OPTIMIZATION UNDER UNCERTAINTY:
SENSITIVITY ANALYSIS AND REGRET

By

CHIN HON TAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2012

I dedicate this dissertation to my parents.

ACKNOWLEDGMENTS

I would like to gratefully and sincerely thank Dr. Joseph C. Hartman for his patience, encouragement and support in my pursuit of problems that excite me.I would also like to thank my committee members Dr. Joseph P. Geunes, Dr. Jean Philippe Richard and Dr. Gary Koehler for their valuable comments that improved the content and presentation of this work. Last, but not least, I would like to thank my parents for their support and understanding throughout my graduate studies. This work is supported in part by the National Science Foundation under grant CMMI-0813671.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS 4

LIST OF TABLES 7

LIST OF FIGURES 8

ABSTRACT 9

CHAPTER

1 INTRODUCTION 11

2 SENSITIVITY ANALYSIS 14

2.1 Types of Sensitivity Analysis 14
2.2 Dynamic Programming and Markov Decision Processes 17
2.3 Linear Programs and Markov Decision Processes 21

3 SENSITIVITY ANALYSIS IN MARKOV DECISION PROCESSES WITH UNCERTAIN REWARD PARAMETERS 24

3.1 Stationary Rewards 26
3.1.1 Single Parameter Sensitivity Analysis 30
3.1.2 Tolerance Approach 33
3.2 Non-Stationary Rewards 34
3.3 Tolerance Gap 37

4 REGRET 43

4.1 Multivariate Stochastic Dominance 45
4.2 Regret in Normative Decision Analysis 48

5 EQUIPMENT REPLACEMENT ANALYSIS WITH AN UNCERTAIN FINITE HORIZON 53

5.1 Minimizing Expected Cost 56
5.2 Minimizing Maximum Regret 59
5.3 Leasing Options to Further Mitigate Risk 63

6 OPTIMAL PRODUCT LINE FOR CONSUMERS THAT EXPERIENCE REGRET 78

6.1 Model Description 82
6.1.1 Consumer Choice Model 83
6.1.2 Problem Formulation 84
6.2 Optimal Number of Products to Offer 86
6.2.1 Limiting Cases 88
6.2.2 Experimental Results 91

6.2.3 Upper Bound on $\mathbb{E}[|\pi^*|]$ 93

7 SUMMARY AND FUTURE RESEARCH 97

7.1 Sensitivity Analysis 97
7.2 Regret 99

REFERENCES 103

BIOGRAPHICAL SKETCH 113

LIST OF TABLES

Table		page
3-1	Single parameter sensitivity analysis and τ.	42
3-2	Computing υ.	42
5-1	Conditional probabilities of the various scenarios	70
5-2	Boundary conditions at $t = 8$	70
5-3	Boundary conditions at $t = 6$	70
5-4	Discounted costs through periods 6, 7 and 8 periods with an asset of age i	70
5-5	x^*, y^* and z^* values for policy (i)	70
5-6	x^*, y^* and z^* values for policy (ii)	70
5-7	Optimal policies: $P(T = 6) = 0.6$, $P(T = 7) = 0.2$, $P(T = 8) = 0.2$ and $n = 4$	71
5-8	Optimal policies: $P(T = 6) = 0.2$, $P(T = 7) = 0.2$, $P(T = 8) = 0.6$ and $n = 4$	72
5-9	Optimal policies: $P(T = 6) = 0.6$, $P(T = 7) = 0.2$, $P(T = 8) = 0.2$ and $n = 8$	73
5-10	Optimal policies: $P(T = 6) = 0.2$, $P(T = 7) = 0.2$, $P(T = 8) = 0.6$ and $n = 8$	74
6-1	Respective u_{ijs} values	95
6-2	Estimated expected optimal number of products to offer, $\hat{\mathbb{E}}[|\pi^*|$	96

LIST OF FIGURES

Figure page

5-1 Visualization of dynamic programming formulation 75

5-2 Discount rate and option price 76

5-3 Number of lease periods and option price 77

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

OPTIMIZATION UNDER UNCERTAINTY:
SENSITIVITY ANALYSIS AND REGRET

By

Chin Hon Tan

May 2012

Chair: Joseph C. Hartman
Major: Industrial and Systems Engineering

Solving for the optimal solution of any problem under uncertainty is generally challenging. This dissertation explores optimization under uncertainty from the perspective of sensitivity analysis and regret.

Sequential decision problems can often be modeled as Markov decision processes. Classical solution approaches assume that the parameters of the model are known. However, model parameters are usually estimated and uncertain in practice. As a result, managers are often interested in determining how estimation errors affect the optimal solution. We illustrate how sensitivity analysis can be performed directly for a Markov decision process with uncertain reward parameters using the Bellman equations. In particular, we consider problems involving (i) a single stationary parameter, (ii) multiple stationary parameters and (iii) multiple non-stationary parameters. We illustrate the applicability of this work on a capacitated stochastic lot-sizing problem.

In sensitivity analysis, we study the stability or robustness of an optimal solution with respect to uncertainties in the model. If the optimal solution is constant across all possible scenarios, the uncertainties in the model parameters can be ignored. However, these uncertainties need to be addressed if the optimal solution differs under different possible scenarios.

Research in psychology and behavioral decision theory suggests that regret plays an important role in shaping preferences in decisions under uncertainty. Despite their

relevance, feelings of regret are often ignored in the optimization literature and the representation of regret, when present, is limited. Regret theory describes choice preferences based on the rewards received and opportunities missed. We show that regret-theoretic choice preferences are described by multivariate stochastic dominance, present regret-based risk measures and illustrate how they can be adopted within the mean-risk framework.

Research also suggest that people are willing to exchange direct material gain for regret reduction. We consider an equipment replacement problem under horizon uncertainty. We present stochastic dynamic programming formulations and explore solutions which minimize either expected costs or maximum regret. We identify the critical time period where optimal decisions diverge for different horizon realizations and design a lease option contract such that owners can lower the regret that may result from a given horizon realization, while opening a possible source of revenue for a leasor.

Finally, we also study the optimal number of products to offer under various conditions in a heterogeneous market. In the absence of regret, consumers are happier when presented with more choices and a company that wishes to capture a broad market share needs to provide a rich product line. However, the relationship between the optimal number of products and the number of market segments are reversed when consumers are regret averse. In general, the product line should be narrow when outcomes are uncertain and consumers experience regret. In addition, it is almost surely optimal for the firm to offer a single product when the outcomes of choices are highly uncertain and/or consumers are highly regret averse. We also show that the optimal number of products to offer is non-increasing when the cost of introducing variety into the product line increases uniformly and obtain a tight upper bound on the expected optimal number of products to offer when fixed costs are uniform. However, interestingly, the optimal number of products to offer can increase when regret aversion increases.

CHAPTER 1
INTRODUCTION

For more than half a century, operations researchers have employed various optimization tools, including mathematical programming and dynamic programming, to solve a variety of problems. Advances in optimization techniques and computing power have allowed operations researchers to tackle problems at a scale previously unimaginable. Despite these advances, solving for the optimal solution under uncertainty is generally still very challenging. This dissertation explores optimization under uncertainty from the perspective of sensitivity analysis and regret.

In sensitivity analysis, we study the stability or robustness of an optimal solution with respect to uncertainties in the model. If the optimal solution is constant across all possible scenarios, the uncertainties in the model parameters can be ignored. However, these uncertainties need to be addressed if the optimal solution differs under different possible scenarios.

When the optimal solution differs across different scenarios, the decision maker may seek solutions that maximize expected rewards received (or minimize expected costs incurred). A solution that maximizes expected rewards outperforms the other solutions in the long run. However, such solutions may not be appropriate for single-event problems. In addition, such solutions do not account for the risk preference of the decision maker. Krokhmal et al. (2011) provides an extensive review on different approaches to optimization under risk in mathematical programming models. In addition, researchers have highlighted that the probabilities of possible outcomes may not be available. Under this condition, researchers propose finding robust solutions that are insensitive to uncertainties in the model. The interested reader is referred to Beyer and Sendhoff (2007) and Bertsimas et al. (2011) for a survey on robust optimization. In this dissertation, we focus on the use of Markov decision processes (MDPs) to solve sequential decision problems under risk.

Research in psychology and economics suggests that satisfaction is influenced by feelings of regret (Connolly and Zeelenberg 2002). Furthermore, anticipated regret can affect the decision making process (see Zeelenberg (1999) and Engelbrecht-Wiggans and Katok (2009)). For example, managers who are concerned about the opinions of shareholders and superiors may be more interested in decisions that minimize regret resulting from lost opportunities, rather than options that maximize profits. Despite their relevance, feelings of regret are often ignored in the optimization literature. In this dissertation, we identify the gaps, illustrate the importance of regret consideration by showing how regret can drastically change the optimal solution and highlight promising areas of future research.

In Chapter 2, we discuss different concepts of sensitivity analysis and review the literature on sensitivity analysis in dynamic programs and MDPs. In addition, we highlight the relationship between linear programs and MDPs and highlight the need for sensitivity analysis approaches strictly for MDPs. In Chapter 3, we illustrate how sensitivity analysis can be conducted for a MDP with uncertain reward parameters. A single parameter approach and a tolerance approach are proposed for problems involving single and multiple uncertain reward parameters, respectively. In addition, we extend the tolerance approach to problems involving non-stationary rewards and investigate the sensitivity of optimal solutions for different problems.

In Chapter 4, we give a brief overview of regret theory. We obtain necessary and sufficient conditions for regret-theoretic choice preferences and illustrate how they can be used to aid the decision making process by eliminating necessarily inferior solutions. In addition, we illustrate how regret can be viewed as a measure of risk and be adopted within the mean-risk framework. We consider an equipment replacement problem where the horizon is uncertain in Chapter 5. We highlight that the policy that minimizes expected cost and the policy that minimizes maximum regret can differ. Van de Ven and Zeelenberg (2011) recently showed that people are willing to exchange direct material gain for regret reduction. We illustrate how an option contract, which allows the decision maker to lease the equipment across the uncertain horizon at a favorable rate, can reduce regret for the decision maker and generate revenue for the leasor. In Chapter 6, we consider the effects of regret on consumer behavior. We present a choice model where consumer satisfaction is affected by feelings of regret but choice preferences are independent of anticipated regret. We obtain theoretical and experimental results for the optimal number of products to offer in a heterogeneous market under our proposed choice model. In particular, we show that the relationship between the optimal number of products and the number of market segments are reversed when regret intensity is high and outcomes are variables.

