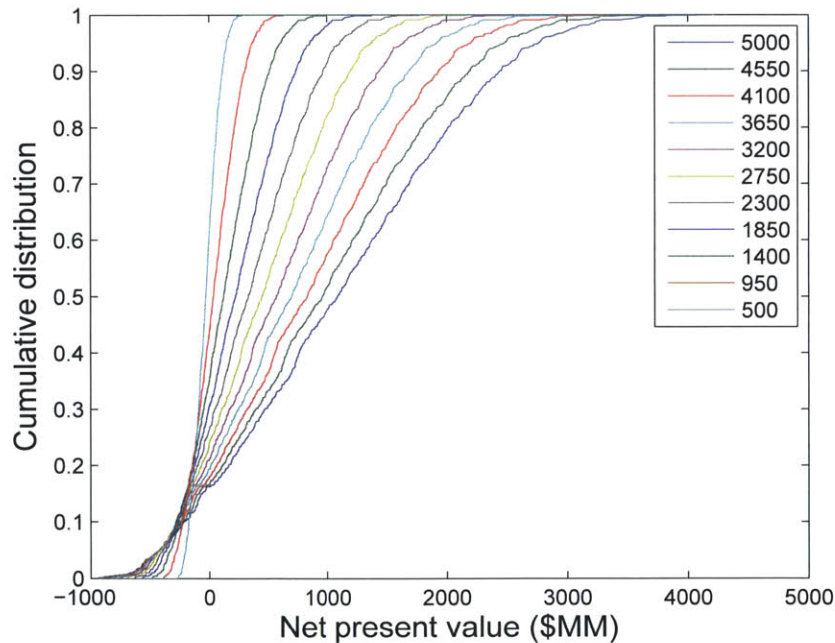
Figure 8-12: Cumulative distribution for all the decisions in the SSD efficient set

177
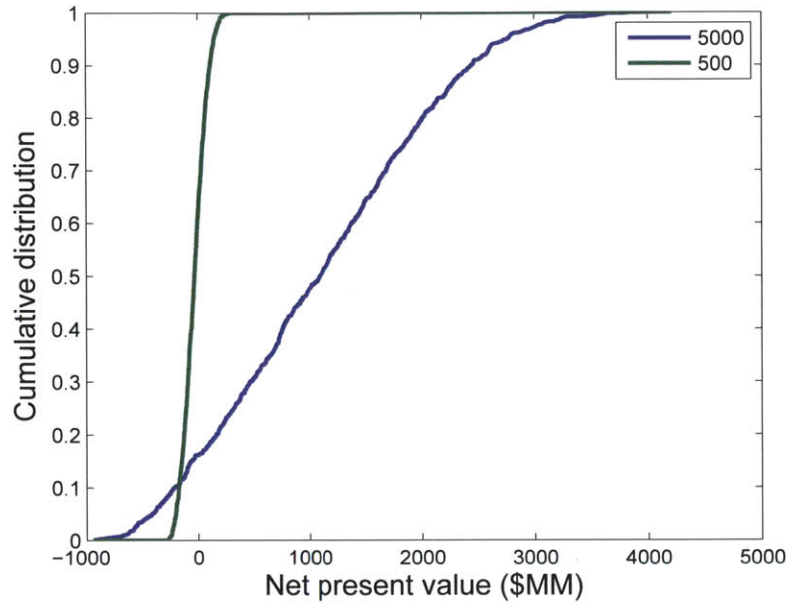
Figure 8-13: Cumulative distribution for two decisions in the SSD set corresponding to the expected value and worst case decision

Objective	Number of Decisions
FSD	11
SSD	11
Mean-Variance	111

Table 8.11: Number of decisions in the Stochastic Dominance and Mean-Variance efficient set with reduced decision variables

crease the number of simulations. Eliminating both variables reduces the decision space by a factor of 22 and enabling us to run 10,000 simulations of the remaining decision variables. Table 8.11 shows the sizes of the efficiency sets of the different objectives. As we see, the size of the uncertain sets for both the FSD and SSD sets remain unchanged while the size of the mean-variance efficient set reduces by 11. The breakdown of the new MV sets are shown in Table 8.12 and Figure 8-14.

The results presented here however mask one very important variable - the uncertainty in the supply of the feed. So far, we have assumed that once a capacity has been determined, there is corresponding certainty in the amount of biomass generated. In order to see how the biomass uncertainty affected this, we amended the models to include this uncertainty. To do this, we assumed that during the design, a nominal yield of 15 wet tons/hectare was assumed for the farm size that will produce the required inlet capacity. Once this was determined, the overall plant cost was fixed. Next, we assumed that for the real plant, the farm size stayed constant but the yield from it varied according to a uniform distribution that ranged from 7 tons per hectare to about 23 tons per hectare. This varying yield was then used to

178

Variable	Number in set	Variable	Number in set
EF Gasifier	74	Autothermal reformer	78
CFB Gasifier	37	No reformer	33
Co catalyst	85	One pass through	0
Fe catalyst	26	Two pass through	111

Table 8.12: Breakdown of process structure variables in the Mean-variance efficient set
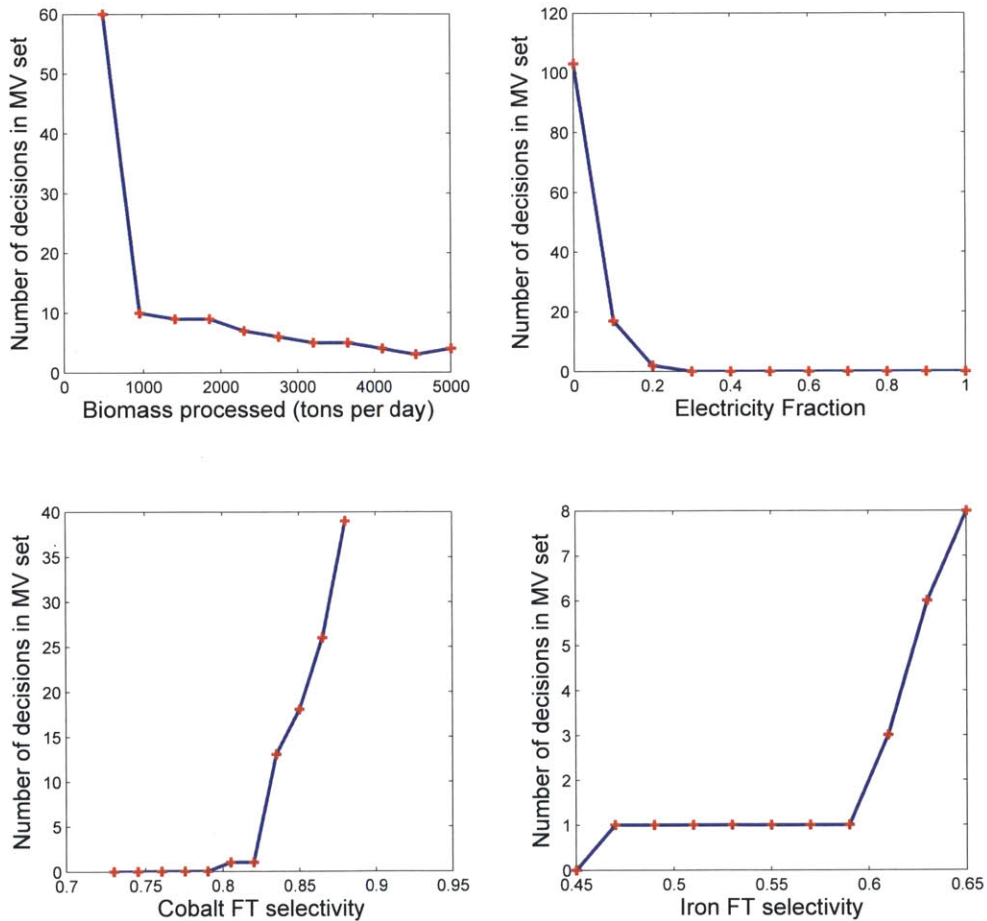
Figure 8-14: Breakdown of continuous variables in the Mean-variance efficient set

Objective	Number of Decisions
FSD	11
SSD	11
Mean-Variance	175

Table 8.13: Number of decisions in the Stochastic Dominance and Mean-Variance efficient set with uncertainty in feed
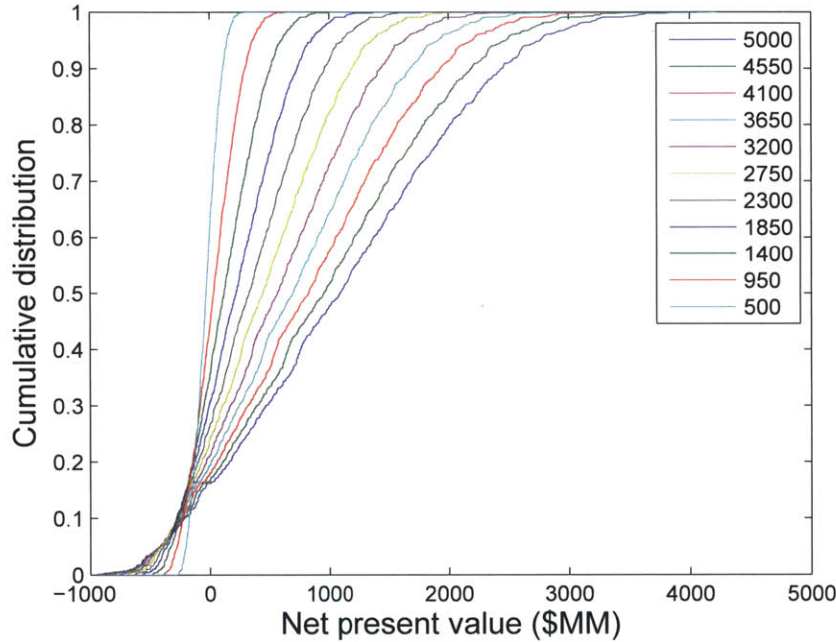
Figure 8-15: Cumulative distribution for all the decisions in the SSD efficient set with uncertainty in feed supply

calculate the actual capacity the plant will process and it was this cost that was used in determining the operating and miscellaneous costs of the plant and not the overall deterministic cost.

The resulting sizes of the different efficient sets obtained are shown in Table 8.13. The FSD and SSD sets remain teh same but the composition of the mean-variance set changes a bit as shown in Table 8.14 and Figure 8-17. The cumulative distributions of all the decisions in the SSD efficient set are shown in Figure 8-15 and the expected value and worst-case decision distributions are shown in Figure 8-16.

8.6   From efficient sets to decisions

Like the previous case studies, we examined different ways of further pruning the efficient set.

180
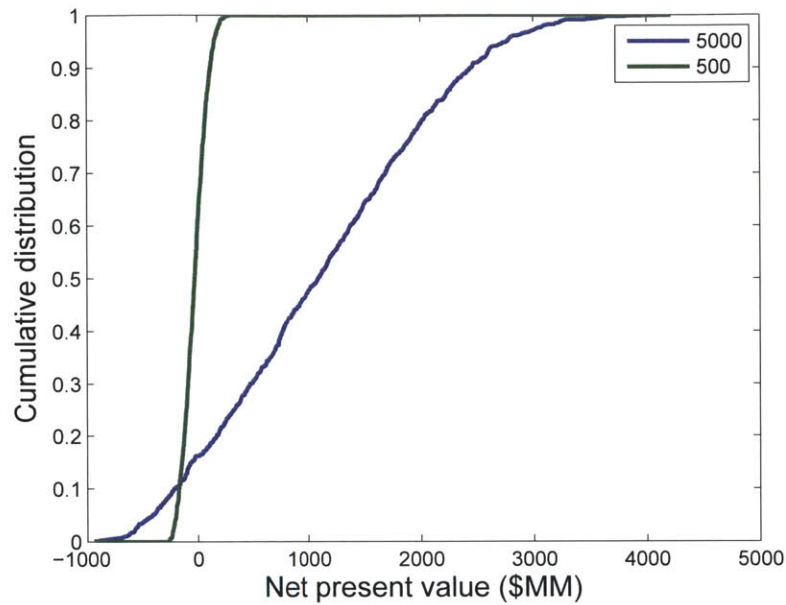
Figure 8-16: Cumulative distribution for two decisions in the SSD set corresponding to the expected value and worst case decision with uncertainty in feed supply

Variable	Number in set	Variable	Number in set
EF Gasifier	74	Autothermal reformer	78
CFB Gasifier	37	No reformer	33
Co catalyst	85	One pass through	0
Fe catalyst	26	Two pass through	111

Table 8.14: Breakdown of process structure variables in the Mean-variance efficient set when uncertainty in feed is introduced
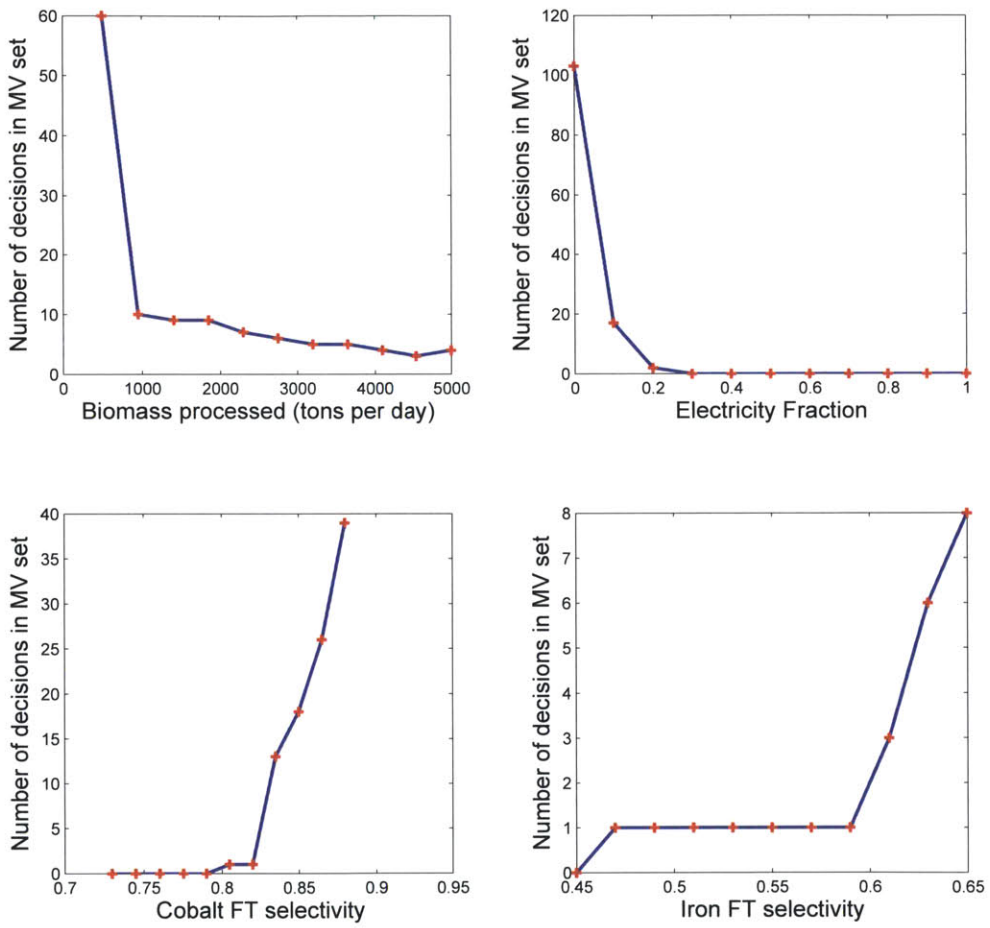
Figure 8-17: Breakdown of continuous variables in the Mean-variance efficient set
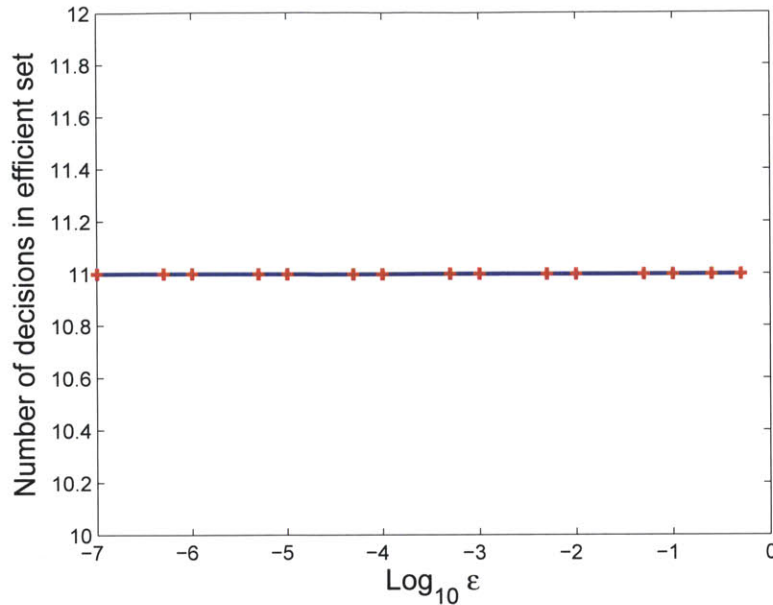
Figure 8-18: Using ASSD to screen the BTL SSD efficient set yields no change in the number of decisions present in them

Almost Stochastic Dominance   Running the ASD algorithm on the efficient set gives the plot shown in Figure 8-18. There is no reduction observed in the size of the efficient set because the intersections between the (integrals of) the distributions occur at places with significant probabilities of occurence. As a result, the value of ϵ required to observe a significant reduction in the size of the efficient set is much larger and may now be very dependent on the particular decision maker.

The use of alternative metrics   As was done in the previous case-study, we used the profitability index as an alternative metric in order to reduce the size of the efficient set. The decisions present in the SSD efficient set are similar to the decisions present in the efficient set using the net present value metric. However, the major difference is that instead of the cobalt catalyst in use in the FT reactor, the iron catalys is largely used. This would have implied no common decision in the set - but there was also the decision to build the 5000 tons/day option also using the cobalt catalyst. The structure is outlined in Table 8.15.

Variable	Value	Variable	Value
Gasifier type	Entrained Flow	Gasifier type	Entrained Flow
Syngas fraction for electricity	0	Syngas fraction for electricity	0
FT Catalyst	Cobalt	FT Catalyst	Cobalt
FT selectivity	0.6 (iron maximum)	FT selectivity	0.88 (cobalt maximum)
Autothermal reformer	Yes	Autothermal reformer	Yes
One or 2 passes through FT	2 passes	One or 2 passes through FT	2 passes
Mass of biomass processed	2750 - 5000 tons/day	Mass of biomass processed	5000 tons/day

Table 8.15: Structure of BTL process represented in the SSD set using the profitability index

Parameter	Sensitivity	Parameter	Sensitivity
Biomass Price	0.00	EF carbon efficiency	0.00
Gasoline Price	0.00	CFB syngas efficieny	0.00
Diesel Price	0.53	CFB carbon efficiency	0.00
Electricity price	0.04	Discount rate	0.02
Moisture content	0.00	FT 1 pass efficency	0.00
Cost multiplier	0.01	FT 2 pass efficiency	0.02
Electricity conversion efficiency	0.00	Biomass yield	0.42
EF syngas efficiency	0.00	-	-

Table 8.16: Average, total sensitivity indices for the different parameters and for the decisions in the SSD efficient set

Thus the common decision in both efficient sets is to use the structure obtained from the net present value metric (using cobalt instead of iron), building a plant with a capacity of 5000 tons per day.

Sensitivity analysis In further examining the SSD efficient set, it is first useful to determine which input uncertainties are driving the uncertainties in the outputs as a means to figuring out possible next steps. We thus carry out a total sensitivity analysis (discussed in Chapter 3) on all the decisions in the set using all the parameters in the study. The average values of the sensitivities across the decisions are given in Table 8.16

This sensitivities were calculatd using 10,000 samples - enough for convergence. As can be seen, the two most significant uncertainties - driving essentially all the variation in the output are the diesel price and the biomass yield. The petrol price is absent because in our model, we have assumed it to be perfectly correlated with the diesel price. The break down of the sensitivities for both these parameters are shown in Figure 8-19. It should be emphasized that this sensitivities were calculated using the decisions in the SSD efficient set, hence the seeming irrelevance of many of the other parameters - as most of the variables that depend on these parameters were not varied.

The sensitivity of the biomass yield is seen to grow with biomass feed capacity while that of the diesel price falls almost the same way. Both of these together account for more than 90% of the variation in the net present value of the designs. Thus, any attempts focused on reducing the uncertainty in the outcome will do especially well by targeting those two variables. Biomass yield may be reduced for by getting another source of carbon to balance low yields - for example a combined input that allows for natural gas processing, instead of relying solely on biomass. This profitability of this approach will have to be considered in conjunction with whatever potential tax credits and benefits (sustainability related) that might have accrued to building a purely biomass to liquids plant.

The variation in energy prices can be mitigated by negotiating long term buy-sell contracts where possible, the likelihood of rising energy prices due to increasing global
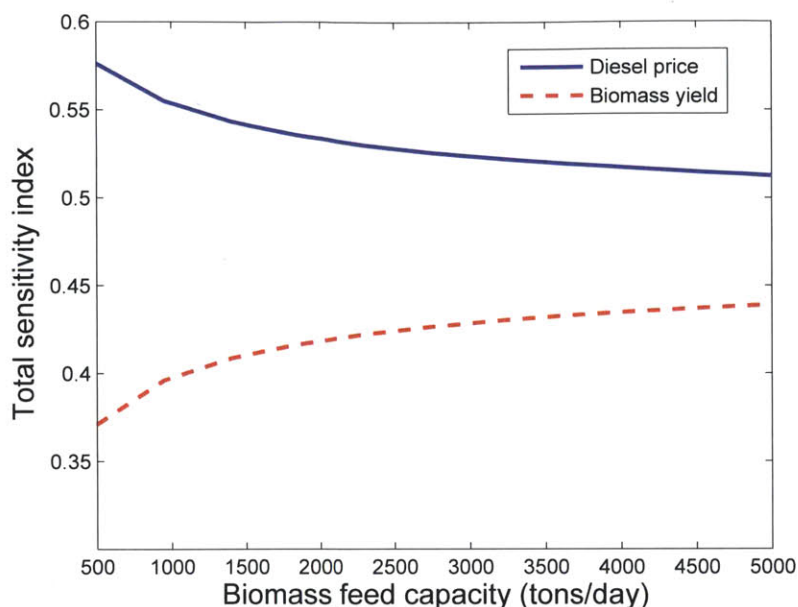
Figure 8-19: Variation of total sensitivity indices of the diesel price and biomass with biomass feed capacity for designs in the SSD efficient set

demand making this a tenable proposition to prospective buyers.

8.7 Conclusions

This case-study focused on the design of a facility that converts biomass to liquid fuels and electricity. A range of decision variables were incorporated into the problem - with the goal to select both a structure and a processing capacity for the facility. Our analysis in both the deterministic and the uncertain case showed that the best structural design for the facility was one that used an entrained flow gasifier, a cobalt catalyst for the FT process and maximized the production of fossil fuels over electricity by reforming methane generated in gasifier and recycling the products of the FT reactor. This overall process structure is identical to that in Figure 8-5. This is reproduced below in Figure 8-20. The difference in the two figures is simply that the capacity choice is not fixed for the uncertain case as discussed below.

In the deterministic case, the best design was the maximum possible capacity of 5000 tons per day as the combination of economies of scale and production output contributed to a very positive net present value. In the uncertain case however, the capacity choice was the one and only design variable that varied in the set - with the full range of capacities present. From the distributions, we could see that even though the NPV grew with capacity, the worst possible outcome also worsened and so, as is typical with the SSD analysis, the degree to which the possibility of loss can be tolerated by the decision maker will ultimately determine the capacity selected.