We conclude with a summary and list future research directions in Chapter 7.

CHAPTER 2
SENSITIVITY ANALYSIS

The validity of any solution depends on the accuracy of the model and the data for a given instance. However, the parameters of the model are often uncertain and estimated in practice. Hence, the robustness or stability of the solution with respect to changes in the model parameters is of interest. One approach is to solve the problem for different parameters and try to infer the relationship between the model parameters and the optimal solution from the resulting solutions (see, for example, Puumalainen (1998) and Tilahun and Raes (2002)). However, this approach has two drawbacks. First, when the size of the problem or the number of scenarios to analyze is large, this "brute force" approach can be very time-consuming and may not be practical in practice. Second, unless the properties regarding the relationship between the model parameters and the optimal solutions are known (for example, continuous, monotonic, etc.), the solution for a set of parameter values may not provide insights on the solution for another set of parameter values. For example, suppose that ρ is some model parameter. One cannot guarantee that the optimal objective value for $\rho = 2$ must be between the optimal objective values for $\rho = 1$ and $\rho = 3$. From a theoretical and practical perspective, more sophisticated approaches for conducting sensitivity analysis are desired.

2.1 Types of Sensitivity Analysis

In this section, we present various concepts of sensitivity analysis in the optimization literature through a multi-stage problem. Let S and Π denote the set of states and the set of policies, respectively. Let $A(s)$ denote the set of actions available with $s \in S$. We consider problems of the following form:

$$V^{\pi} = \sum_{t=0}^{T} \gamma^t r^{\pi(s_t^{\pi})}, \tag{2–1}$$

where V^{π} is the value function of the system under policy $\pi \in \Pi$ through the horizon T. T can be either finite or infinite. $s_t^{\pi} \in S$ is the state of the system at time t under

π and $\pi(s) \in A(s)$ is the action that is taken at state s under π. Note that the action that is selected depends solely on the state and is independent of t. This stationary policy assumption is not restrictive since additional states can always be declared to address the dependency of π on t. r^a is the reward associated with action a and γ is the per-period discount factor. If the state transitions are stochastic, s_t^π is random and the value function is expressed as:

$$V^\pi = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t r^{\pi(s_t^\pi)}\right]. \tag{2–2}$$

Let π^* denote an optimal policy and V denote the optimal value function. The objective is to find a policy that maximizes the value function,

$$V = V^{\pi^*} = \max_{\pi \in \Pi} V^\pi. \tag{2–3}$$

We are interested in how V and π^* change when model parameters vary. We first consider the case where the variations in the model parameters can be expressed by a single parameter ρ (i.e., univariate). For example, suppose that γ is allowed to vary between 0.85 and 0.95. We model this by setting $\gamma = 0.9 + \rho$, where $\rho \in [-0.05, 0.05]$. This modeling approach allows for dependencies between the model parameters and is particularly useful when modeling simple variations in transition probabilities, where the sum of probabilities must add to one.

Suppose that the problem is solved for $\rho = \rho^{(0)}$ and the optimal value function and optimal policy, $V^{(0)}$ and $\pi^{(0)}$, are obtained. Sensitivity analysis aims to answer one or more of the following questions:

1. What is $\frac{dV^{\pi^{(0)}}}{d\rho}|_{\rho=\rho^{(0)}}$?
2. What is the range of ρ values for which $\pi^{(0)}$ is optimal?
3. What is V if $\rho = \rho^{(0)} + \delta$, for some $\delta \in \mathbb{R}$?
4. What is π^* if $\rho = \rho^{(0)} + \delta$, for some $\delta \in \mathbb{R}$?

Questions 1 and 2 are specific to $\pi^{(0)}$ and can only be answered after $\pi^{(0)}$ has been identified. Hence, these questions are addressed in a *posterior analysis* (see, Fernandez-Baca and Venkatachalam (2007)) or *postoptimal analysis* (see, Sotskov et. al. (1995) and Wallace (2000)). The answers to questions 3 and 4 do not require advanced knowledge of $V^{(0)}$ and $\pi^{(0)}$. However $V^{(0)}$ and $\pi^{(0)}$ can often be used to obtain or approximate V and π^* (see Section 2.2) and knowledge of $V^{(0)}$ and $\pi^{(0)}$ may provide insights to the answers of questions 3 and 4.

Besides posterior analysis, researchers have also considered *prior analysis* where knowledge of $V^{(0)}$ and $\pi^{(0)}$ is not required. For example, inverse parametric optimization aims to find the ρ value where a policy is optimal or good (see, Eppstein (2003), Sun et al. (2004) and Kim and Kececioglu (2007)). Prior analysis will not be discussed but the interested reader is referred to Fernandez-Baca and Venkatachalam (2007) for a discussion on this topic. In addition to uncertainties in the model parameters, researchers have also pointed out that the model is generally an approximation of the actual problem and hence there is doubt on the structure of the model itself. Literature that address model uncertainties include Briggs et al. (1994), Chatfield (1995), Draper (1995) and Walker and Fox-Rushby (2001).

The analysis can also be classified as being either *local* or *global* (McKay, 1979). Question 1 takes a local perspective and is concerned with changes in the ε-neighborhood of $\rho^{(0)}$. Wagner (1995) considers question 2 to be local as well as it is concerned with the properties of $\pi^{(0)}$. In contrast, questions 3 and 4 explore variations across a wider range of ρ values and possibly different π^*.

In the univariate problem, the stability of $V^{(0)}$ and $\pi^{(0)}$ are addressed by questions 1 through 4. However, the notion of stability is more complicated in the multivariate problem (i.e., multiple parameters). The typical approach to sensitivity analysis for a multivariate problem is to vary one parameter at a time while holding the other parameters constant. This is sometimes referred to as *ordinary sensitivity analysis* (see Wendall (2004) and Filippi (2005)). However, ordinary sensitivity analysis does not allow for simultaneous variations of different parameters. Bradley et al. (1977) provided restrictions on the variations in the right-hand-side (RHS) terms and objective-function coefficients to guarantee the optimality of the current optimal basis (i.e., optimal decision variables) in a linear program. They term these restrictions as the *100 percent rule*. Wendell (1985) was interested in the range that each parameter was allowed to vary without violating optimality. His approach, which he termed the *tolerance approach*, is different from ordinary sensitivity analysis in that simultaneous variations of different parameters are considered. The tolerance approach has been applied to different problems by various researchers (see, for example, Wendell (1985), Hansen et al. (1989), Ravi and Wendell (1989), Wondolowski (1991) and Filippi (2005)).

2.2 Dynamic Programming and Markov Decision Processes

Dynamic programming (DP) has been successfully applied to a wide range of problems in a variety of fields (see, for example, Held and Karp (1962), Pruzan and Jackson (1967), Elton and Gruber (1971) and Yakowitz (1982)) since it was proposed by Richard Bellman in the middle of the last century (Bellman 1952). One of the main advantages of DP lies in its ability to account for the dynamics of a system and it is often used by operations researchers to solve problems involving sequential decisions. In addition, unlike most optimization techniques, DP is able handle a wide range of cost and/or reward functions. Given its versatility, it is not surprising that DP is used across various disciplines, including artificial intelligence, control, economics and operations research. However, DP does have its drawbacks as well. The size of a dynamic program grows rapidly with the dimension of the problem. This is commonly referred to as the curse of dimensionality and has been the main challenge in applying DP approaches to solving real world problems. In light of the computational difficulties that arise from the curse of dimensionality, researchers have proposed a variety of approximation approaches, including state aggregation and value function approximation, to solve large dynamic programs. The interested reader is referred to Powell (2007) for a detailed presentation of approximate DP.

A dynamic program comprises a set of states S, with a set of decisions $A(s)$, available at each $s \in S$. Each $a \in A(s)$ transits the system from a state s to another state s' and results in a reward of r^a. In the deterministic problem (Equation (2–1)), the state transitions are known with certainty. In the stochastic problem, the state transitions are uncertain (Equation (2–2)) and the probability of transiting from state s to state s' under $a \in A(s)$ is denoted by $P^a(s')$. The state transitions are Markovian (i.e., only dependent on the state s' and action a) and stochastic dynamic programs are often referred to as Markov decision processes, or MDPs. We focus on the stochastic problem. However, we note that the discussion and results also apply to deterministic DP since the deterministic problem is merely a special case of the stochastic problem.

Most textbooks divide MDPs into two broad categories: i) finite horizon and ii) infinite horizon (see, White (1993), Puterman (1994) and Powell (2007)). Finite horizon problems are usually modeled as acyclic graphs where each node corresponds to a particular s at a particular t. The common solution approach for finite horizon problems is to define the boundary conditions $V_T(s)$ for each state at time T and recursively solve for V in a backward manner:

$$V_{t-1}(s) = \max_{a \in A(s)} \left\{ r^a + \gamma \sum_{s' \in S} P^a(s') V_t(s') \right\} \quad t = 0, 1, ..., T-1, \ \forall s \in S, \tag{2–4}$$

where $V_t(s)$ is the value function of state s at time t. Boundary conditions do not exist for infinite horizon problems. Infinite horizon problems are expressed as cyclic graphs and value and/or policy iteration approaches can be used to obtain the optimal solution. These approaches are based on the optimality conditions that are common referred to as the Bellman equations (see, Puterman (1994)):

$$V(s) = \max_{a \in A(s)} \left\{ r^a + \gamma \sum_{s' \in S} P^a(s') V(s') \right\} \quad \forall s \in S. \tag{2–5}$$

It is clear from Equations (2–1) to (2–3) that V is continuous in γ. Altman and Shwartz (1991) showed that V remains continuous in γ in a constrained MDP. In addition, it also follows from Equations (2–1) to (2–3) that V is continuous and monotone with respect to each individual reward. However, the rewards may not be independent from each other. For example, the cost of producing 10 items may be twice the cost of producing 5 items. In these cases, we are interested in the changes in the optimal solution with respect to the unit cost of production, rather than the cost of producing 10 items specifically. In these problems, the rewards are expressed as:

$$r^a = c^a(\boldsymbol{\rho}),$$

where $\boldsymbol{\rho}$ is a vector that represents the parameters that are of interest and $c^a : P \to \mathbb{R}$, where P is the set containing $\boldsymbol{\rho}$. White and El-Deib (1986) showed that if r^a and $V_T(s)$ are affine in $\boldsymbol{\rho}$ for all $a \in A$ and $s \in S$, the optimal value function of each state, including V, is piecewise affine and concave in $\boldsymbol{\rho}$.