In terms of the comparison with the different efficient sets, we observed that the

186
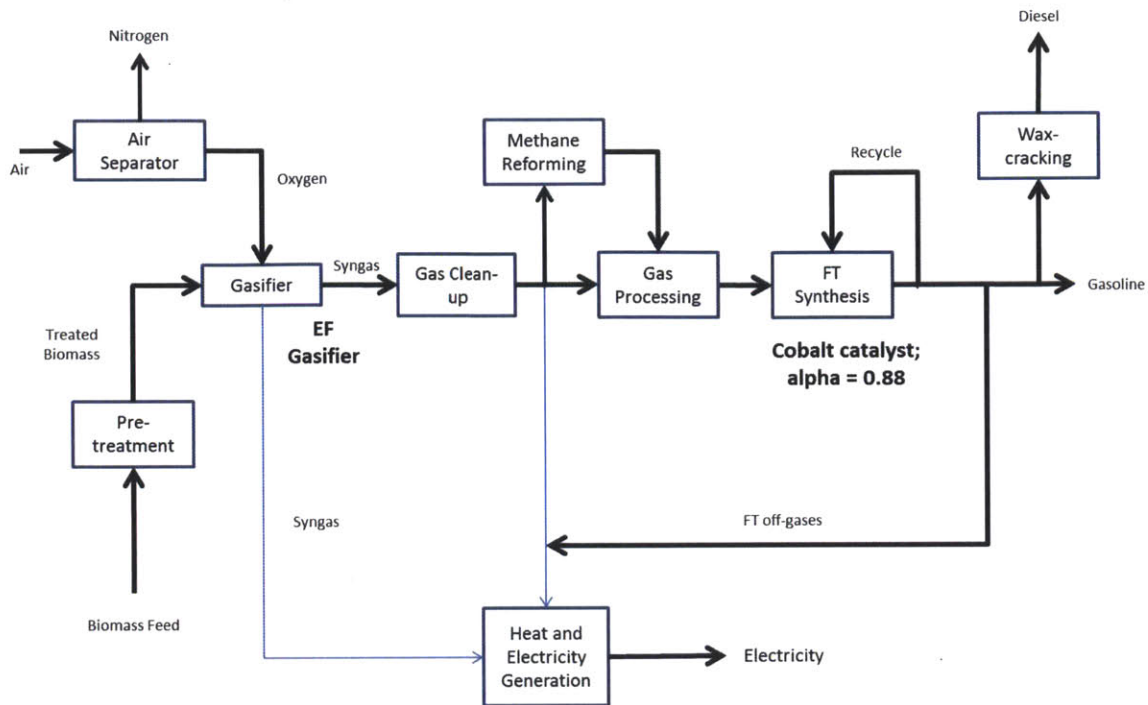
Figure 8-20: Best overall process structure for the conversion of Biomass to liquid fuels and electricity

FSD and SSD sets produced identical decisions but were by far more effective in pruning the initial feasible set than the mean-variance approach was. Both produced an efficient set of 11 decisions as compared with the 175 decisions produced by the mean-variance approach.

The use of almost Stochastic dominance made no difference to the size of the efficient set while the use of the profitability index had the common decision with the NPV metric to produce 5000 tons of biomass per day with the cobalt catalyst.

In analyzing the key parameters driving the SSD set, it was discovered that the energy price and the biomass yield had far more effect than any of the other variables. Thus if uncertainty reduction were part of the next stage of analysis, these two variables would be the most important to focus on.

Thus we conclude from the study that the optimal BTL structure is one that maximizes the production of biofuels over the production of electricity. The structure to be used is that of an entrained flow gasifier, a cobalt catalyst in the FT reactor, an autothermal reformer (for reforming the methane gas produced) and multiple passes through the reactor to maximize the conversion of the syngas to biofuels. While the use of an alternative metric - the profitability index - had the common decision for the capacity to be designed for, the original SSD set and the results of the sensitivity study suggest the gathering of more information about biomass yields and the energy prices in order to better decide the capacity.

187

Chapter 9

Conclusions and Future work

This thesis focused on the application of Stochastic Dominance as a tool for project screening and selection in the presence of uncertainty and demonstrating its performance on a number of case studies. First we give a summary of the background issues behind the need for a new method for project screening, discuss lessons learned from the case studies and highlight the key contributions of this thesis. We then focus on some of the different directions that could be explored in furthering the use of the framework developed in this thesis.

9.1 Summary

In laying the groundwork for the method, we discussed uncertainty and its effects on the decision problem. The presence of uncertainty significantly changes the nature and structure of the design problem and it is important that it is well characterized and properly incorporated into process design. In the absence of uncertainty, the decision-making problem reduces to one of proper selection of scoring metrics and being able to find a commensurable basis for which to evaluate multiple dimensions of the decision problem (example safety, profitability, market share, etc). For projects that can be evaluated on a single dimension, the decision problem becomes an optimization problem and once a mathematical representation for the problem can be formulated, it is all done. Uncertainty makes even the simplest of these problems more complicated by adding the element of risk - defined informally as the failure to achieve the desired objectives - and learning to make good decisions that take this into account is what great firms learn to do in practice. .

This led us to decision theory which provides an axiomatic base for evaluating decisions in the presence of uncertainty. We discussed how, for objective that satisfy the axioms of transitivity, completeness, continuity and independence, expected utility theory provides the correct framework as it is equivalent to all four axioms. It recommends that a utility function be constructed over the outcomes and an expectation (a probability weighted average) be taken and used as a score of each decision. The decision with the highest expected utility should be taken.

The challenge with using expected utility theory is that the utility function has

Study	No. of Feasible Options	Mean-Variance	FSD	SSD
Reactor-Separator	44	25	13	10
Crop-selection	88	22	45	21
Biomass-to-Liquids	21000	175	11	11

Table 9.1: Summary of the results from the different case-studies

to be determined. Elicitation techniques where decision questions are posed to the decision maker and the results used to fit an objective function suffer from a number of errors [122] and it can be difficult to get consistent measurements. However, we pointed out that often we know the preference group of the decision maker, even when we lack information on the specific form of the utility function. The question that naturally arises from that point is whether we can use this knowledge to screen options and eventually make decisions and, if we could, how to go about doing so. That was where Stochastic Dominance was introduced as the framework that allows us to do so. Stochastic Dominance has been shown to be a general method for incorporating risk preferences into the decision-making process. It is consistent with classical decision theory, it makes minimal assumptions of the structure of the utility functions of the decision makers and of the nature of the distributions of the uncertainty and under certain conditions can be shown to be equivalent to the other objectives.

We demonstrated the method on a number of case-studies which we discuss in the subsequent section. Stochastic Dominance acts primarily as a screen for decision as it produces an efficient set of decisions - a group of decisions that the decision maker will find favourable - rather than a specific decision. We therefore investigated different ways for reducing the efficient set and also summarize their results below.

9.2 Lessons from Case Studies

As mentioned earlier, we applied the framework developed to a set of case studies: a reactor-separator design problem, crop selection for a biomass farmer and the design of an industrial facility for the production of liquid fuels from biomass. In each of these case-studies we implemented the Stochastic Dominance framework and compared it primarily with the Mean-Variance approach. Table 9.1 summarized the outcomes of the three studies

As we see, Stochastic dominance was able to reduce the size of the efficient set in all cases - although the degree of reduction varied, with the most significant being the Biomass-to-Liquids case which saw a reduction of more than 99% of the feasible set. In the crop-selection case where there is not much difference between the SSD and mean-variance, we suggest that the nature of outcome distributions may be the issue, as, under conditions discussed in Chapter 5, Stochastic Dominance can give results that are similar to the mean-variance method.

As mentioned, we used utility on the above demonstrations and proceeded to examine different approaches to reducing the efficient set. The Tables 9.2 and 9.3 summarize two of the methods used to do this

Study	SSD	ASSD
Reactor-separator	10	1
Crop-selection	21	10
Biomass-to-Liquids	11	11

Table 9.2: Using ASSD to reduce the Stochastic Dominance efficient set

Study	SSD	With another metric
Reactor-separator	10	3
Crop-selection	21	4
Biomass-to-Liquids	11	1

Table 9.3: Using an alternative metric to reduce the Stochastic Dominance efficient set

From both tables, we see that the use of ASSD can give improvements in the efficient set - although this is case dependent. The use of an alternative-metric also gives a reduction in the efficient set, although, depending on the particular alternative metric, this can have some overall bias in the particular decision left. For example, the initial metric in the crop-selection study was the NPV which takes scale into account. The alternative metric was the use of the benefit-cost ratio (profitability index) and this reduced the efficient set to the smallest farms in initial set. In that case, those results may be less than satisfactory since the size of the farm may not be sufficient to produce enough biomass for the ultimate use of the facility.

In addition to the two methods discussed above, we showed how sensitivity analysis can be used to identify the key input uncertainties that are responsible for most of the uncertainty in the output. This is important because process design and project selection is often iterative and identifying the uncertainties helps to illuminate the areas where the most value can be obtained in future stages focused on uncertainty reduction. Once these uncertainties are identified and eliminated, they can serve to reduce the size of the efficient set. For the case studies the key input uncertainties are summarized in Table 9.4[1].

To demonstrate the effect of reducing the uncertainties on the size of the resulting efficient set, we modelled new uncertain distributions (as though we had perfomed the step of getting more information) for the key parameters in the second case study

---

[1] We should note that the input uncertainties in the first case-study were preselected before the initial analysis

Study	Key uncertainties
Reactor-separator	Reaction-rate, product selling price
Crop-selection	Growth cost and selling price of crops
Biomass-to-Liquids	Energy price, biomass yiels

Table 9.4: Key uncertainties in the different case-studies

and were able to show that it resulted in a reduction from 21 to 8 options in the efficient set.

Overall, we believe the SD framework is a useful approach to screening and ultimately selecting among different projects. There are, however, a number of areas where the method can be improved or extended and we offer some suggestions in the next section

9.3 Future work

This thesis has demonstrated the applicability of Stochastic Dominance as a tool for screening and selecting among projects in the presence of uncertainty. However, much still remains to be done in terms of refining the method and making it more generally applicable. From our present study, a list of promising directions for investigation emerge and they are discussed below

9.3.1 Optimization algorithms for frontier determination

In this work, in evaluating the dominance of certain alternatives within a set of feasible options, we used the pairwise comparison approach that defines traditional Stochastic dominance - albeit modified to reduce the number of alternatives we compared. The pairwise approach is very simple to use but is computationally intensive. Convex Stochastic Dominance used linear programming methods to eliminate dominated alternatives from a set of options although no study has been carried out to determine what the relative computational gain over the traditional approach is (you may need to do that).

In addition, both traditional SD and convex SD still require the discretization of continuously variable options (like volume in the reactor-separator example) so that a finite set of alternatives can be evaluated for pairwise dominance. This is unlike the mean-variance approach that can allow the use of continuously varying alternatives under certain conditions. Under the assumptions of the capital asset pricing model, the mean-variance efficient frontier is known and can be determined without discretizing the alternatives.

The Stochastic Dominance framework developed will thus be vastly improved if ways of implementing the methodology in a continuous fashion were developed as it will allow us take advantage of some of the vast machinery that already exists in solving these kinds of problems.

9.3.2 Decision statistics and significance

Decision making under uncertainty requires the knowledge of the probability distributions that govern the input uncertainties. Often times, these distributions are not known precisely - their parameters have to be estimated from data or inferred using theoretical principles. With elicitation from data comes the potential for estimation errors which can be propagated through the model to the distributions obtained for

the outputs. Since these output distributions are used in the evaluation of dominance, it is possible that dominance may be determined where none exists, or two projects may be deemed to be in the efficient set where, in reality, one decison does dominate another.

Mcfadden [72] discusses and develops some statistical tests for Stochastic Dominance that may be useful in resolving this issue and may potentially be able to help with setting the value of ϵ in the Almost Stochastic Dominance framework by relating it in some way to the errors inherent in the process of estimating the distributions. some of the issues related to thetesting Stochastic Dominance.

9.3.3 Using black box models

All the case-studies in this thesis were developed using equation-based models i.e. we knew the constitutitve equations that related the models input to the outputs and used these, in conjunction with the uncertain distributions in the parameters to determine the output distributions and then subsequently determine dominance. However, nothing in the Stochastic Dominance framework developed requires that we know the functional forms of the models. Rather, we only need some relationship - a black box - that can take in process inputs and produce the corresponding outputs. The process design space in Chemical Engineering uses these kinds of models a lot and it will definitely be a boost to the appeal of the framework we developed to demonstrate its applicability on those kinds of models too.

9.3.4 Dynamic uncertainty and dynamic decision making

Our focus in this work was on static decisions and static uncertainty i.e. the decisions regarding the process configuration and capacity are made at the beginning of the project phase and don't change throughout the project life. The same goes for the uncertainties affecting the process - they have distributions that remain fixed and don't change with time.

Though this can sometimes be an adequate representation of certain decisions and uncertainties (like technical and process uncertainties), it is by no means a complete one as we recognize that this assumption introduces some simplification into the problem. Market uncertainties - like product prices, raw material costs, etc - are dynamic in nature. Their uncertainties follow stochastic processes and can be a bit more involved to model. In addition, companies rarely make decisions and stick by them, come what may. They respond to changing market conditions by altering their decisions - scaling down capacity or shutting down an exisiting facility in bad market conditions and expanding and entering new markets in good ones. This flexibility is a valued asset in decision scenarios and its presence in certain cases can greatly increase the value of an engineering design. Options theory - both in financial and real assets - has largely developed to analyze and correctly value flexibility.

Using Stochastic Dominance to value flexibility may introduce new insights to the field currently not provided by the methods currently in use. In particular it may highlight how the inclusion of preferences affects the valuations and may help

designers make better informed decisions as to what flexibilities are ultimately worth incorporating in the long run. Readers interested in learning more about financial options are encouraged to see the book by Dixit and Pindyck [31] while those more interested in real options in understanding and valuing flexibility in engineering designs as well as more discussion on valuation approaches to flexibility and design are referred to the book by de Neufville and Scholtes [29].

9.4 Thesis Contributions

To the best of our knowledge, we believe this thesis has made the following contributions

1. **Presented Stochastic Dominance as a viable framework for chemical engineering project selection and process design under uncertainty.** Stochastic dominance (in particular Second order SD) is little known and not in much use in the Chemical Engineering field as a decisin making framework. We have, by this work, introduced the concept more comprehensively[2] and showed it is a viable alternative for screening and selecting among project alternatives

2. **Demonstrated its application via the use of a number of case-study examples and compare with other metrics presently in use.** Stochastic dominance is a general objective function framework for incorporating risk under the expected utility decision framework (Rothschild and Stiglitz, 1970). We implemented these algorithms in a tool we develop on different project selection scenarios and compare the performance of alternative implementations of the algorithms and explore their use to the different project selection scenarios.

3. **Demonstrated ways in which it can be further refined to give unique decisions when needed.** We have done this by exploring how the methodology can be embedded in the stage-gate decision making framework that is used in engineering project decisions.

---

[2]First order stochastic dominance is briefly mentioned in [112] but in a cursory, illustrative idea of the concept using a very simple example and it doesn't explore the concept in depth

Appendix A

Models for Reactor Separator Problem

This appendix contains the technical and economic models used for the Reactor-Separator Problem. The model structure follows that developed by Kiss ??.

A.1 Technical Model

A.1.1 First Order reaction A $\longrightarrow$ C

We build a simple model of a reactor and a separator with the system displayed in Figure A-1. The system is for the conversion of a reactant A to a product C via an irreversible first order reaction given by

$$A \longrightarrow C$$
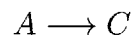
We represent the total number of moles (molar flow rate) in each stream by F_i and the mole fraction of component A in the reation by x_{Ai}, where i is a subscript that indicates the particular stream. With these equations, we can write material and component balances for each unit (the mixing point, reactor and the separator as follows:

For the mixing point, we have the total balance as

$$F_1 + F_2 = F_3 \tag{A.1}$$

and the component balance as:

$$F_1 x_{A1} + F_2 x_{A2} = F_3 x_{A3} \tag{A.2}$$

And for the reactor, we have,

$$F_3 = F_4 \tag{A.3}$$

The above equation holds because of the reactor stoichiometric ratio: for every mole of reactant A that reacts, one mole of C is formed keeping the total constant.
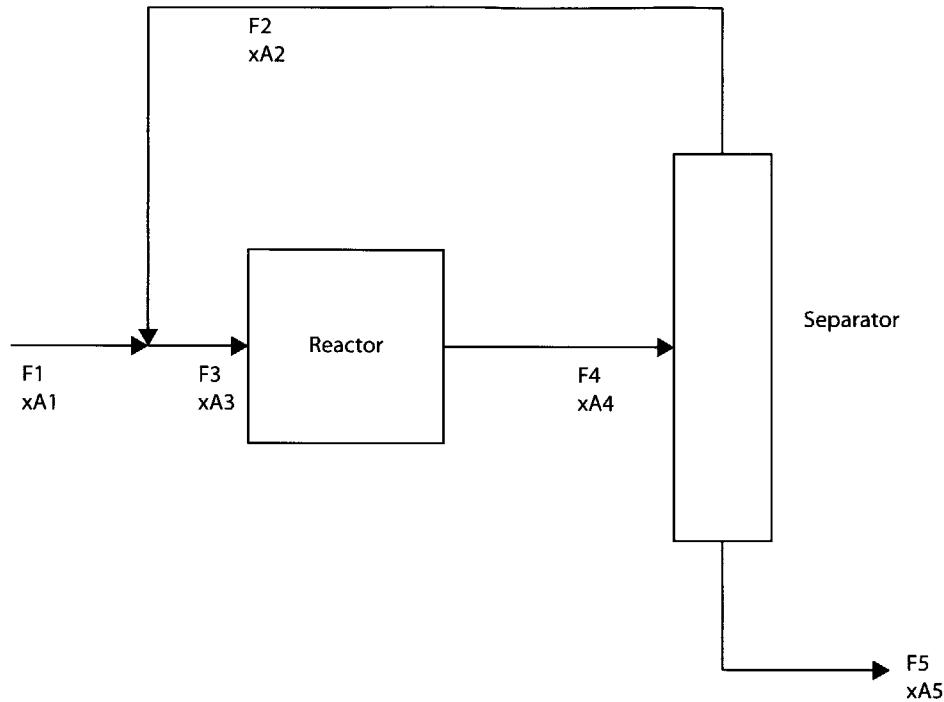
Figure A-1: Model of reactor-separator system for a single reaction

The component A balance however is given by

$$F_3 x_{A3}(1 - X) = F_4 x_{A4} \qquad (A.4)$$

where X is the conversion (per pass) for the reactor. And finally, the balance for the separator is given by the equations

$$F_4 = F_2 + F_5 \qquad (A.5)$$

And the component balance as:

$$F_4 x_{A4} = F_2 x_{A2} + F_5 x_{A5} \qquad (A.6)$$

In the above system of reactions, we have 11 variables (5 streams, components and the conversion) and 6 equations relating them. As a result, we have $11 - 6 = 5$ degrees of freedom or variables that we have to fix to fully specify the system. First, we choose to specify the following 5 variables: F_1, x_{A1}, x_{A2}, x_{A5} and X. We will revisit the specification of the conversion when we analyze particular reactors and reaction kinetics.

We can combine equations A.3 A.1 and A.5 to get the overall balance and show that

$$F_5 = F_1 \qquad (A.7)$$

196

And use equation A.3 in equation A.4 to get

$$x_{A4} = x_{A3}(1 - X) \tag{A.8}$$

Using these new relationships to eliminate F_3 and F_5, we get the following set equations

$$F_1(x_{A1} - x_{A2}) = F_4(x_{A3} - x_{A2}) \tag{A.9}$$

$$F_1(x_{A5} - x_{A2}) = F_4(x_{A4} - x_{A2}) \tag{A.10}$$

From the second equation in A.10, we can write F_4 as

$$F_4 = \frac{F_1(x_{A5} - x_{A2})}{x_{A3}(1 - X) - x_{A2}} \tag{A.11}$$

where we have substituted the value of x_{A4} from equation A.8. We can then substitute equation A.11 into the first equation in A.10 and simplify to obtain an equation in terms of x_{A3} or

$$x_{A3} = \frac{x_{A2}(x_{A1} - x_{A5})}{(1 - X)(x_{A1} - x_{A2}) - x_{A5} + x_{A2}} \tag{A.12}$$

Thus we obtain equations for all the non-specified variables in terms of the specified ones.

So far, we have not assumed any particular reactor or reaction kinetic model and as such the above set of equations (being based on mass balances alone) can be adapted to any system that has just two components. We will first focus on adapting it to a CSTR.

CSTR For a CSTR, we have the relationship between the conversion and the volume, V of the reactor given as

$$V = \frac{F_{Ao}X}{-r_A} \tag{A.13}$$

where F_{Ao} is the entering molar flowrate of reactant A and $-r_A$ is the reaction rate - the rate of consumption of reactant A. For a first order reaction (which we assume our present reaction system is), we hvae that

$$-r_A = kC_A \tag{A.14}$$

where k is the reaction rate constant and C_A is the concentration of the reactant at the exit [36]. For our reactor separator system, we have

$$F_{Ao} = F_3 x_{A3} \tag{A.15}$$

We assume that overall the system operates at constant density so that we can

write

$$C_A = C_{Ao}(1 - X) \tag{A.16}$$

With this system of equations, if we specify the volume (instead of the conversion as we did previously), we can substitute the various expressions into the reactor equation A.13 and simplify to obtain

$$V = \frac{F_3 X}{kC_1(1 - X)} \tag{A.17}$$

where C_1 is the concentration of A in the initial flow stream. Defining the plant Damkohler number [57] as

$$Da = \frac{kC_1 V}{F_1} \tag{A.18}$$

and substituting for F_3 using equation A.3 and equation A.11, we transform equation A.17 to obtain

$$\frac{1}{X} = \frac{x_{A5} - x_{A2}}{Da[x_{A3}(1 - X) - x_{A2}]} + 1 \tag{A.19}$$

which, when we recall equation A.12 and substitute for x_{A3} in equation A.19 gives

$$\frac{1}{X} - 1 = \frac{x_{A5} - x_{A2}}{Da\left[\frac{x_{A2}(1-X)(x_{A1}-x_{A5})}{(1-X)(x_{A1}-x_{A2})-x_{A5}+x_{A2}} - x_{A2}\right]} \tag{A.20}$$

Equation A.20 is a non-linear equation in one variable, X, and can be solved to obtain a relationship between X and Da that is more general than that given in the literature by Andrieu [57]. To obtain their solution for perfect separation (i.e. pure recycle and pure product), we have that $x_{A5} = 0$, $x_{A2} = 1$ and $x_{A1} = 1$ and substituting into the above equation and simplifying gives

$$\frac{1}{X} - 1 = \frac{1}{XDa} \tag{A.21}$$

which can be rewritten to give the expression they obtain, i.e.

$$X = \frac{Da - 1}{Da} \tag{A.22}$$

We plot the stability for a range of Da for both the case above and the case where $x_{A5} = 0.2$, $x_{A2} = 0.8$ and $x_{A1} = 1$.

$$X = \frac{Da}{1 + Da} \tag{A.23}$$

which is stable for all values of Da.