Modeling the uncertainties in the transition probabilities is complicated by the additional requirement that transition probabilities must add to one. Researchers have modeled the uncertainties in the transition probabilities in a variety of ways (see, for example, Satia and Lave (1973), White and Eldeib (1994), Givan et al. (2000), Kalyanasundaram et al. (2002) and Iyengar (2005)). Muller (1997) showed that the optimal value function is monotone and continuous when appropriate stochastic orders and probability metrics are used.

Clearly, V is dependent on T. In particular, if all the rewards are non-negative (non-positive), V will be monotonically non-decreasing (non-increasing) in T. However, the former is not a necessary condition and it is possible to prove the monotonicity of V for certain problems, even if the rewards differ in sign. For example, Tan and Hartman (2010) showed that the cost of owning and operating equipment is non-decreasing in T

if the discount factor and maintenance cost are non-negative and the salvage value is non-increasing with age.

Although the value of V is sensitive to changes in the model parameters, their impact on π^* is not as apparent. In fact, the optimal policy is often robust to minor deviations in the model parameters. White and El-Deib (1986) considered a MDP with imprecise rewards and were interested in finding the set of policies that are optimal for some realization of the imprecise parameters. They provided a successive approximation procedure for an acyclic finite horizon problem and a policy iteration procedure for the infinite horizon problem. Harmanec (2002) computed the set of non-dominated policies for a finite horizon problem with imprecise transition probabilities by generalizing the classical backward recursion approach (Equation (2–4)).

Researchers have highlighted that the current decision is the only decision that the decision-maker has to commit to in many cases and hence he or she might be solely interested in the optimality of the current decision (i.e., decision at time zero). Hopp (1988) was interested in the tolerance on the value functions of the later stages. In particular, he derived lower bounds on the maximum allowable perturbations in the state value of the future stages such that the time zero optimal decision remain unchanged. His numerical results indicate that the tolerance of the value functions (i.e., allowable perturbation) increases geometrically with time. In addition, researchers have also proposed finding a forecast horizon such that the optimal time zero decision is not affected by data beyond that horizon. Researchers who have studied this problem include Bean and Smith (1984), Hopp (1989), Bean et al. (1992) and Cheevaprawatdomrong et al. (2003).

In most cases, it is not necessary to re-solve the problem when the model parameters change. Very often, the new π^* and V can be efficiently computed or approximated from the current solution. For example, the optimal solutions to shortest path problems can be easily revised when the arc lengths change (Ahuja et al., 1993).

Topaloglu and Powell (2007) provided an efficient approach to approximate the change in the optimal value function for a dynamic fleet management model. Their approach was based on the special structure that resulted from their value function approximation. For MDP problems with no special structure, the Bellman equation (Equation (2–5)) can be used to verify the optimality of the current solution. If the optimality conditions are violated, the policy iteration approach can be used to determine the new solution.

2.3 Linear Programs and Markov Decision Processes

There is a rich literature on sensitivity analysis for linear programs (see Gal and Greenberg (1997) and Bazaraa et al. (2005)). In addition, it has been long recognized that a MDP can be formulated and solved as a linear program. The LP method for MDPs was first proposed by Manne (1960) and elaborated upon by Derman (1962) and Puterman (1994). It can be shown that a MDP can always be formulated by the following linear program (Powell, 2007):

$$\begin{aligned} \min \sum_{s \in S} & \quad v_s & & \\ \text{s.t. } v_s - \gamma \sum_{s' \in S} P^a(s') v_{s'} & \geq r^a & & \forall s, a \\ v_s & \in \mathbb{R} & & s = 1, 2, \dots, |S|. \end{aligned}$$

The optimal value function V corresponds to $v^*_{s_0}$, where s_0 is the current state (i.e. state at time 0) and π^* is defined by the constraints that are binding at optimality. Manne's observation has an important theoretical implication. It highlights that an optimal solution of a dynamic program will have the same properties as an optimal solution to its associated linear program. This is particularly insightful as sensitivity analysis is well-established in LP.

For every linear program, which we will refer to as the primal problem, there is another associated linear program, which is commonly referred to as the dual problem. The relationship between the two problems provides insight on the stability of the solution in the primal problem. For each constraint in the primal problem, there is an associated variable in the dual problem. The latter is commonly referred to as the dual variable. The complementary slackness theorem, which was first proposed by Goldman and Tucker (1956) and extended by Williams (1970), highlights the relationship between the variables and constraints of the primal and dual problems at optimality. In particular, if a variable in one problem is positive, the corresponding constraint in the other problem must be binding. If the constraint is not binding, the corresponding variable in the other problem must be zero. The values of the optimal dual variables, also known as shadow prices, correspond to the marginal change in the objective value when the RHS of the associated constraint is perturbed. This result can be applied directly to DP problems where the rewards of individual actions are independent. If a small perturbation to a reward changes some state value function v_s, it follows from the complementary slackness theorem that the associated action must be optimal. If the action is not optimal, its associated reward can be perturbed by a small amount without changing the value of any v_s.

The complementary slackness theorem and economic interpretation of the shadow prices that are presented are only valid when the rewards are independent of each other. When the rewards are dependent, the individual rewards are expressed as functions of independent parameters (Equation (2–5)) and parametric analysis approaches can be used to analyze the sensitivity of the solution with respect to these independent parameters (see Ward and Wendell (1990), Gal and Greenberg (1997) and Bazaraa et al. (2005)).

Although LP-based sensitivity analysis approaches can be applied to MDPs, there are a couple of reasons for a separate study of sensitivity analysis in MDPs. First, the number of constraints in the LP formulation of a MDP is large (Powell, 2007) and hence MDPs are rarely formulated and solved as linear programs in practice. Second, the type of sensitivity analysis that is required for the two problem classes can differ. For example, the analyst may be interested in the optimality of the current decision (i.e., decision at time zero) rather than the optimality of a basis in the linear program.

In the next chapter, we illustrate how the marginal change in the objective and the tolerances of the individual uncertain parameters can be obtained directly from the well-known Bellman equations. Our proposed approach is general and can be applied to problems involving multiple non-stationary reward parameters.

CHAPTER 3
SENSITIVITY ANALYSIS IN MARKOV DECISION PROCESSES WITH UNCERTAIN REWARD PARAMETERS

The Markov decision process (MDP) framework has been used by researchers to model a variety of sequential decision problems because it can account for the dynamics of a complex system and handle a wide range of reward functions. A MDP is defined by a set of states, with a set of potential actions associated with each state. Classical solution approaches assume that the parameters of the model, including rewards, transition probabilities and the discount factor, are known (see Powell (2007), Puterman (1994) and White (1993)). However, these are often estimated and uncertain in practice. For example, it is difficult to quantify the cost of not having an item in the store upon the arrival of a customer (stock-out cost).

White and El-Deib (1986) identified optimal policies for some realization of the imprecise parameters, termed non-dominated policies, for a MDP with imprecise rewards. Harmanec (2002) studied a similar problem where the imprecision was defined in the transition probabilities, rather than the rewards. However, a decision maker can only implement a single policy in practice. One approach is to assume that the imprecision is resolved in the most pessimistic scenario (see Iyengar (2005) and Nilim and El Ghaoui (2005)). This is often referred to as the max-min policy. However, it has been highlighted that max-min policies can be overly conservative and may not be practical in reality (Wallace, 2000).

Managers are often interested in how an optimal solution changes with deviations in the model parameters. The typical approach to answering this question is to solve the problem for different values of the uncertain parameter, but this can be very time-consuming when the problem is large. For example, Topaloglu and Powell (2007) were interested in the benefits of adding an extra vehicle or load in a dynamic fleet management model. Sandvik and Thorlund-Petersen (2010) were interested in the conditions where there is at most one critical risk tolerance value, such that the knowledge of whether the individual's risk tolerance is above or below that value is sufficient for identifying the preferred decision.

In this chapter, we consider a MDP where rewards are expressed as affine functions of uncertain parameters. Problems of this form abound in the MDP literature including, the lot-sizing problem (Muckstadt and Sapra, 2010), the equipment replacement problem (Tan and Hartman, 2010), the sequential search problem (Lim et al., 2006) and various resource allocation problems (see, for example, Erkin et al. (2010), Charalambous and Gittins (2008) and Glazebrook et al. (2004)). Bounds on the perturbations in the state values for a given policy are computed in Mitrophanov et al. (2005). We are interested in the maximum range parameters are allowed to vary such that a policy remains optimal. Hopp (1988) derived bounds on the minimum perturbations in the future state values required to change the current optimal decision (i.e., at time 0) and extended the results to perturbations in the rewards at each state. Our model allows for dependencies between the uncertainties in the rewards associated with different actions and states. Another important difference is that we are not deriving bounds, but computing the actual range of values our parameters are allowed to vary.

A single parameter analysis provides insight on the stability of the solution with respect to a particular parameter. However, estimation errors can exist for multiple parameters. Wendell (1985) proposed finding a tolerance level which indicates the maximum percentage parameters are allowed to vary from their base value such that the optimal basis of a linear program remains optimal. We illustrate how the maximum tolerance can be obtained for our MDP when multiple uncertain parameters are allowed to vary simultaneously. In addition, we allow these parameters to be non-stationary.

First, we obtain the range in which a single parameter and multiple parameters are allowed to vary while maintaining the optimality of the current solution (Propositions 3.1 and 3.2). Second, we illustrate how the maximum allowable tolerance can be computed when uncertain parameters are non-stationary (Proposition 3.3) and show that it cannot be greater than the allowable tolerance of the stationary problem (Theorem 3.3). Third, we derive the conditions where the tolerances of the stationary and non-stationary rewards problem are the same (Corollary 2) and the conditions where they differ (Theorem 3.4). In particular, we show that, under mild assumptions, the tolerances of lot-sizing problems with uncertain ordering costs and backlog penalties differ when the maximum allowable tolerance is associated with an action that changes the reorder point (Theorem 3.5).

In the next section, we describe our stationary rewards model and illustrate how single parameter sensitivity analysis can be performed for this problem. In addition, we demonstrate how the maximum allowable tolerance can be computed when the uncertain parameters are allowed to vary simultaneously. Next, we illustrate how the maximum allowable tolerance can be computed for the non-stationary rewards problem. In Section 3.3, we study and discuss the difference in the maximum allowable tolerance of the two problems.