From equation A.22, we see that the system is stable for values of $Da \geq 1$ as values of $Da < 1$ results in an unrealistic reaction system (negative conversion). We will see later that for imperfect separation (i.e. values of $x_{A2} < 1$ or $x_{A5} > 0$, there is
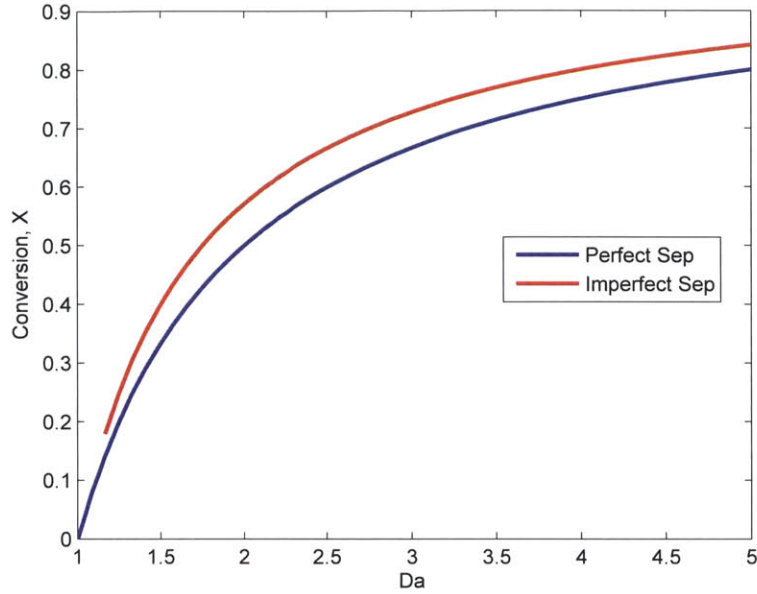
Figure A-2: A plot of reactor conversion versus plant Damkohler number for the CSTR reactor separator for both perfect separation and imperfect separation

also an upper bound to the value that the Damkohler number can take else it will also lead to stability. This is in contrast to the case of a single CSTR with no separator attached to it where the reaction system is stable for all values of the Damkohler number as the conversion-Damkohler number equation for that case is [57]

PFR Most of the equations that we developed for the CSTR remain valid for the plug flow reactor (PFR) since they largely involve the same material balances. Where they differ however is the equation that relates the volume and the conversion, which for the PFR is given by [36]

$$V_{PFR} = F_{Ao} \int \frac{dX}{-r_A} \tag{A.24}$$

where, in this case, F_{Ao} is the inlet molar flowrate of the reactant, and $-r_A$ is the reaction rate. Substituting the relevant expressions from the CSTR model, we have

$$V_P FR = \frac{F_3 x_{A3}}{kC_{A3}} \int \frac{dX}{1-X} = \frac{F_3 x_{A3}}{kC_{A3}} ln \frac{1}{1-X} \tag{A.25}$$

Note that the concentration and the molar flowrate of a particular stream are related by the equation

$$C_A = \frac{F_{Ai}}{\nu_i} \tag{A.26}$$

where ν_i is the volumetric flowrate of stream i. By making the assumptions about

199

volumetric flowrates (being directly proportional to molar flowrates given that total moles are conserved), we get

$$\nu_i = \nu_1 \frac{F_i}{F_1} \qquad \text{(A.27)}$$

where F_1 is the inlet flowrate of the stream to the plant. Using the above relationships, remembering that $F_{Ai} = F_i x_{Ai}$, we can substitute into the equation A.25 and rearrange terms to get

$$X = 1 - exp(\frac{-kC_{A1}V}{F_3}) \qquad \text{(A.28)}$$

Also, by recalling the equaton for F_3 and substituting for the Damkohler number defined in equation A.18 we have that

$$X = 1 - exp(\frac{-Da}{\frac{(x_{A5}-x_{A2})}{x_{A3}(1-X)-x_{A2}}}) \qquad \text{(A.29)}$$

In which we can finally substitute the expression developed for x_{A3} in equation A.12 and then solve the nonlinear equation for X using a nonlinear solver in MATLAB

A.1.2   Second order reaction A + B $\longrightarrow$ D

The equations for the second order reaction are quite similar to that for the first order reaction but with more variables to keep track of. In the first case, our model had 11 variables in total and we could write 6 independent equations which meant we had to fix 5. As shown in Figure A-3, there are 7 streams, each with 3 variables (flowrate and two components to track, A and B) as well as conversion which makes 22 variables in all. We can write 12 independent equations - total flowrate and component balances for A and B around mixers 1 and 2 as well as around the reactor and the separator. So we have to specify 10 variables in order to fully solve the system. In this model development, we will begin (like the previous one for the single reaction order) that we can specify the conversion and then when we analyze the individual reactor types, we will then transfer the specification from the reactor to the volume which is the decision variable.

For the mixers and the reactor, the three equations follow the order

$$F_i = F_j + F_k \qquad \text{(A.30)}$$
$$F_i x_{Ai} = F_j x_{Aj} + F_k x_{Ak} \qquad \text{(A.31)}$$
$$F_i x_{Bi} = F_j x_{Bj} + F_k x_{Bk} \qquad \text{(A.32)}$$

where the subscripts represent the different streams going in and out of the unit. In such a way, relationships can be written that relates streams 7, 1 and 2; streams 4, 5 and 6 and streams 3, 5, and 7 and account for 9 out of the twelve equations for the system we are trying to build. For the reactor, the big difference between this system
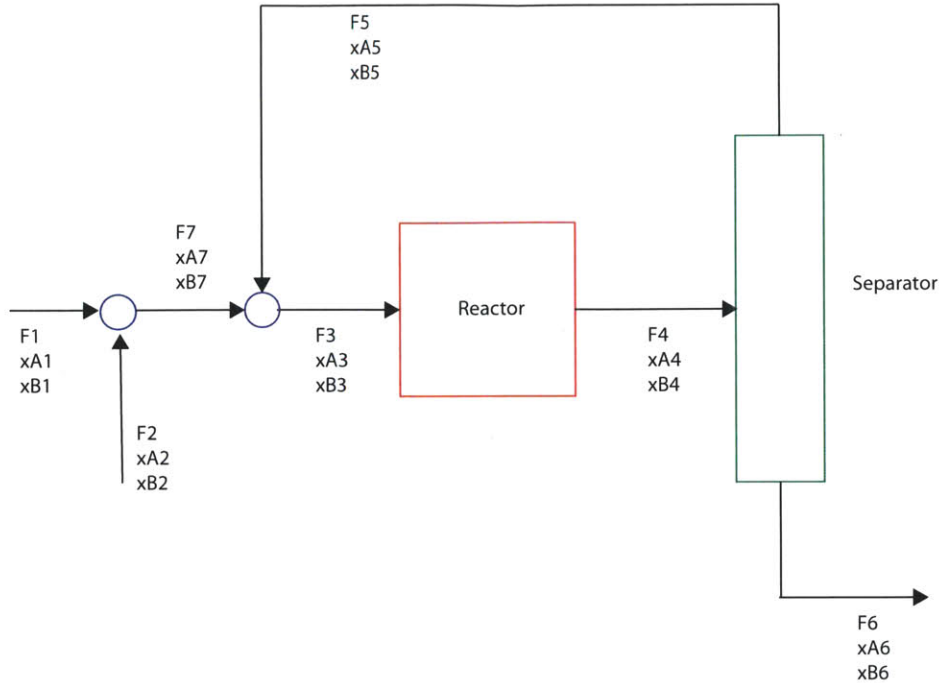
Figure A-3: Reactor-Separator flow sheet for the reaction case A + B $\longrightarrow$ D

and the earlier first order system is that total number of moles is not conserved in the reaction. As a result, we need to carefully do balance. For components A and B, the amount of reactant coming in must equal the amount entering plus any consumed by the reaction which leads to the equations

$$F_4 x_{A4} = F_3 x_{A3}(1 - X) \tag{A.33}$$

$$F_4 x_{B4} = F_3 (x_{B3} - x_{A3} X) \tag{A.34}$$

Note that for this system, we define conversion with respect to the reactant A, which we assume to be the limiting reactant. For the flowrates, we do a total balance which yields

$$F_4 = F_3 (1 - x_{A3} X) \tag{A.35}$$

To fully solve the system of equations, we need to specify 10 variables. From physical insight, we select streams 1 and 2 and all the components (six variables), conversion (which we can get from knowledge of the volume), x_{A5}, x_{B5} and x_{A6}. Once all of these are specified, we can write out the equations for the entire system for the unknown variables as follows:

$$F_7 = F_1 + F_2 \tag{A.36}$$

$$x_{A7} = \frac{F_1 x_{A1} + F_2 x_{A2}}{F_7} \tag{A.37}$$

$$x_{B7} = \frac{F_1 x_{B1} + F_2 x_{B2}}{F_7} \tag{A.38}$$

$$F_5 = \frac{F_7(x_{A7}(1 - X + x_{A6}X) - x_{A6})}{x_{A5}X(1 - x_{A6})} \tag{A.39}$$

$$F_3 = F_7 + F_5 \tag{A.40}$$

$$x_{A3} = \frac{F_7 x_{A7} + F_5 x_{A5}}{F_3} \tag{A.41}$$

$$x_{B3} = \frac{F_7 x_{B7} + F_5 x_{B5}}{F_3} \tag{A.42}$$

$$F_4 = F_3(1 - x_{A3}X) \tag{A.43}$$

$$x_{A4} = \frac{F_3 x_{A3}(1 - X)}{F_4} \tag{A.44}$$

$$x_{B4} = \frac{F_3(x_{B3} - x_{A3}X)}{F_4} \tag{A.45}$$

$$F_6 = F_7 - F_3 x_{A3}X \tag{A.46}$$

$$x_{B6} = \frac{F_4 x_{B4} - F_5 x_{B5}}{F_6} \tag{A.47}$$

The equations developed is valid for any two-component reactor separator system with the variables specified as highlighted and applies to both the CSTR and the PFR. The difference in the model for the two reactor types lies in the equation for reactor volume. For second order kinetics, we assume the reaction model

$$-r_A = kC_A^2 \tag{A.48}$$

Combining this with equation A.13, we obtain the volume of a reactor needed for a given conversion to be

$$V_{CSTR} = \frac{F_{Ao}X}{kC_{Ao}^2} \tag{A.49}$$

where Ao indexes the entry conditions for the limiting reactant. Substituting for the appropriate expressions for F_{Ao} and C_{Ao} from Figure A-3 gives the reactor volume equation as

$$V = \frac{F_3 x_{A3}X}{kC_{A4}^2} \tag{A.50}$$

The PFR, on the other hand has the reactor volume equation for second order kinetics as

$$V_{PFR} = \frac{F_{Ao}}{kC_{Ao}^2} \frac{X}{1-X} \qquad (A.51)$$

This equation is obtained by integrating the expression obtained from substituting the rate equation A.48 into the expression in equation **??**. Subsituting for the appropriate expressions for F_{Ao} and C_{Ao} as we did for the CSTR case gives the reactor volume equation as

$$V = \frac{F_3 x_{A3} X}{kC_{A3}^2(1-X)} \qquad (A.52)$$

Equations A.50 and A.52 are then incorporated into the overall mass balance systems developed for the general reactor-separator system to obtain the design equation for the particular system under study.

A.2 Economic model

In addition to the technical model, we need to build an economic model on which we can run the decision framework. Over the lifetime, L, of the plant, we have the total cost (and for the moment, we are going to ignore discounting)

$$Cost_T = Cost_{reactor} + Cost_{separator} + Cost_{reactant} \qquad (A.53)$$

and the total revenue from the system is given by the sales of the product i.e

$$Revenue = salesproduct \qquad (A.54)$$

And the total profit is given as

$$Profit = Revenue - Cost \qquad (A.55)$$

To thus be able to obtain the profit (and later be able to do a proper Net Present Value analysis when we incorporate discounting), we need to determine each of the various components of equations and we will estimate them accordingly

Reactor cost From elementary principles, we determine the cost of the CSTR to be

$$C_{cstr} \propto V^{2/3} \implies C_{cstr} = k_1 V^{2/3} \qquad (A.56)$$

where V is the volume of the reactor and k_1 is the constant of proportionality. For a PFR, the reactor being more similar to a cylinder than a sphere, we have that the cost scales linearly with the volume or

$$C_{pfr} \propto V \implies C_{pfr} = k_2 V \qquad (A.57)$$

Also from first principles, we determine the cost of a separator as

203

$$C_{sep} \quad \propto \quad f_1(N) \times f_2(D) \implies C_{sep} = k_2 \times f_1(N) \times f_2(D) \qquad \text{(A.58)}$$

where N is the number of trays in the column and D is the diameter and f_1 and f_2 are functions. The number of trays, for example, affects the total height of the column. A simplified equation for the total number of trays in binary distillation is given by Fenske's equation [71] which is given as

$$N = \frac{ln[x_{A2}(1 - x_{A5}/x_{A5}(1 - x_{A2}]}{ln\alpha_{ac}} \qquad \text{(A.59)}$$

where α_{ac} is the average relative volatility of the components. To correct for the fact that this is an ideal number of trays, we can account for real trays by using an efficiency i.e

$$N_r = \frac{N_F}{\epsilon} \qquad \text{(A.60)}$$

where ϵ is the efficiency of separation and N_r is the total number of trays. Given the number of trays, we can then proceed to calculate the height of the distillation column as

$$H_C = H_{min} + h \times N_r \qquad \text{(A.61)}$$

where H_C is the height of the column, H_{min} is the total height of the column and h is the average tray separation.

From [], we see that the diameter of a column is proportional to the square root of the flowrate within it. Taking the inlet flowrate as a first order estimate of the total flowrate gives

$$D \quad \propto \quad F_4^{1/2} \implies D = D_{min} + k_3 \times F_4^{1/2} \qquad \text{(A.62)}$$

where D_{min} is the minimum diameter, k_3 the constant of proportionality and F_4 the inlet flowrate.