3.1 Stationary Rewards

Consider a finite state, finite action, infinite horizon MDP. Let S and $A(s)$ denote the set of states and the set of actions available with $s \in S$, respectively. Each $a_s \in A(s)$ transitions the system from state s to state s' with probability $P^{a_s}(s')$. Let $\tilde{r}^{a_s}$ denote the reward associated with a_s, expressed as an affine function of N uncertain parameters:

$$\tilde{r}^{a_s} = \lambda_0^{a_s} + \boldsymbol{\lambda}^{a_s}\tilde{\mathbf{x}},$$

where $\lambda_0^{a_s}$ is some known constant, $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N)'$ represents the uncertain parameters and $\boldsymbol{\lambda}^{a_s} = (\lambda_1^{a_s}, \lambda_2^{a_s}, \dots, \lambda_N^{a_s})$ the respective known coefficients. We assume that an estimation of $\tilde{x}_i$ is available for $i = 1, 2, \dots, N$. Let $\mathbf{x} = (x_1, x_2, \dots, x_N)'$ denote the vector of estimated parameter values. In addition, let r^{a_s} denote the estimated reward associated with a_s:

$$r^{a_s} = \lambda_0^{a_s} + \boldsymbol{\lambda}^{a_s}\mathbf{x}.$$

Let $\boldsymbol{\rho} = (\rho_1, \rho_2, \dots, \rho_N)'$ denote the corresponding estimation error, defined as follows:

$$\rho_i = \frac{\tilde{x}_i - x_i}{x_i},$$

and $\boldsymbol{\Delta}^{a_s} = (\Delta_1^{a_s}, \Delta_2^{a_s}, \dots, \Delta_N^{a_s})$ the corresponding coefficient:

$$\Delta_i^{a_s} = \lambda_i^{a_s} x_i,$$

such that $\tilde{r}^{a_s}$ can be re-expressed as follows:

$$\tilde{r}^{a_s} = r^{a_s} + \boldsymbol{\Delta}^{a_s}\boldsymbol{\rho}.$$

As in sensitivity analysis, we are interested in the stability of the solution obtained using the estimated parameters $\mathbf{x}$. In this section, we obtain the relationship between the estimation errors and the total reward received. In addition, we compute the range of error values where the current solution remains optimal. Here, we consider a problem where the value of $\boldsymbol{\rho}$ is uncertain but stationary. Since the rewards are stationary (i.e., do not vary with time), there must exist a stationary optimal policy where the action is determined solely by the state of the process (Puterman 1994). Let γ and Π denote the periodic discount factor and the set of all possible stationary policies, respectively. We assume that r^{a_s} is bounded and $\gamma < 1$ to ensure that the value function is finite. Let $\pi(s) \in A(s)$ denote the action that is taken at state s under $\pi \in \Pi$ and let $V_s^{\pi}(\boldsymbol{\rho})$ denote the value function of state s under policy π for a given $\boldsymbol{\rho}$. The value function of a state can be expressed by the following recursive equation:

$$V_s^{\pi}(\boldsymbol{\rho}) = r^{\pi(s)} + \boldsymbol{\Delta}^{\pi(s)}\boldsymbol{\rho} + \gamma \sum_{s' \in S} P^{\pi(s)}(s') V_{s'}^{\pi}(\boldsymbol{\rho}) \qquad \forall s, \pi. \tag{3–1}$$

Note that $V_s^{\pi}(\boldsymbol{\rho})$ depends on the value functions of the other states. Hence, it is convenient to express Equation (3–1) in matrix form.

Let $\mathbf{V}^{\pi}(\boldsymbol{\rho})$ and $\mathbf{r}^{\pi} = \left(r^{\pi(1)}, r^{\pi(2)}, \cdots, r^{\pi(|S|)}\right)'$ denote a vector of state values and rewards, respectively. In addition, let $\boldsymbol{\Delta}^{\pi}$ and $\mathbf{P}^{\pi}$ denote a matrix of uncertain reward parameters and transition probabilities, respectively:

$$\mathbf{V}^{\pi}(\boldsymbol{\rho}) = \left(V_1^{\pi}(\boldsymbol{\rho}), V_2^{\pi}(\boldsymbol{\rho}), \cdots, V_{|S|}^{\pi}(\boldsymbol{\rho})\right)',$$

$$\mathbf{r}^{\pi} = \left(r^{\pi(1)}, r^{\pi(2)}, \cdots, r^{\pi(|S|)}\right)',$$

$$\boldsymbol{\Delta}^{\pi} = \begin{pmatrix} \Delta_1^{\pi(1)} & \Delta_2^{\pi(1)} & \cdots & \Delta_N^{\pi(1)} \\ \Delta_1^{\pi(2)} & \Delta_2^{\pi(2)} & \cdots & \Delta_N^{\pi(2)} \\ \vdots & \ddots & \ddots & \vdots \\ \Delta_1^{\pi(|S|)} & \Delta_2^{\pi(|S|)} & \cdots & \Delta_N^{\pi(|S|)} \end{pmatrix}$$

and

$$\mathbf{P}^{\pi} = \begin{pmatrix} P^{\pi(1)}(1) & P^{\pi(1)}(2) & \cdots & P^{\pi(1)}(|S|) \\ P^{\pi(2)}(1) & P^{\pi(2)}(2) & \cdots & P^{\pi(2)}(|S|) \\ \vdots & \ddots & \ddots & \vdots \\ P^{\pi(|S|)}(1) & P^{\pi(|S|)}(2) & \cdots & P^{\pi(|S|)}(|S|) \end{pmatrix}.$$

$\mathbf{V}^{\pi}(\boldsymbol{\rho})$ can be expressed as:

$$\begin{aligned} \mathbf{V}^{\pi}(\boldsymbol{\rho}) &= \mathbf{r}^{\pi} + \boldsymbol{\Delta}^{\pi}\boldsymbol{\rho} + \gamma\mathbf{P}^{\pi}\mathbf{V}^{\pi}(\boldsymbol{\rho}) \\ (\mathbf{I} - \gamma\mathbf{P}^{\pi})\,\mathbf{V}^{\pi}(\boldsymbol{\rho}) &= \mathbf{r}^{\pi} + \boldsymbol{\Delta}^{\pi}\boldsymbol{\rho} \\ \mathbf{V}^{\pi}(\boldsymbol{\rho}) &= (\mathbf{I} - \gamma\mathbf{P}^{\pi})^{-1}(\mathbf{r}^{\pi} + \boldsymbol{\Delta}^{\pi}\boldsymbol{\rho}) \\ \mathbf{V}^{\pi}(\boldsymbol{\rho}) &= (\mathbf{I} - \gamma\mathbf{P}^{\pi})^{-1}\mathbf{r}^{\pi} + (\mathbf{I} - \gamma\mathbf{P}^{\pi})^{-1}\boldsymbol{\Delta}^{\pi}\boldsymbol{\rho} \\ \mathbf{V}^{\pi}(\boldsymbol{\rho}) &= \mathbf{V}^{\pi}(\mathbf{0}) + (\mathbf{I} - \gamma\mathbf{P}^{\pi})^{-1}\boldsymbol{\Delta}^{\pi}\boldsymbol{\rho}. \end{aligned} \tag{3–2}$$

Let $\mathbf{V}(\boldsymbol{\rho}) = \max_{\pi}\mathbf{V}^{\pi}(\boldsymbol{\rho})$. For a given $\boldsymbol{\rho}$ (including $\boldsymbol{\rho} = \mathbf{0}$), the policy that maximizes $\mathbf{V}^{\pi}(\boldsymbol{\rho})$ can be obtained through value and/or policy iteration approaches (Puterman 1994). Let $\widetilde{\pi}$ denote a policy that maximizes $\mathbf{V}^{\pi}(\mathbf{0})$. It follows from Equation (3–2) that, within the region where $\widetilde{\pi}$ is the optimal policy, the marginal change in $\mathbf{V}(\boldsymbol{\rho})$ is $\left(\mathbf{I} - \gamma\mathbf{P}^{\widetilde{\pi}}\right)^{-1}\boldsymbol{\Delta}^{\widetilde{\pi}}$.

In linear programs, sensitivity analysis is performed by deriving a set of necessary and sufficient conditions for optimality based on the reduced cost of each variable and finding the range of values for which these conditions hold Bazaraa et al. (2005). In theory, MDPs can be formulated as linear programs (Manne, 1960) and the allowable ρ values can be obtained by applying results from parametric linear programming (see Gal and Greenberg (1997) or Ward and Wendell (1990)) on the dual of the associated linear program (Tan and Hartman, 2011). However, the set of necessary and sufficient conditions (i.e., Bellman equations) is readily available for MDPs (Bellman 1957). We state the Bellman equations for this problem, re-express them in a compact form and use it to obtain the maximum allowable error for the single parameter and multiple parameter problem in Sections 3.1.1 and 3.1.2, respectively.

Let $\mathbf{P}^{a_s} = (P^{a_s}(1), P^{a_s}(2), ..., P^{a_s}(|S|))$. Note that $\mathbf{P}^{\pi(s)}$ is the s^{th} row of $\mathbf{P}^{\pi}$. The Bellman equations for the stationary rewards problem are:

$$V_s(\boldsymbol{\rho}) = \max_{a_s \in A(s)} \{r^{a_s} + \boldsymbol{\Delta}^{a_s}\boldsymbol{\rho} + \gamma \mathbf{P}^{a_s}\mathbf{V}(\boldsymbol{\rho})\} \quad \forall s \in S,$$

and $\widetilde{\pi}$ is optimal if and only if:

$$r^{\widetilde{\pi}(s)} + \boldsymbol{\Delta}^{\widetilde{\pi}(s)}\boldsymbol{\rho} + \gamma \mathbf{P}^{\widetilde{\pi}(s)}\mathbf{V}^{\widetilde{\pi}}(\boldsymbol{\rho}) \geq r^{a_s} + \boldsymbol{\Delta}^{a_s}\boldsymbol{\rho} + \gamma \mathbf{P}^{a_s}\mathbf{V}^{\widetilde{\pi}}(\boldsymbol{\rho}) \qquad \forall\, s \in S,\ a_s \in A(s). \tag{3–3}$$

Define:

$$c^{\widetilde{\pi},a_s} = r^{\widetilde{\pi}(s)} - r^{a_s} + \gamma \left(\mathbf{P}^{\widetilde{\pi}(s)} - \mathbf{P}^{a_s}\right)\left(\mathbf{I} - \gamma \mathbf{P}^{\widetilde{\pi}}\right)^{-1}\mathbf{r}^{\widetilde{\pi}},$$

and:

$$\mathbf{b}^{\widetilde{\pi},a_s} = \boldsymbol{\Delta}^{\widetilde{\pi}(s)} - \boldsymbol{\Delta}^{a_s} + \gamma \left(\mathbf{P}^{\widetilde{\pi}(s)} - \mathbf{P}^{a_s}\right)\left(\mathbf{I} - \gamma \mathbf{P}^{\widetilde{\pi}}\right)^{-1}\boldsymbol{\Delta}^{\widetilde{\pi}}.$$

Note that $c^{\widetilde{\pi},a_s}$ is the marginal decrease in the estimated reward that results from a single perturbation of the action at s, while $\mathbf{b}^{\widetilde{\pi},a_s}$ is the marginal change in the estimation error that results from that action perturbation. Using our definitions of $c^{\widetilde{\pi},a_s}$ and $\mathbf{b}^{\widetilde{\pi},a_s}$, the necessary and sufficient optimal conditions expressed in Equation (3–3) can be rewritten as:

$$c^{\widetilde{\pi},a_s} + \mathbf{b}^{\widetilde{\pi},a_s}\boldsymbol{\rho} \geq 0 \qquad \forall\, s \in S,\ a_s \in A(s). \tag{3–4}$$

Let H denote the region where $\widetilde{\pi}$ is optimal:

$$H = \{\boldsymbol{\rho} : V^{\widetilde{\pi}}(\boldsymbol{\rho}) \geq V^{\pi}(\boldsymbol{\rho}),\ \forall \pi \in \Pi\}.$$

Theorem 3.1. *Given Equation (3–2), H is closed and convex.*

Proof. It follows from Equation (3–4) that H is the intersection of closed half-spaces. Hence, H is closed and convex. □

3.1.1 Single Parameter Sensitivity Analysis

In single parameter sensitivity analysis, we are interested in the set of ρ_i values where $\widetilde{\pi}$ remains optimal when $\rho_{j \neq i} = 0$. It follows from Theorem 3.1 that there exist constants $\rho_i^{(l)}, \rho_i^{(u)} \in \mathbb{R}$ such that $\widetilde{\pi}$ remains optimal when $\rho_{j \neq i} = 0$ and $\rho_i \in [\rho_i^{(l)}, \rho_i^{(u)}]$.

Proposition 3.1. *Given Equation (3–2),*

$$\rho_i^{(l)} = \begin{cases} \infty & b_i^{\widetilde{\pi}, a_s} \leq 0 \quad \forall\, s \in S,\ a_s \in A(s) \\ \max_{b_i^{\widetilde{\pi}, a_s} > 0, \forall s, a_s} -\frac{c^{\widetilde{\pi}, a_s}}{b_i^{\widetilde{\pi}, a_s}} & \text{otherwise,} \end{cases}$$

and:

$$\rho_i^{(u)} = \begin{cases} -\infty & b_i^{\widetilde{\pi}, a_s} \geq 0 \quad \forall\, s \in S,\ a_s \in A(s) \\ \min_{b_i^{\widetilde{\pi}, a_s} < 0, \forall s, a_s} -\frac{c^{\widetilde{\pi}, a_s}}{b_i^{\widetilde{\pi}, a_s}} & \text{otherwise.} \end{cases}$$

Proof. Let $b_i^{\widetilde{\pi}, a_s}$ denote the i^{th} entry of $\mathbf{b}^{\widetilde{\pi}, a_s}$. Setting $\rho_{j \neq i} = 0$ and using our definitions of $c^{\widetilde{\pi}, a_s}$ and $b_i^{\widetilde{\pi}, a_s}$, we obtain the following necessary and sufficient optimality conditions from Equation (3–4):

$$\rho_i \geq -\frac{c^{\widetilde{\pi}, a_s}}{b_i^{\widetilde{\pi}, a_s}} \quad \text{when } b_i^{\widetilde{\pi}, a_s} > 0, \tag{3–5}$$

and:

$$\rho_i \leq -\frac{c^{\widetilde{\pi}, a_s}}{b_i^{\widetilde{\pi}, a_s}} \quad \text{when } b_i^{\widetilde{\pi}, a_s} < 0. \tag{3–6}$$

Given that these hold for all values of ρ_i at optimality, the extreme values are of interest and the proposition follows from Equations (3–5) and (3–6). □

Next, we illustrate how single parameter sensitivity analysis can be conducted for a capacitated stochastic lot-sizing problem with uncertain ordering cost and backlog penalty. The interested reader is referred to Muckstadt and Sapra (2010) for an introduction to lot-sizing problems.

Example 1: Capacitated stochastic lot-sizing problem. Consider a lot-sizing problem where the probability distribution of demand in each period is stationary and given by $P(D = 0) = P(D = 1) = P(D = 2) = 1/3$. Each item sells for \$150. The inventory capacity of the system is 3 and backlogging is allowed. There are a total of 5 states, $S = \{-1, 0, 1, 2, 3\}$. The index of each state represents the amount of inventory that is available at the beginning of the period. Orders are placed at the start of the period and an order must be placed if there is no inventory. Hence, the set of feasible actions (i.e., order quantity) for each state is $A(-1) = \{2, 3, 4\}$, $A(0) = \{1, 2, 3\}$, $A(1) = \{0, 1, 2\}$, $A(2) = \{0, 1\}$ and $A(3) = \{0\}$. We assume that the stock arrives at the end of the period in which it is ordered and the demand is also realized at the end of the period. The production cost of each item is \$20 and the holding cost of each item is \$5 per period. The value of the ordering cost and backlog penalty are unclear, but believed to be \$40 and \$100, respectively. We analyze each uncertain parameter independently such that $N = 1$ for each case. The problem extends across an infinite horizon with a per-period discount factor of 0.9.

Let ρ_1 and ρ_2 denote the estimation error in the order cost and backlog penalty, respectively. It follows that the expected reward associated with $\pi(s)$ is:

$$r^{\pi(s)} = \begin{cases} (0.9)(150)(1/3 + 2/3) - 20\pi(s) - 40 - 100 & \text{if } s = -1 \\ (0.9)(150)(1/3 + 2/3) - 5s - 20\pi(s) - 40 I_{\pi(s)} & \text{otherwise.} \end{cases}$$

In addition, we define the following indicator vector:

$$I_{\pi(s)} = \begin{cases} 1 & \text{if } \pi(s) \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\mathbf{\Delta}^{\pi} = \begin{pmatrix} -l_{\pi(-1)} & -100 \\ -l_{\pi(0)} & 0 \\ -l_{\pi(1)} & 0 \\ -l_{\pi(2)} & 0 \\ -l_{\pi(3)} & 0 \end{pmatrix}.$$

The transition probabilities are defined by:

$$P^{a_s}(s') = \begin{cases} 1/3 & \text{if } -2 \le s' - (s + a_s) \le 0 \\ 0 & \text{otherwise.} \end{cases}$$

Solving the MDP described above with the policy iteration approach, we obtain $\widetilde{\pi} = (4, 3, 2, 0, 0)$ and $\mathbf{V}^{\widetilde{\pi}}(\mathbf{0}) = (\$723.5, \$843.5, \$858.5, \$908, \$928.5)'$. It follows from Equation (3–2) that the marginal change in $\mathbf{V}^{\widetilde{\pi}}(\boldsymbol{\rho})$ is $(-196, -196, -196, -168, -156)'\rho_1 + (-100, 0, 0, 0, 0)'\rho_2$.

There are two uncertain parameters in this problem. We analyze the optimal region of $\widetilde{\pi}$ with respect to ρ_1 by setting $\rho_2 = 0$. The corresponding $c^{\widetilde{\pi}, a_s}$ and $b_i^{\widetilde{\pi}, a_s}$ values of each action are listed in Table 3-1. At state -1, the optimal decision is to order 4 units when $\boldsymbol{\rho} = \mathbf{0}$. However, if the ordering cost decreases by more than 46% (with ρ_2 remaining 0), ordering 3 units will result in higher expected profits than ordering 4 units. In addition, the decision to order 4 units at state -1 is better than that of ordering 2 units so long as the ordering cost does not decline by 200% of its estimated value (i.e., $\$40(1 - 2.00) = -\$40$). Assuming that the ordering cost must be nonnegative, it follows that $a_{-1} = 2$ is suboptimal when $\rho_2 = 0$. It follows from Table 3-1 that $\rho^{(l)} = \max\{-2.00, -0.46, -1.23, -\infty\} = -0.46$ and $\rho^{(u)} = \min\{0.04, \infty\} = 0.04$. Hence, $\widetilde{\pi}$ remains optimal for all $\rho_1 \in [-0.46, 0.04]$ given that $\rho_2 = 0$.

In a similar fashion, we obtain that $\widetilde{\pi}$ remains optimal for all $\rho_2 \in [-0.03, \infty]$, given that $\rho_1 = 0$, from the results in Table 3-1. □

3.1.2 Tolerance Approach

When there are multiple uncertain parameters, Wendell (1985) proposed finding a tolerance level τ, where τ is the maximum ratio uncertain parameters are allowed to vary from their base value such that the optimal basis of a linear program remains optimal. Note that τ is, by definition, non-negative. Geometrically speaking, this entails finding the largest hypercube that is contained in the critical region (i.e., H). Wendell showed that the maximum allowable tolerance, which we denote by τ^*, can be obtained by finding the maximum tolerance with respect to each constraint independently. Following a similar approach, we illustrate how the tolerance level can be computed for a MDP with uncertain rewards. Let τ_{s,a_s} denote the maximum tolerance allowable by Equation (3–4) for state s and action a_s:

$$\tau_{s,a_s} = max\{y : c^{\widetilde{\pi},a_s} + \mathbf{b}^{\widetilde{\pi},a_s}\boldsymbol{\rho} \geq 0 \text{ and } |\rho_i| \leq y \text{ for } i = 1, 2, ..., N\}.$$

Proposition 3.2. *Given Equation (3–2),*

$$\tau^* = \min_{s,a_s} \frac{c^{\widetilde{\pi},a_s}}{\sum_{i=1}^{N} |b_i^{\widetilde{\pi},a_s}|}.$$

Proof. To find the maximum allowable tolerance for each constraint expressed by Equation (3–4), we consider the worst case scenario where:

$$\rho_i = \begin{cases} \tau_{s,a_s} & \text{if } b_i^{\widetilde{\pi},a_s} \leq 0 \\ -\tau_{s,a_s} & \text{otherwise.} \end{cases} \tag{3–7}$$

Since $c^{\widetilde{\pi},a_s}$ is, by definition, non-negative, we obtain from Equation (3–4) the following expression for τ_{s,a_s}:

$$\tau_{s,a_s} = \frac{c^{\widetilde{\pi},a_s}}{\sum_{i=1}^{N} |b_i^{\widetilde{\pi},a_s}|}.$$

Since τ^* cannot be larger than any of the individual tolerances τ_{s,a_s}:

$$\tau^* = \min_{s,a_s} \tau_{s,a_s} = \min_{s,a_s} \frac{c^{\widetilde{\pi},a_s}}{\sum_{i=1}^{N} |b_i^{\widetilde{\pi},a_s}|}.$$

□

We reconsider Example 1, allowing for simultaneous perturbations in the ordering cost and backlog penalty, as follows.