The total amount of raw material we assume is purchased over the lifetime, L, of the plant is given by

$$C_{reactant} = s_A \times F_1 \times L \qquad \text{(A.63)}$$

where s_A is the selling price of a unit amount of A, F_1 is the inlet flowrate - all proper units taken into consideration. And total revenue from the sale of the product is given by

$$Revenue = s_C \times F_5 \times L \qquad \text{(A.64)}$$

where F_5 is the flowrate of the product and s_C is the selling price per unit of product. Putting this all together, we have that the total profit received from the reactor-separator system over its lifetime is given by

Parameter	Name	Unit	Value
C_1	Concentration of A	kmol/m^3	1
k_{CSTR}	Rate constant (CSTR)	day^{-1}	220
k_{PFR}	Rate constant (PFR)	day^{-1}	200
k_1	CSTR Cost Factor	\$/m^3	20000
k_2	PFR Cost Factor	\$/m^3	35000
k_{sep}	Dist. Column cost Factor	\$/m^3	5500
H_{min}	Minimum column height	m	2
D_{min}	Minimum column diameter	m	1
h_{tray}	tray separation	m	0.3
ϵ	tray efficiency	-	0.7
α	relative volatility	-	2
sp	product C selling price	\$/kmol	12
sc	reactant A cost price	\$/kmol	6

Table A.1: Some key parameters in the Reactor-Separator Example

$$P = L(s_C F_5 - s_A F_3) - k_1 V^{2/3} - k_2(D_{min} + k_3 \times F_4^{1/2})(H_{min} + h \times N_r) \quad \text{(A.65)}$$

where P is the profit. If there is a budget constraint on the capital cost, then we can include the following restriction

$$k_1 V^{2/3} - k_2(D_{min} + k_3 \times F_4^{1/2})(H_{min} + h \times N_r) \leq B \quad \text{(A.66)}$$

where B is the budget for the design

The parameters for the problem are given in Tables A.1 and A.2 and table A.3 describes the uncertain parameters used for the simulations.

Parameter	name	Unit	Value
C_1	Concentration of A	$kmol/m^3$	2
C_2	Concentration B	$kmol/m^3$	2
k_{CSTR}	Rate constant (CSTR)	day^{-1}	220
k_{PFR}	Rate constant (PFR)	day^{-1}	200
k_1	CSTR Cost Factor	$\$/m^3$	25000
k_2	PFR Cost Factor	$\$/m^3$	50000
k_{sep}	Dist. Column cost Factor	$\$/m^3$	5000
H_{min}	Minimum column height	m	2
D_{min}	Minimum column diameter	m	1
h_{tray}	tray separation	m	0.3
ϵ	tray efficiency	-	0.7
α	relative volatility	-	1.75
sp	product D selling price	$\$/kmol$	17.6
sc_1	Reactant A cost price	$\$/kmol$	8
sc_2	Reactant B cost price	$\$/kmol$	6

Table A.2: Some key parameters in the Reactor-Separator Example

Parameter	Distribution type	Distribution parameters
k_{CSTR}, product C	Lognormal	$\mu = 220$, $\sigma = 20$
k_{PFR}, product C	Lognormal	$\mu = 200$, $\sigma = 12$
selling price, C	Lognormal	$\mu = 15.5$, $\sigma = 0.7$
k_{CSTR}, product D	Lognormal	$\mu = 220$, $\sigma = 25$
k_{PFR}, product D	Lognormal	$\mu = 200$, $\sigma = 15$
selling price, C	Lognormal	$\mu = 17.6$, $\sigma = 1.4$

Table A.3: Description of uncertain parameters for the reactor-separator

Appendix B

Algorithms and Computational Cost

In this chapter, we include some of the algorithms that were implemented for the uncertainty analysis perfomed in the case studies. We also discuss the compuational cost associated with a number of the objectives that we considered in our analysis.

B.1 Overall Structure

All of the case-studies required a the creation of a deterministic model that takes in a set of decision variables and a set of parameters and generates the output we want. For example, in the crop selection model, a deterministic model was created which took in the crop choices and the farm radius as the decision variables and the crop selling price, growth cost, yield, establishment costs and the transportation cost as parameters. With the variables and the parameters, the model returns the net present value for the particular crop and farm radius. This set of outputs form the basis for the analysis of the deterministic case.

For the uncertain case, the parameters are allowed to vary according to prescribed distributions (according to the principles described in Chapter 3). A number of parameter sets are generated, up to the total number of predetermined parameter sets, with each parameter in that set sampled from its distribution. The collection of parameters form the parameter matrix which are then used to generate the distribution of outputs for each decision variable. Each parameter set is then used to generate an output for each decision, a 'sample' from the decision's output distribution. This way, a collection of output samples for each decision is generated. The uncertainty analysis using the different uncertain objectives (stochastic dominance, mean-variance, etc.) is then carried out using the uncertainty matrix, indexed by decision.

The algorithms that follow all assume that this matrix, which we call F, has already been generated. Each column represents a specific decision - such as the choice of crop and radius in the crop selection case study. The first few rows of each column contain the details of the decision while the output samples follow in the remaining rows.

B.2 The dual-objective Pareto Algorithm

This thesis analyzed a number of dual-objective approaches to selecting projects in the presence of uncertainty and we often had to determine the efficient set for these dual-objectives. To do this, we required two numbers calculated from the samples of each decision's output distribution: the mean of the decisions M_j, and the risk-metric that is used in particular, R_j[1]. Once these numbers are calculated and stored in the respective vectors M and R, the algorithm below is used to determine the efficient set

1. Start and set $j = 1$

2. Take the respective elements of the mean and risk vectors to get the mean-risk pair, $[M_j, R_j]$

3. For the remaining decisions to the last decision, take the respective mean-risk pair - generically labelled $[M_i, R_i]$

4. Compare M_j and M_i, and R_1 and R_i

   (a) If $M_j > M_i$ and $R_j < R_i$, decision j dominates decision i. Eliminate the rows in M and R corresponding to decision i and move to the next decision, $i + 1$ and repeat step 3.

   (b) If $M_j < M_i$ and $R_j > R_i$, decision i dominates decision j. Eliminate the rows in M and R corresponding to decision j. Move to decision $j + 1$ and go to step 5.

5. Proceed to the next decision $j + 1$ and return to step 2. Repeat till all the elements of M and R have been analyzed.

6. Return the decisions corresponding to the remaining indices in M and R. These decisions form the efficient set of the the mean-risk dual objective.

7. Stop

B.3 Algorithm for Mean-Variance

With the matrix F and an algorithm for calculating mean-risk efficient sets, the algorithm for calculating the mean-variance efficient set follows directly. It is presented more explicitly below:

1. Start

2. Calculate the mean, M_j and standard deviation, R_j for each decision in F. Collect these in vectors M and R

---

[1]This can be the standard deviation, the absolute semi-deviation or anyone of the many risk metrics that were highlighted in Chapter 2.

3. Run the mean-risk algorithm using the vectors M and R.

4. Return the decisions corresponding to the remaining indices in M and R. This is the mean-variance efficient set

5. Stop

B.4 Algorithm for FSD and SSD

The algorithm used to obtain the efficient sets for First and Second order Stochastic Dominance are given below. Because the SSD efficient set is a subset of FSD efficient set, the algorithm generates both efficient sets together - but uses the FSD efficient set as the springboard for the SSD set. This allows fewer computations because the initial set for the SSD simulations is the FSD set - often much smaller than the initial feasible. Like the mean-risk analysis, this set begins with the uncertainty matrix, F. First we present the FSD algorithm and then the SSD portion of the algorithm.

B.4.1 FSD algorithm

1. Start

2. Sort the samples in each column of F in an ascending order - smallest to the largest sample - to produce a matrix, G

3. Determine the mean value of each column, $\hat{G}_j$ and re-order the columns with respect to their means in a descending order such that $\hat{G}_1 > \hat{G}_2 > ... > \hat{G}_j > \hat{G}_{N-1} > \hat{G}_N$

4. For $j = 1$ and i going from $j + 1$ to N

   (a) Get the cumulative distributions of G_j and G_i (cumulative distribution algorithms are described in the next section) - CD_j and CD_i

   (b) For $k = 1 : N$, compare $CD_{j,k}$ and $CD_{i,k}$

   (c) If $CD_{j,k} <= CD_{i,k}$ for all k, and $CD_{j,k} < CD_{i,k}$ for at least one k, then decision G_j dominates G_i. Eliminate column i from G and move to the next decision, G_{i+1}. Otherwise, leave G_i in the efficient set and proceed to the next decision[2].

   (d) Continue till G_j has been compared with all the remaining decisions G_{j+1} to G_N

5. Set j = j+1 and return to Step 4. Continue to $j = N$ and stop

6. All the decisions left in G form the FSD efficient set.

---

[2]Note that we don't have to check that i dominates j because we have already sorted the decisions based on means and a lower mean distribution cannot dominate a higher mean distribution as discussed in Chapter 5.

7. Stop

B.4.2 SSD algorithm

1. Start with the matrix of FSD efficient set decisions and samples, H

2. Set $j = 1$. For $i : j + 1$ to N

    (a) Get the cumulative distributions of H_j and H_i - CD_j and CD_i

    (b) Get the cumulative integral of cumulative distributions of CD_j and CD_i - FD_j and FD_i. This is obtained by integrating the cumulative distribution, CD from 1 to l with respect to the outcome space - where l ranges from $k = 2$ to N. For $k = 1$, $CD_{j,1} = FD_{j,1}$

    (c) For $k = 1 : N$, compare $FD_{j,k}$ and $FD_{i,k}$

    (d) If $FD_{j,k} <= FD_{i,k}$ for all k, and $FD_{j,k} < FD_{i,k}$ for at least one k, then decision H_j dominates H_i. Eliminate column i from H and move to the next decision, H_{i+1}. Otherwise, leave H_i in the efficient set and proceed to the next decision[3].

    (e) Continue till H_j has been compared with all the remaining decisions H_{j+1} to H_N

3. Set j = j+1 and return to Step 2. Continue to $j = N$ and stop

4. All the decisions left in H form the SSD efficient set.

5. Stop

The algorithms for higher order stochastic dominance are similar although Step 4 is a litle more intricate for TSD. Interested readers are referred to [64] for more details about the subtle nuances in the algorithms.

B.5   Cumulative Distribution Algorithm

At the heart of the FSD and SSD algorithm lies the generation of cumulative distributions (and integrals of cumulative distributions) needed for pairwise comparison. These distributions are generated empirically and the algorithms presented here describe how we do so from the samples obtained from running the models. Because we use the cumulative distributions for pairwise comparison, our analysis presents the generation of cumulative distributions for a pair of vectors. The advantage of this approach is that it minimizes the storage requirements needed during the determination of the efficient sets - leading to speed improvements as well as the capacity to handle much larger sample sizes.

---

[3]As in the SSD, we don't have to check that i dominates j for the same reason highlighted

210

This algorithm needs two vectors, A and B which, in general, need not be the same size. For our analysis, we let l be the number of elements in A, m be the number of elements in B and n be the number of elements in C. Because some elements in A may be in B, $n <= l + m$. Further we assume that the samples from each vector have equal probabilities - $\frac{1}{l}$ for samples in A and $\frac{1}{m}$ in B.

1. Start

2. Merge the samples in A and B into a single output sample, C

3. Sort the samples in A, B and C in ascending order

4. Select ordered matrix A

5. For $k = 1 : m$

   (a) $\hat{i} = 1$

   (b) For $i = \hat{i} : l$

   (c) if $C_k < A_i$ then $CD_{A,k} = \frac{i-1}{l}$, set $\hat{i} = i$, increment k go to Step 5

   (d) Else $i = i + 1$, return to previous step

6. if $C_k >= A_l$, then $CD_{A,k:m} = 1$

7. Select vector B and return to step 4

8. Stop

B.6 Computational cost

Below we discuss the derivation of the expressions that we gave Table 5.7 in the discussion of the computational complexity of the different objectives in the thesis.