Example 2: Tolerance approach. Consider Example 1 again. When the ordering cost and backlog penalty are allowed to perturb simultaneously, it follows from Table 3-1 and Proposition 3.2 that $\tau^* = 0.02$ and is associated with the action $a_1 = 0$. This implies that $\tilde{\pi}$ will remain optimal so long as the ordering cost and backlog penalty do not deviate from their current estimates by more than 2%. In particular, it is suboptimal to order when $s = 1$ if we underestimate the ordering cost and overestimate the penalty cost by more than 2% each. □

3.2 Non-Stationary Rewards

In this section, we consider the non-stationary rewards problem where uncertain parameters are allowed to vary at each period. Let υ denote the tolerance for the non-stationary rewards problem. In addition, let $\boldsymbol{\omega}$ represent the estimation error in the non-stationary reward problem, where $\omega_{s,i,t}$ denotes the value of ρ_i at state s at period t. We say that $\boldsymbol{\omega}$ is stationary if $\omega_{s,i,t_1} = \omega_{s,i,t_2}$ for all t_1, t_2, s and i. If $\boldsymbol{\omega}$ is stationary, ρ_i depends only on the state that the process is in and we denote the value of ρ_i at state s by $\omega_{s,i}$. Let Ω_υ^{NS} and Ω_υ^{ST} denote the sets containing all non-stationary $\boldsymbol{\omega}$ and stationary $\boldsymbol{\omega}$ for a given υ, respectively. Let $P^\pi_{s,t}(i)$ denote the probability of being in state i at time t under policy π given initial state s. The value function of state s given π and $\boldsymbol{\omega}$ is:

$$V_s^\pi(\boldsymbol{\omega}) = \sum_{t=0}^{\infty} \sum_{i \in S} \gamma^t P^\pi_{s,t}(i) \left[r^{\pi(i)} + \sum_{j=1}^{N} \Delta_j^{\pi(i)} \omega_{i,j,t} \right]. \tag{3–8}$$

We say that $\tilde{\pi}$ is υ-optimal if:

$$V_s^{\tilde{\pi}}(\boldsymbol{\omega}) \geq V_s^\pi(\boldsymbol{\omega}) \qquad \forall s \in S, \pi \in \Pi, \boldsymbol{\omega} \in \Omega_\upsilon^{NS}. \tag{3–9}$$

The condition must hold for all possible sets of scenarios across the infinite horizon.

Since, there are infinitely many elements in Ω_υ^{NS}, it is impossible to evaluate the υ-optimality of a policy with Conditions (3–9). Theorem 3.2 highlights that we can limit our analysis to stationary $\boldsymbol{\omega}$.

Theorem 3.2. *Given Equation (3–8), $\widetilde{\pi}$ is υ-optimal if and only if:*

$$V_s^{\widetilde{\pi}}(\boldsymbol{\omega}) \geq V_s^{\pi}(\boldsymbol{\omega}) \quad \forall s \in S, \pi \in \Pi, \boldsymbol{\omega} \in \Omega_\upsilon^{ST}. \tag{3–10}$$

Proof. First, we prove that $\widetilde{\pi}$ is υ-optimal if Conditions (3–10) hold. We prove this by contradiction. Assume that Condition (3–10) holds and $\widetilde{\pi}$ is not υ-optimal. This implies that there must exist some $\boldsymbol{\omega}' \in \Omega_\upsilon^{NS} \setminus \Omega_\upsilon^{ST}$, $s \in S$ and $\pi' \in \Pi$ such that $V_s^{\widetilde{\pi}}(\boldsymbol{\omega}') - V_s^{\pi'}(\boldsymbol{\omega}') < 0$. It follows from Equation (3–8) that

$$V_s^{\widetilde{\pi}}(\boldsymbol{\omega}') - V_s^{\pi'}(\boldsymbol{\omega}') = \\ \sum_{t=0}^{\infty}\sum_{i\in S}\gamma^t\left[\left(P_{s,t}^{\widetilde{\pi}}(i)r^{\widetilde{\pi}(i)} - P_{s,t}^{\pi'}(i)r^{\pi'(i)}\right) + \sum_{j=1}^{N}\left(P_{s,t}^{\widetilde{\pi}}(i)\Delta_j^{\widetilde{\pi}(i)} - P_{s,t}^{\pi'}(i)\Delta_j^{\pi'(i)}\right)\omega_{i,j,t}\right].$$

We construct a stationary $\boldsymbol{\omega}''$ by setting:

$$\omega''_{i,j} = \begin{cases} \upsilon & \text{if } P_{s,t}^{\widetilde{\pi}}(i)\Delta_j^{\widetilde{\pi}(i)} - P_{s,t}^{\pi'}(i)\Delta_j^{\pi'(i)} < 0 \\ -\upsilon & \text{otherwise.} \end{cases}$$

Note that $\boldsymbol{\omega}'' \in \Omega_\upsilon^{ST}$. In addition, $V_s^{\widetilde{\pi}}(\boldsymbol{\omega}'') - V_s^{\pi'}(\boldsymbol{\omega}'') \leq V_s^{\widetilde{\pi}}(\boldsymbol{\omega}') - V_s^{\pi'}(\boldsymbol{\omega}') < 0$, contradicting Conditions (3–10). Therefore, Conditions (3–10) imply that $\widetilde{\pi}$ is υ-optimal. The proof for the reverse direction is straightforward and follows from the observation that $\Omega_\upsilon^{ST} \subseteq \Omega_\upsilon^{NS}$. □

Theorem 3.2 provides a set of conditions that can be used to evaluate the υ-optimality of a policy. However, the number of policies in Π can grow rapidly with the size of the problem. Corollary 1 provides a more compact set of conditions. First, we make the following definitions:

$$\mathbf{V}^{\pi}(\boldsymbol{\omega}) = \left(V_1^{\pi}(\boldsymbol{\omega}), V_2^{\pi}(\boldsymbol{\omega}), \cdots, V_{|S|}^{\pi}(\boldsymbol{\omega})\right)' \text{ and } \boldsymbol{\omega}_s = (\omega_{s,1}, \omega_{s,2}, \cdots, \omega_{s,N})'.$$

Corollary 1. *Given Equation (3–8), $\widetilde{\pi}$ is υ-optimal if and only if the following Bellman equations are satisfied:*

$$r^{\widetilde{\pi}(s)} - r^{a_s} + \left(\boldsymbol{\Delta}^{\widetilde{\pi}(s)} - \boldsymbol{\Delta}^{a_s}\right)\boldsymbol{\omega}_s + \gamma\left(\mathbf{P}^{\widetilde{\pi}(s)} - \mathbf{P}^{a_s}\right)\mathbf{V}^{\widetilde{\pi}}(\boldsymbol{\omega}) \geq 0 \qquad \forall\, s \in S,\ a_s \in A(s), \boldsymbol{\omega} \in \Omega_{\upsilon}^{ST}.$$

Proof. For a given $\boldsymbol{\omega}$, the Bellman equations are necessary and sufficient:

$$\begin{aligned} &V_s^{\widetilde{\pi}}(\boldsymbol{\omega}) \geq V_s^{\pi}(\boldsymbol{\omega}) \ \ \forall s \in S, \pi \in \Pi \qquad \Leftrightarrow \\ &\quad r^{\widetilde{\pi}(s)} - r^{a_s} + \left(\boldsymbol{\Delta}^{\widetilde{\pi}(s)} - \boldsymbol{\Delta}^{a_s}\right)\boldsymbol{\omega}_s + \gamma\left(\mathbf{P}^{\widetilde{\pi}(s)} - \mathbf{P}^{a_s}\right)\mathbf{V}^{\widetilde{\pi}}(\boldsymbol{\omega}) \geq 0 \ \ \forall s \in S,\ a_s \in A(s). \end{aligned}$$

This implies that:

$$\begin{aligned} &V_s^{\widetilde{\pi}}(\boldsymbol{\omega}) \geq V_s^{\pi}(\boldsymbol{\omega}) \ \ \forall s \in S, \pi \in \Pi, \boldsymbol{\omega} \in \Omega_{\upsilon}^{ST} \qquad \Leftrightarrow \\ &\quad r^{\widetilde{\pi}(s)} - r^{a_s} + \left(\boldsymbol{\Delta}^{\widetilde{\pi}(s)} - \boldsymbol{\Delta}^{a_s}\right)\boldsymbol{\omega}_s + \gamma\left(\mathbf{P}^{\widetilde{\pi}(s)} - \mathbf{P}^{a_s}\right)\mathbf{V}^{\widetilde{\pi}}(\boldsymbol{\omega}) \geq 0 \ \ \forall s \in S,\ a_s \in A(s), \boldsymbol{\omega} \in \Omega_{\upsilon}^{ST} \end{aligned}$$

and the corollary follows from Theorem 3.2. □

Next, we illustrate how υ^*, the maximum allowable tolerance for the non-stationary problem, can be obtained from Corollary 1. Let $(\mathbf{I} - \gamma\mathbf{P}^{\pi})^{-1}_{s,i}$ denote the entry in the s^{th} row and i^{th} column of the matrix $(\mathbf{I} - \gamma\mathbf{P}^{\pi})^{-1}$. For stationary $\boldsymbol{\omega}$, the value function of a state can be expressed as:

$$V_s^{\pi}(\boldsymbol{\omega}) = \sum_{i \in S} (\mathbf{I} - \gamma\mathbf{P}^{\pi})^{-1}_{s,i}\left(r^{\pi(i)} + \sum_{j=1}^{N} \Delta_j^{\pi(i)}\omega_{i,j}\right). \tag{3–11}$$