Expected Value For the expected value objective, the determination of the eventual decision takes place in two phases: first, the expected value of all the decisions are calculated - which involves taking the arithmetic mean of m samples for N different vectors for a total Nm operations. After this, the maximum of the list of N expected values has to be obtained leading to an additional $\frac{3N}{2}$ comparison operations. Thus the total compuatational time for the expected value objective is:

$$C_{EV} = Nm + \frac{3N}{2} \tag{B.1}$$

Mean Variance For the mean-variance approach, we first have to calculate the mean and the variance of each of the vectors. As we saw in the previous paragraph, the computation of the mean requires $N \times m$ calculations while that for the standard deviation requires $3Nm + N$ calculations - Nm operations to square each element of the matrix, Nm operations to take the difference from the respective mean (already calculated), Nm operations to average the squared sums and N square root operations. Once the means and standard deviations have been computed, we need to compare them in order obtain the efficient set. This comparison can be broken down into two parts.

- First the total number of comparisons across the entire list of projects is $\frac{N(N-1)}{2}$ which is simply the number of uinique pairwise combinations that can be done in a list of N object.

- Next within each pairwise comparison, two comparisons are made: a comparison of the means and a comparison of the standard deviations

All of these give a total cost for the mean-variance approach as:

$$C_{MV} = Nm + 3Nm + 2 \times \frac{N(N-1)}{2} = 4Nm + N(N-1) \tag{B.2}$$

First order Stochastic Dominance The FSD procedure takes place in a number of steps highlighted below

1. To start we first calculate the means of the different distributions and rank the means from highest to lowest (as discussed in the algorithm section). This takes a total of $Nm + N$ steps - Nm to calculate the mean and N as the average complexity of ranking a list [].

2. Each of the ranked vectors needs to be sorted. Sorting each vector of m samples takes $m \log m$ steps and so the total number of steps to sort all the vectors is $Nm \log m$

3. Next we carry out pairwise comparisons of the different vectors in the feasible set. There are $\frac{N(N-1)}{2}$ possible such combinations. Each vector comparison requires that we compare the samples within each vector. This gives m comparisons for each pair wise comparison - or a total of $\frac{mN(N-1)}{2}$ comparisons for this step.

This thus, yields a total number of steps for the FSD method as

$$C_{FSD} = Nm + N \log N + Nm \log m + \frac{Nm(N-1)}{2} \tag{B.3}$$

Second order Stochastic Dominance As can be seen in the algorithm section, the steps required for the computation of second order stochastic dominance are very similar to those for FSD. However, there is one key difference and that is in computing the vectors for comparison, we not only sort (as is required for FSD) but also have to numerically integrate the sorted samples to arrive at the final vector for comparison. In the previous discussion, we saw that the sorting step for each vector takes $Nm \log m$. The numerical integration of the ranked sample is a form of growing average requiring roughly the sum of j terms for the jth sample. Thus in total, the evaluation of the average requires about $\frac{m(m-1)}{2}$ operations [4].

Thus, if we replace the computational cost of step 2 in the FSD with the modified number for the SSD - $m \log m + \frac{m(m-1)}{2}$, we get the total number of steps required for the SSD as

$$C_{SSD} = Nm + N \log N + N(m \log m + \frac{m(m-1)}{2}) + \frac{Nm(N-1)}{2} \qquad \text{(B.4)}$$

---

[4]This is simply the sum of the growing series of numerical computations $1, 2, ..., j, ..., m-1, m$

Bibliography

[1] Swithgrass for biomass, November 2009. http://www.uky.edu/Ag/NewCrops/introsheets/switchgrass.pdf. Accessed October 2012.

[2] Miscanthus best practice guidelines. Technical report, Teagasc, Crops Research Centre; Agri-Food Bioscience Institute, Northern Ireland, April 2011. http://www.teagasc.ie/publications/2011/315/MiscanthusBestPractice.pdf. Web report. Accessed October 11, 2012.

[3] Aspentech, 2012. http://www.aspentech.com/.

[4] Gasification technologies, 2012. http://www.thyssenkrupp-uhde.de/fileadmin/documents/brochures/gasification_technologies.pdf. Accessed, January 2013.

[5] Lawrence P. Abrahamson, Timothy A. Volk, Lawrence B. Smart, and Edwin H. White. Short-rotation willow for bioenergy, bioproducts, agroforestry and phytoremediation in the northeastern united states. Technical report, IEA Bioenergy, 2012. http://142.150.176.36/task43/library/promisingresources/IEA_Bioenergy_Task43_PR2012_01.pdf. Accessed October 2012.

[6] Sijesh C. Aravindhakshan, Francis M. Epplin, and Charles M. Taliaferro. Economics of switchgrass and miscanthus relative to coal as feedstock for generating electricity. *Biomass and Bioenergy*, 34:1375–1383, 2010.

[7] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[8] Ahmed Aseeri and Miguel J. Bagajewicz. New measures and procedures to manage financial risk with applications to the planning of gas commercialization in asia. *Computers and Chemical Engineering*, 28:2791 – 2821, 2004.

[9] M. Bagajewicz. On the use of net present value in investment capacity planning models. *Industrial and Engineering Chemistry Research*, 47:9413–9416, 2008.

[10] Kenneth L. Barker, John H. Stek, Walter W. Wessel, and Ronald Youngblood, editors. *The Holy Bible (New International Version)*. Zondervan, 1984.

[11] Paul I. Barton. Mixed integer and non-convex optimization. Class notes, Spring 2011.

[12] Vijay S. Bawa. Optimal Rules for Ordering Uncertain Prospects. *Journal of Financial Economics*, 2:95–121, 1975.

[13] Vijay S. Bawa, James N. Jnr. Bodurtha, M. R. Rao, and Hira L. Suri. On determination of stochastic dominance optimal sets. *The Journal of Finance*, 40(2):417–431, June 1985.

[14] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust Optimization*. Princeton University Press, 2009.

[15] Ingrid Berkelmans. *Development and Applicaton of a Framework for Technology and Model Selection under Uncertainty*. PhD thesis, Massachusetts Institute of Technology, June 2010.

[16] Lorenz T. Biegler and Ignacio E. Grossmann. Retrospective on optimization. *Computers & Chemical Engineering*, 28:1169–1192, 2004.

[17] Harold Boerrigter. Economy of Biomass-to-Liquids (BTL) plants. Technical report, Energy research Centre of the Netherlands (ECN) - Biomass Coal and Environmental research, May 2006. ftp://question.ecn.nl/pub/www/library/report/2006/c06019.pdf.

[18] Richard A. Brealey and Stewart C. Myers. *Principles of Corporate Finance*. McGraw Hill/Irwin, 2003.

[19] Robert C. Brown. *Thermochemical Processing of Biomass - Conversion into Fuels, Chemicals and Power*. John Wiley & Sons, 2011.

[20] S. L. Brunelle and R.G Vickson. A unified approach to stochastic dominance. In *Stochastic Optimization Models in Finance*. New York: Academic Press, 1975.

[21] J. A. Cano-Ruiz and G. J. McRae. Environmentally conscious chemical process design. *Annual Review of Energy and the Environment*, 23:499–536, November 1998.

[22] Allison An Chang. *Integer Optimization Methods for Machine Learning*. PhD thesis, Massachusetts Institute of Technology, June 2012.

[23] Steven C. Chapra and Raymond P. Canale. *Numerical Methods for Engineering*. McGraw Hill, 2002.

[24] A. Charnes and W.W. Cooper. Chance-constrained programming. *Management Science*, 6(1):73–79, October 1959.

[25] L. Cheng, E. Subrahmanian, and A.W. Westerberg. Design and Planning under Uncertainty: issues on problem formulation and solution. *Computers and Chemical Engineering*, 27:781–801, 2003.

[26] John Clifton-Brown, Steve Renvoize, Yu-Chung Chiang, Yasushi Ibaragi, Richard Flavell, Joerg Greef, Lin Huang, Tsai Wen Hsu, Do-Soon Kim, Astley Hastings, Kai Schwarz, Paul Stampfl, John Valentine, Toshihiko Yamada, Qingguo Xi, and Iain Donnison. Developing miscanthus for bioenergy. In Nigel G. Halford and Angela Karp, editors, *Energy crops*, chapter 15, pages 301 321. Royal Society of Chemistry, 2010. `http://pubs.rsc.org/en/content/chapter/bk9781849730327-00301/978-1-84973-204-8`. Accessed October 2012.

[27] Keith Conrad. Probability distributions and maximum entropy. `http://www.math.uconn.edu/~kconrad/blurbs/analysis/entropypost.pdf`. Accessed July 2012.

[28] Mark Daichendt and Ignacio E. Grossmann. Integration of hierarchical decomposition and mathematical programming for the synthesis of process flowsheets. *Computers & Chemical Engineering*, 22(1-2):147–175, 1997.

[29] Richard de Neufville and Stefan Scholtes. *Flexibility in Engineering Design*. The MIT Press, 2011.

[30] Urmila M. Diwekar. Greener by design. *Environmental Science and Technology*, 37:5432–5444, 2003.

[31] Avinash K. Dixit and Robert S. Pindyck. *Investment under Uncertainty*. Princeton University Press, 1994.

[32] Michael Duffy and Virginie Y. Nanhou. Costs of producing switchgrass for biomass in southern iowa. Technical report, Iowa State University - University Extension, http://www.extension.iastate.edu/Publications/PM1866.pdf, April 2001. `http://www.extension.iastate.edu/Publications/PM1866.pdf`. Accessed October 2012.

[33] Gary D. Eppen, R. Kipp Martin, and Linus Schrage. A scenario approach to capacity planning. *Operations Research*, 37:517–527, 1989.

[34] Peter C. Fishburn. Convex Stochastic Dominance with Continuous Distribution Functions. *Journal of Economic Theory*, 7:143–158, 1974.

[35] Peter C. Fishburn. Stochastic dominance and moments of distributions. *Mathematics of Operations Research*, 5(1):94–100, 1980.

[36] H. Scott Fogler. *Elements of Chemical Reaction Engineering*. Prentice Hall, Fourth edition, 2009.

[37] Itzhak Gilboa. *Rational Choice*. The MIT Press, 2010.

[38] Bo Gong. *Methodology for Technology Evaluation under Uncertainty and its Application in Advanced Coal Gasification Processes*. PhD thesis, Massachusetts Institute of Technology, 2011.

[39] John R. Graham and Campbell R. Harvey. The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics*, 60:187–243, 2001.

[40] R. L. Graham, R. Nelson, J. Sheehan, R. D. Perlack, and L. L. Wright. Current and potential u.s. corn stover supplies. *Agronomy*, 99:1–11, 2007.

[41] Ignacio E. Grossmann and Mark M. Daichendt. New trends in optimization-based approaches to process synthesis. *Computers & Chemical Engineering*, 20:665–693, 1995.

[42] M.H. Halabi, M. H. J. M. de Croon, J. van der Schaaf, P.D. Cobden, and J.C. Schouten. Modeling and analysis of autothermal reforming of methane to hydrogen in a fixed bed reformer. *Chemical Engineering Journal*, 137:568–578, 2007.

[43] Carlo N. Hamelinck, Andre P. C. Faajj, Herman den Uil, and Harold Boerrigter. Production of FT transportation fuels from biomass: technical options, process analysis and optimisation, and development potential. *Energy*, 29:1743–1771, 2004.

[44] Dunja Hoffmann and Martin Weih. Limitations and improvement of the potential utilisation of woody biomass for energy derived from short-rotation woody crops in sweden and germany. *Biomass and Bioenergy*, 28:267–279, 2005.

[45] William W. Hogan and James M. Warren. Computation of the efficient frontier in the e-s portfolio selection model. *Journal of Financial and Quantitative Analysis*, 7:1881–1896, 1972.

[46] Toshimitsu Homma and Andrea Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52:1–17, 1996.

[47] Kenneth T. Hu. *Bayesian Design of Experiments for Complex Chemical Systems*. PhD thesis, Massachusetts Institute of Technology, 2011.

[48] D. Humbird and A. Aden. Biochemical production of ethanol from corn stover: 2008 state of technology model. Technical Report NREL/TP-520-46214, National Renewable Energy Laboratory, August 2009. http://www.nrel.gov/biomass/pdfs/46214.pdf. Accessed September 30, 2012.

[49] E.T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.

[50] Jeremy Johnson. *Technology Assessment of Biomass Ethanol: A multi-objective, life cycle approach under uncertainty*. PhD thesis, Massachusetts Institute of Technology, June 2006.

[51] Philippe Jorion. *Value at risk: the new benchmark for controlling market risk.* Irwin Professional Pub., 1997.

[52] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, March 1979.