Define:

$$f_{i,j}^{\widetilde{\pi},a_s} = \begin{cases} (1 + G_s^{\widetilde{\pi},a_s})\Delta_j^{\widetilde{\pi}(s)} - \Delta_j^{a_s} & \text{if } i = s \\ G_i^{\widetilde{\pi},a_s}\Delta_j^{\widetilde{\pi}(i)} & \text{otherwise,} \end{cases}$$

where $G_i^{\widetilde{\pi},a_s}$ denote the i^{th} entry of $\gamma\left(\mathbf{P}^{\widetilde{\pi}(s)} - \mathbf{P}^{a_s}\right)\left(\mathbf{I} - \gamma\mathbf{P}^{\widetilde{\pi}}\right)^{-1}$. Substituting Equation (3–11) into Corollary 1 and using our definitions of $c^{\widetilde{\pi},a_s}$ and $f_{i,j}^{\widetilde{\pi},a_s}$:

$$c^{\widetilde{\pi},a_s} + \sum_{i \in S}\sum_{j=1}^{N} f_{i,j}^{\widetilde{\pi},a_s}\omega_{i,j} \geq 0 \qquad \forall\, s \in S,\ a_s \in A(s), \boldsymbol{\omega} \in \Omega_{\upsilon}^{NS}. \tag{3–12}$$

Proposition 3.3. *Given Equation (3–8),*

$$\upsilon^* = \min_{s,a_s} \frac{c^{\widetilde{\pi},a_s}}{\sum_{j=1}^{N} d_j^{\widetilde{\pi},a_s}}.$$

Proof. Let υ_{s,a_s} denote the maximum tolerance allowable by Equation (3–12) for state s and action a_s:

$$\upsilon_{s,a_s} = max\{y : c^{\widetilde{\pi},a_s} + \sum_{i \in S} \sum_{j=1}^{N} f_{i,j}^{\widetilde{\pi},a_s} \omega_{i,j} \geq 0 \text{ and } |\omega_{i,j}| \leq y \text{ for } i \in S, j = 1, 2, ..., N\}.$$

Similar to the stationary reward case, we consider the worst case scenario and obtain the following expression for υ_{s,a_s} from Equations (3–12):

$$\upsilon_{s,a_s} = \frac{c^{\widetilde{\pi},a_s}}{\sum_{j=1}^{N} d_j^{\widetilde{\pi},a_s}},$$

where $d_j^{\widetilde{\pi},a_s} = \sum_{i \in S} |f_{i,j}^{\widetilde{\pi},a_s}|$. Hence, the maximum allowable tolerance for the non-stationary rewards problem is:

$$\upsilon^* = \min_{s,a_s} \upsilon_{s,a_s} = \min_{s,a_s} \frac{c^{\widetilde{\pi},a_s}}{\sum_{j=1}^{N} d_j^{\widetilde{\pi},a_s}}.$$

□

We re-examine our capacitated stochastic lot-sizing problem in Example 3 by allowing the ordering cost and backlog penalty to vary over time.

Example 3: Non-stationary rewards. Consider Example 1 again. When the ordering cost and backlog penalty are allowed to perturb simultaneously at each period, it follows from Table 3-2 and Proposition 3.3 that $\upsilon^* = 0.01$ and is associated with the action $a_1 = 0$. This implies that $\widetilde{\pi}$ remains optimal so long as the ordering cost and backlog penalty do not deviate from their current estimates by more than 1% across all periods. □

3.3 Tolerance Gap

At the start of this chapter, we claimed that the maximum allowable tolerance obtained under the assumption that the parameters are stationary may be overly optimistic if the values of the uncertain parameters vary across the horizon. In this section, we provide a formal proof for this statement and highlight the conditions where the tolerances of the stationary and non-stationary rewards problem are the same and the conditions where they differ.

If the current decision $\widetilde{\pi}(s)$ is replaced by a_s, $|b_j^{\widetilde{\pi},a_s}|$ and $d_j^{\widetilde{\pi},a_s}$ represent the marginal changes in the estimation error of the stationary and non-stationary rewards problem, respectively. Recall that τ_{s,a_s} and υ_{s,a_s} denote the maximum allowable tolerance for state s and action a_s in the stationary and non-stationary problem, respectively. τ^* and υ^* denote the maximum allowable tolerance for the stationary and non-stationary problem, respectively. It follows from Propositions 3.2 and 3.3 that $\tau^* = \upsilon^*$ if $|b_j^{\widetilde{\pi},a_s}| = d_j^{\widetilde{\pi},a_s}$ for all a_s and j. If that is not true, the allowable tolerances of the stationary and non-stationary rewards problem may differ. In particular, $\tau^* > \upsilon^*$ if $|b_j^{\widetilde{\pi},a_s}| < d_j^{\widetilde{\pi},a_s}$ for some s and a_s where $\tau_{s,a_s} = \tau^*$.

Lemma 1. *$|b_j^{\widetilde{\pi},a_s}| \leq d_j^{\widetilde{\pi},a_s}$.*

Proof. Note that $b_j^{\widetilde{\pi},a_s}$ and $d_j^{\widetilde{\pi},a_s}$ are the sums of the values and absolute values of $f_{i,j}^{\widetilde{\pi},a_s}$ across all states, respectively. Hence, it follows that:

$$|b_j^{\widetilde{\pi},a_s}| = |\sum_{i \in S} f_{i,j}^{\widetilde{\pi},a_s}| \leq \sum_{i \in S} |f_{i,j}^{\widetilde{\pi},a_s}| = d_j^{\widetilde{\pi},a_s}.$$

□

Theorem 3.3. *$\tau^* \geq \upsilon^*$.*

Proof. Follows directly from Lemma 1 and Propositions 3.2 and 3.3. □

Lemma 1 highlights that $|b_j^{\widetilde{\pi},a_s}|$ cannot be greater than $d_j^{\widetilde{\pi},a_s}$. Theorem 3.3 states that the maximum allowable tolerance obtained for a stationary rewards problem is at least as great as (i.e., more optimistic) the maximum allowable tolerance obtained when the stationary reward parameters assumption is relaxed.

Lemma 2. $|b_j^{\widetilde{\pi},a_s}| < d_j^{\widetilde{\pi},a_s}$ *if and only if there exist* $s_1, s_2 \in S$ *where* $f_{s_1,j}^{\widetilde{\pi},a_s}$ *is positive and* $f_{s_2,j}^{\widetilde{\pi},a_s}$ *is negative.*

Proof. If there exist $s_1, s_2 \in S$ where $f_{s_1,j}^{\widetilde{\pi},a_s}$ is positive and $f_{s_2,j}^{\widetilde{\pi},a_s}$ is negative,

$$|b_j^{\widetilde{\pi},a_s}| = |\sum_{i\in S} f_{i,j}^{\widetilde{\pi},a_s}| < \sum_{i\in S} |f_{i,j}^{\widetilde{\pi},a_s}| = d_j^{\widetilde{\pi},a_s}.$$

If $|b_j^{\widetilde{\pi},a_s}| < d_j^{\widetilde{\pi},a_s}$, $|\sum_{i\in S} f_{i,j}^{\widetilde{\pi},a_s}| < \sum_{i\in S} |f_{i,j}^{\widetilde{\pi},a_s}|$ and there must exist $s_1, s_2 \in S$ where $f_{s_1,j}^{\widetilde{\pi},a_s}$ is positive and $f_{s_2,j}^{\widetilde{\pi},a_s}$ is negative. □

Theorem 3.4. $\tau_{s,a_s} > \upsilon_{s,a_s}$ *if and only if there exists some parameter* j *and states* $s_1, s_2 \in S$ *where* $f_{s_1,j}^{\widetilde{\pi},a_s}$ *is positive and* $f_{s_2,j}^{\widetilde{\pi},a_s}$ *is negative.*

Proof. If there exists some parameter j and states $s_1, s_2 \in S$ where $f_{s_1,j}^{\widetilde{\pi},a_s}$ is positive and $f_{s_2,j}^{\widetilde{\pi},a_s}$ is negative, it follows from Lemma 2 that $|b_j^{\widetilde{\pi},a_s}| < d_j^{\widetilde{\pi},a_s}$. Since $|b_j^{\widetilde{\pi},a_s}| \leq d_j^{\widetilde{\pi},a_s}$ for all j (Lemma 1), $\tau_{s,a_s} > \upsilon_{s,a_s}$.

If $\tau_{s,a_s} > \upsilon_{s,a_s}$, it implies that there exists some parameter j where $|b_j^{\widetilde{\pi},a_s}| < d_j^{\widetilde{\pi},a_s}$ and it follows from Lemma 2 that there exist states $s_1, s_2 \in S$ where $f_{s_1,j}^{\widetilde{\pi},a_s}$ is positive and $f_{s_2,j}^{\widetilde{\pi},a_s}$ is negative. □

Theorem 3.4 provides a set of necessary and sufficient conditions for there to be a difference in the tolerances between the stationary and non-stationary rewards problem. In particular, the allowable tolerance associated with action a_s differs if a_s increases and decreases the effect of parameter j on the rewards associated with two different states. In Example 4, we illustrate how this condition can be checked without performing the actual computations. Next, we present a set of conditions where the marginal changes in the estimation error of the stationary and non-stationary rewards problem are equal.

Corollary 2. $|b_j^{\widetilde{\pi},a_s}| = d_j^{\widetilde{\pi},a_s}$ *if there exists* $k \in S$ *where* $\Delta_j^{\pi(i)} = 0$ *for all* $i \neq k$ *and* $\pi \in \Pi$.

Proof. If $\Delta_j^{\pi(i)} = 0$ for all $i \neq k$ and $\pi \in \Pi$, there can be at most one non-zero $f_{i,j}^{\widetilde{\pi},a_s}$ term and it follows from Lemma 1 and Lemma 2 that $|b_j^{\widetilde{\pi},a_s}| = d_j^{\widetilde{\pi},a_s}$. □

Corollary 2 provides a sufficient condition for $|b_j^{\widetilde{\pi},a_s}| = d_j^{\widetilde{\pi},a_s}$. In particular, when the contribution of ρ_i to the value function is restricted to just a single state, the impact of ρ_i on the allowable tolerance is the same, regardless of whether ρ_i is stationary or not. The validity of Theorems 3.3 and 3.4 and Corollary 2 are illustrated in the following example.