[53] M. Kasas, Z. Kravanja, and Z. Novak Pintaric. Suitable modelling for process flow sheet optimization using the correct economic criterion. *Industrial and Engineering Chemistry Research*, 50:3356 – 3370, 2011.

[54] Tim Keddy. Short-rotation woody crops for energy production. Technical report, Canadian Wood Fibre Center. http://www.gov.mb.ca/agriculture/agrienergy/pdf/20120308h.pdf - accessed October 2012.

[55] Ralph L. Keeney and Howard Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* Cambridge University Press, 1976.

[56] Madhu Khanna, Basanta Dhungana, and John Clifton-Brown. Costs of producing miscanthus and switchgrass for bioenergy in illinois. *Biomass and Bioenergy*, 32:482 493, 2008.

[57] Anton Alexandru Kiss. *Design and Control of Recycle Systems by Nonlinear Analysis.* PhD thesis, University of Amsterdam, 2004.

[58] Hiroshi Konno and Hiroaki Yamazaki. Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market. *Management Science*, 37:519 531, 1991.

[59] Kelly Launder. Energy crops and their potential development in michigan. Technical report, Michigan Biomass Energy Program, August 2002. http://www.michigan.gov/documents/eco-impact_89541_7.pdf. Accessed October 2012.

[60] William F. Lazarus. Minnesota crop cost and return guide for 2011. Technical report, University of Minnesota Extension, November 2010. http://faculty.apec.umn.edu/wlazarus/documents/cropbud.pdf. Accessed October 2012.

[61] Rocky Lemus and David J. Parrish. Herbaceous crops with potential for biofuel production in the usa. *Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 4(057):1 – 23, 2009.

[62] Moshe Leshno and Haim Levy. Preferred by "All" and Preferred by "Most Decision Makers: Almost Stochastic Dominance. *Management Science*, 48(8):1074–1085, August 2002.

[63] Haim Levy. Stochastic Dominance and Expected Utility: Survey and Analysis. *Management Science*, 38(4):555–593, April 1992.

[64] Haim Levy. *Stochastic Dominance: Investment Decision Making under Uncertainty.* Springer, 2nd edition, 2006.

[65] Andrey M. Lizyayev. *Stochastic Dominance in Portfolio Analysis and Asset Pricing.* PhD thesis, Erasmus University of Rotterdam, 2010.

[66] Andrey M. Lizyayev and Andrzej Ruszczynski. Tractable almost stochastic dominance. *European Journal of Operational Research*, 218:448–455, 2012.

[67] Weldon A. Lodwick. A Generalized Convex Stochastic Dominance Algorithm. *IMA Journal of Mathematics Applied in Business & Industry*, 2:225–246, 1990.

[68] Harry Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77–91, March 1952.

[69] Harry Markowitz. *Portfolio Selection.* Basil Blackwell, 2nd edition, 1991.

[70] Andreu Mas-Colell, Michael D. Winston, and Jerry R. Green. *Microeconomic Theory.* Oxford University Press, 1995.

[71] Warren L. McCabe, Julian C. Smith, and Peter Harriott. *Unit Operations of Chemical Engineering.* McGraw Hill, Sixth edition, 2001.

[72] Daniel McFadden. Testing for stochastic dominance. In Thomas B. Fomby and Tae Kun Seo, editors, *Studies in the Economics of Uncertainty.* Springer-Verlag, 1989.

[73] Peter Mckendry. Energy production from biomass (part 1): Overview of biomass. *Bioresource Technology*, 83:37 – 46, 2002.

[74] Samuel B. McLaughlin and Lynn Adams Kszos. Development of switchgrass (panicum virgatum) as a bioenergy feedstock in the united states. *Biomass and Bioenergy*, 28:515–535, 2005.

[75] Gregory J. McRae, James W. Tilden, and John H. Seinfeld. Global sensitivity analysis - a computational implementation of the fourier amplitude sensitivity test (fast). *Computers and Chemical Engineering*, 6:15 – 25, 1982.

[76] Jack Meyer. Choice Among Distributions. *Journal of Economic Theory*, 14:326–336, 1977.

[77] John M. Mulvey, Robert J. Vanderbei, and Stavros A. Zenios. Optimization of large-scale systems. *Operations Research*, 43(2):24–281, March - April 1995.

[78] Wlodzimierz Ogryczak and Andrzej Ruszczynski. From stochastic dominance to mean-risk models: Semi-deviations as risk measures. *European Journal of Operational Research*, 116:33–50, 1999.

[79] Wlodzimierz Ogryczak and Andrzej Ruszczynski. On consistency of stochastic dominance and mean-semideviation models. *Mathematical Programming*, 89:217–232, 2001.

[80] Wlodzimierz Ogryczak and Andrzej Ruszczynski. Dual stochastic dominance and quantile risk measures. *International Transactions in Operational Research*, 9:661–680, 2002.

[81] Wlodzimierz Ogryczak and Andrzej Ruszczynski. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13:60–78, 2002.

[82] Tony O'Hagan. Dicing with the unknown. *Significance*, 1(3):132–133, September 2004.

[83] Richard Perrin, Kenneth Vogel, Marty Schmer, and Rob Mitchell. Farm-scale production cost of switchgrass for biomass. *Bioenergy Resources*, 1:91–97, 2008.

[84] Martin Peterson. *An Introduction to Decision Theory*. Cambridge University Press, 2009.

[85] David Pimentel and Tad W. Patzek. Ethanol production using corn, switchgrass and wood; biodiesel production using soybean and sunflower. *Natural Resources Research*, 14:65–76, 2005.

[86] Z. Novak Pintaric and Z. Kravanja. Selection of the economic objective function for the optimization of process flow sheets. *Industrial and Engineering Chemistry Research*, 45:4222–4232, 2006.

[87] Mark J. Prins, Krzysztof J. Ptasinski, and Frans J.J.G. Janssen. From coal to biomass gasification: Comparison of thermodynamic efficiency. *Energy*, 32(7):1248–1259, July 2007.

[88] John Quiggin. A theory of anticipated utility. *Journal of Economic Behaviour and Organization*, 3:323–343, 1982.

[89] John Quiggin. *Generalized Expected Utility Theory: The Rank-Dependent Model*. Springer, 1993.

[90] Steven F. Rice and David P. Mann. Autothermal reforming of natural gas to synthesis gas. Technical Report KBR Paper 2031, Sandia National Laboratories, April 2007.

[91] R. Tyrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26:1443–1471, 2002.

[92] R.T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.

[93] Michael J. Rogers, Anshuman Gupta, and Costas D. Maranas. Real options based analysis of optimal pharmaceutical research and development portfolios. *Ind. Eng. Chem. Res*, 41:6607–6620, 2002.

[94] Stephen Ross. Introduction to financial economics. Fall 2011 Lecture Notes, Massachusetts Institute of Technology.

[95] Michael Rothschild and Joseph Stiglitz. Increasing Risk I: A Definition. *Journal of Economic Theory*, 2(3):225–243, September 1970.

[96] A.D. Roy. Safety first and the Holding of Assets. *Econometrica*, 20:431–449, 1952.

[97] Andrzej Ruszczynski and Robert J. Vanderbei. Frontiers of stochastically non-dominated portfolios. *Econometrica*, 71(4):1287–1297, July 2003.

[98] Andrea Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145:280–297, 2002.

[99] Andrea Saltelli, Karen Chan, and E. Marian Scott, editors. *Sensitivity Analysis*. John Wiley & Sons, 2000.

[100] Andrea Saltelli, Stefano Tarantola, Francesca Campolongo, and Marco Ratto. *Sensitivity Analysis in Practice: A guide to assessing scientific models*. John Wiley & Sons, 2004.

[101] Andrea Saltelli, Stefano Tarantola, and K. P-S. Chan. A quantitative, model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56, 1999.

[102] Ali Sayigh. *Comprehensive Renewable Energy Volume 5: Biomass and Biofuel Production*. Elsevier, 2012.

[103] Warren D. Seider, J.D. Seader, Daniel R. Lewin, and Soemantri Widagdo. *Product and Process Design Principles*. John Wiley & Sons, third edition, 2009.

[104] Derek Sidders and Rachel Clinch. Biomass opportunities: Short-rotation woody crops. Technical report, Canadian Wood Fibre Centre. http://www.cityofnorthbay.ca/business/presentations/woodPellets/Rachelle%20Clinch%20(CFS)%20Presentation.pdf - Accessed September 29, 2012.

[105] R.K. Sinnott. *Coulson & Richardson's Chemical Engineering*, volume 6. Butterworth-Heinemann, third edition, 2003.

[106] Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93:964–979, 2008.

[107] Ryan M. Swanson. Techno-economic analysis of biomass-to-liquids production based on gasification. Master's thesis, Iowa State University, 2009.

[108] Menner A. Tatang. *Direct Incorporation of Uncertainty in Chemical and Environmental Engineering Systems*. PhD thesis, Massachusetts Institute of Technology, February 1995.

[109] Pradeep J. Tharakan, Timothy A. Volk, Christopher A. Lindsey, Lawrence P. Abrahamson, and Edwin H. White. Evaluating the impact of three incentive programs on the economics of cofiring willow biomass with coal in new york state. *Energy Policy*, 33(3):337–347, February 2005.

[110] Jena Thompson and Wallace E. Tyner. Corn stover for bioenergy production: Cost estimates and farmer supply response. Technical Report RE-3-W, Purdue University - Purdue Extension, 2011. http://www.extension.purdue.edu/extmedia/EC/RE-3-W.pdf. Accessed September 2012.

[111] Michiel J. A. Tijmensen, Andre P. C. Faajj, Carlo N. Hamelinck, and Martin R. M. Harvdeveld. Exploration of the possibilities for production of fischer tropsch liquids and power via biomass gasification. *Biomass and Bioenergy*, 23:129 152, 2002.

[112] Richard Turton, Richard C. Baillie, Wallace B. Whiting, and Joseph A. Shaeiwitz. *Analysis, Synthesis, and Design of Chemical Processes*. Prentice Hall, third edition, 2009.

[113] Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. *Science*, 211(Issue 4481):453 – 458, Jan 1981.

[114] Amos Tversky and Daniel Kahneman. Rational choice and the framing of decisions. *The Journal of Business*, 59:S251 – S278, 1986.

[115] Daniel G. D. Ugarte, Marie E. Walsh, Hosein Shapouri, and Stephen B. Slinsky. The economic impacts of bioenergy crop production on u.s. agriculture. Technical report, The State of Michigan, 2000. http://www.michigan.gov/documents/eco-impact_89541_7.pdf. Accessed September 2012.

[116] Gael D. Ulrich. *A Guide to Chemical Engineering Process Design and Economics*. John Wiley & Sons, 1984.

[117] Karl T. Ulrich and Steven D. Eppinger. *Product Design and Development*. McGraw Hill, fifth edition, 2012.

[118] Tasneem Virani. The economic feasibility of bioenergy production from miscanthus for the ontario grreenhouse industry. Master's thesis, University of Guelph, 2011.

[119] Kenneth P. Vogel. Swithgrass for biomass energy: Status and progress. Technical report, USDA - Lincoln, Nebraska, 2007. http://purl.umn.edu/8030. Accessed Oct 10, 2012.

[120] T.A. Volk, L.P. Abrahamson, C.A. Nowak, L.B. Smart, P.J. Tharakan, and E.H. White. The development of short-rotation willow in the northeastern united states for bioenergy and bioproducts, agroforestry and phytoremediation. *Biomass and Bioenergy*, 30:715 727, 2006.

[121] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behaviour*. Princeton University Press, 1953.

[122] Peter Wakker. *Prospect theory*. Cambridge University Press, 2010.

[123] Joel Weisman and A.G. Holzman. Engineering design optimization under risk. *Management Science*, 19(3):235–249, 1972.

[124] Edwin H. White, Lawrence P. Abrahamson, and C. A Nowak. Commercialization of willow bioenergy. In *Proceedings, Second Biomass Conference of the Americas: Energy, Environment, Agriculture and Industry*, pages 1534–1546, August 1995.

[125] G. A. Whitmore. Third degree stochastic dominance. *The American Economic Review*, 60(3):457–459, June 1970.

[126] G.A. Whitmore. Stochastic dominance for the class of completely monotonic utility functions. In Thomas B. Fomby and Tae Kun Seo, editors, *Studies in the Economics of Uncertainty*, pages 77–90. Springer-Verlag New York, 1989.