Example 4: Tolerance gap. In Examples 2 and 3, we obtain $\tau^* = 0.02$ and $\upsilon^* = 0.01$, respectively. This result is consistent with Theorem 3.3 which states that $\tau^* \geq \upsilon^*$.

Example 2 highlights that $\tau_{1,0} = \tau^*$. τ^* will be strictly greater than υ^* if there is a difference in the allowable tolerance associated with not ordering when inventory is one. Under $\widetilde{\pi}$, the optimal action is to order two units. If no order is placed, an ordering cost is avoided and the probability of entering state 1 remains the same. Hence, $f_{1,1}^{\widetilde{\pi},0}$ is negative. However, the probability of entering state -1 is increased (i.e., $G_{-1}^{\widetilde{\pi},0}$ is negative). Since $\Delta_1^{\widetilde{\pi}(-1)}$ is also negative, $f_{-1,1}^{\widetilde{\pi},0}$ is positive. Hence, it follows from Theorem 3.4 that $\tau^* > \upsilon^*$.

Since $\Delta_2^{\pi(s)} = 0$ for all $s \geq 0$, it follows from Corollary 2 that $|b_2^{\widetilde{\pi},a_s}| = d_2^{\widetilde{\pi},a_s}$ for all a_s. This is validated by comparing the values of $|b_2^{\widetilde{\pi},a_s}|$ and $d_2^{\widetilde{\pi},a_s}$ in Tables 3-1 and 3-2, respectively. This implies that the impact of ρ_2 on the allowable tolerances for the stationary and non-stationary rewards problems are the same and that any reduction in the allowable tolerance of the non-stationary problem is due to the relaxation of the stationary assumption on ρ_1. $\square$

In Example 1, we found that the optimal policy is to bring the inventory level up to three whenever the inventory drops below two. We refer to this as an order-up-to policy. Next, we will show that $\tau^* > \upsilon^*$ if the action associated with τ^* changes the reorder point for general lot-sizing problems under mild assumptions.

Consider a lot-sizing problem where p_i denotes the probability that demand is i. We assume that p_i is stationary (i.e., remains the same across the horizon). There is constant lead time, a discount factor $0 < \gamma < 1$, linear production cost and a convex holding cost. In addition, each order incurs an uncertain ordering cost and each backlog item incurs an uncertain penalty cost. As in Example 1, we model the uncertainties in the ordering cost and backlog penalty by ρ_1 and ρ_2, respectively. The objective is to find the policy that minimizes the long run expected costs.

This problem can be formulated as a MDP where the states represent the amount of inventory available and the actions represent the amount of inventory to order. It is well-known that an optimal order-up-to policy exists for this problem for a given ρ_1 and ρ_2 Veinott and Wagner (1965).

Theorem 3.5. *For the lot-sizing problem described above, $\tau^* > \upsilon^*$ if there exists $\tau_{s,a_s} = \tau^*$ where a_s changes the reorder point and $p_i < \frac{1}{2\gamma}$ for all i.*

Proof. First, we consider the case where $\widetilde{\pi}(s) > 0$ and $a_s = 0$. If no order is placed at s, there must exist some state $s' < s$ where $G_{s'}^{\widetilde{\pi},a_s}$ is negative (i.e., the probability of entering s' is increased as a result of a_s). Since $\widetilde{\pi}$ is an order-up-to policy, $\Delta_1^{a_{s'}}$ is negative and $f_{s',1}^{\widetilde{\pi},0}$ is positive. Recall that $G_s^{\widetilde{\pi},a_s}$ is the discounted expected number of subsequent visits to state s under a_s and $\widetilde{\pi}$. Since, $p_i < \frac{1}{2\gamma}$, $G_s^{\widetilde{\pi},a_s} < 1$. Since $G_s^{\widetilde{\pi},a_s} < 1$, $\Delta_j^{\widetilde{\pi}(s)} < 0$ and $\Delta_j^{a_s} = 0$, $f_{s,1}^{\widetilde{\pi},0}$ is negative. Hence, it follows from Theorem 3.4 that $\tau^* > \upsilon^*$.

When $\widetilde{\pi}(s) = 0$ and $a_s > 0$, there must exist some state $s' < s$ where $G_{s'}^{\widetilde{\pi},a_s}$ is positive and $f_{s',1}^{\widetilde{\pi},0}$ is negative. Since $\Delta_j^{\widetilde{\pi}(s)} = 0$ and $\Delta_j^{a_s} < 0$, $f_{s,1}^{\widetilde{\pi},0}$ is positive. Hence, it follows from Theorem 3.4 that $\tau^* > \upsilon^*$. □

Theorem 3.5 highlights that a tolerance gap exists for lot-sizing problems where the action associated with τ^* changes the reorder point when p_i are bounded from above by $\frac{1}{2\gamma}$. Since $\gamma < 1$, the upper bound is greater than 0.5 for all problems and is, in our opinion, a reasonable assumption for most practical problems.

Table 3-1. Single parameter sensitivity analysis and τ.

s	a_s	$c^{\widetilde{\pi},a_s}$	$b_1^{\widetilde{\pi},a_s}$	$b_2^{\widetilde{\pi},a_s}$	$-c^{\widetilde{\pi},a_s}/b_1^{\widetilde{\pi},a_s}$	$-c^{\widetilde{\pi},a_s}/b_2^{\widetilde{\pi},a_s}$	τ_{s,a_s}
-1	2	40.85	20.40	30.00	-2.00	-1.36	0.81
-1	3	5.50	12.00	0.00	-0.46	-	0.46
-1	4	0.00	0.00	0.00	-	-	-
0	1	40.85	20.40	30.00	-2.00	-1.36	0.81
0	2	5.50	12.00	0.00	-0.46	-	0.46
0	3	0.00	0.00	0.00	-	-	-
1	0	0.85	-19.60	30.00	0.04	-0.03	0.02
1	1	5.50	12.00	0.00	-0.46	-	0.46
1	2	0.00	0.00	0.00	-	-	-
2	0	0.00	0.00	0.00	-	-	-
2	1	34.50	28.00	0.00	-1.23	-	1.23
3	0	0.00	0.00	0.00	-	-	-

Table 3-2. Computing υ.

s	a_s	$c^{\widetilde{\pi},a_s}$	$d_1^{\widetilde{\pi},a_s}$	$d_2^{\widetilde{\pi},a_s}$	υ_{s,a_s}
-1	2	40.85	20.40	30.00	0.81
-1	3	5.50	12.00	0.00	0.46
-1	4	0.00	0.00	0.00	-
0	1	40.85	20.40	30.00	0.81
0	2	5.50	12.00	0.00	0.46
0	3	0.00	0.00	0.00	-
1	0	0.85	60.40	30.00	0.01
1	1	5.50	12.00	0.00	0.46
1	2	0.00	0.00	0.00	-
2	0	0.00	0.00	0.00	-
2	1	34.50	52.00	0.00	0.66
3	0	0.00	0.00	0.00	-

CHAPTER 4
REGRET

It has long been recognized that a decision based solely on the reward received is not adequate for describing how individuals make decisions under uncertainty in practice. In 1738, Daniel Bernoulli highlighted the need to consider the perceived utility of the outcome. Two centuries later, Von Neumann and Morgenstern (1944) proposed a set of axioms that guarantee the existence of a utility function where the expected utility of a preferred choice is greater than the utility of a less preferred alternative. However, researchers have identified instances where predictions from expected utility theory deviate from observations in practice and have proposed alternative theories to explain these observations (Leland 2010).

Research in psychology and behavioral decision theory suggests that regret plays an important role in shaping preferences in decisions under uncertainty (see Zeelenberg (1999), Connolly and Zeelenberg (2002) and Engelbrecht-Wiggans and Katok (2009)). Bell (1982) and Loomes and Sugden (1982) highlighted that gaps between the predictions of expected utility theory and observations in practice can be addressed when feelings of regret and rejoicing are taken into account. They proposed using a modified utility function which depends on both reward and regret to model the satisfaction associated with a decision. Their model, commonly referred to as regret theory, was initially developed for pairwise decisions. Subsequently, it was generalized for multiple feasible alternatives in Loomes and Sugden (1987) and an axiomatic foundation for the theory was presented in Sugden (1993). Quiggin (1990) derived the necessary and sufficient conditions for a pairwise choice to be preferred over another for all regret-theoretic decision makers and provided a set of sufficient conditions for problems involving multiple choices in Quiggin (1994). Recently, Bleichrodt et al. (2010) illustrated how regret theory can be measured and they found that individuals are averse to regret and disproportionately averse to large regret.

Let S and C denote the set of possible scenarios and choices, respectively. The outcome of choice c under scenario s is denoted by $x_c^{(s)}$. Let $\mathbf{x}_c^{(s)}$ represent a vector of outcomes associated with choice c under scenario s:

$$\mathbf{x}_c^{(s)} = \left(x_c^{(s)}, -x_1^{(s)}, -x_2^{(s)}, \ldots, x_{c-1}^{(s)}, -x_{c+1}^{(s)}, \ldots, -x_{|C|}^{(s)}\right).$$

Let $u\left(\mathbf{x}_c^{(s)}\right)$ denote the utility of obtaining $x_c^{(s)}$ and missing out on the other opportunities (i.e., incurring $-x_1^{(s)}, -x_2^{(s)}$, etc.). Since the attractiveness of a choice is generally non-decreasing in the reward received and non-increasing in the opportunities missed, it is reasonable to assume that u is non-decreasing in its arguments. Unlike Von Neumann and Morgenstern's (1944) utility function, the regret-theoretic utility function u is dependent on the outcome of possible choices, rather than the selected choice alone. Let p_s denote the probability of scenario s occurring. In regret theory, choice preferences are described as follows:

$$c_1 \succeq c_2 \Leftrightarrow \sum_{s} p_s \left(u\left(\mathbf{x}_{c_1}^{(s)}\right) - u\left(\mathbf{x}_{c_2}^{(s)}\right)\right) \geq 0,$$

where $c_1 \succeq c_2$ denotes that c_1 is at least as preferred as c_2.

Expected utility theory predicts that $c_1 \succeq c_2$ if the outcomes of c_1 stochastically dominates the outcomes of c_2 (Levy 2006). However, violations of stochastic dominance (i.e., $c_2 \succ c_1$ when the outcomes of c_1 stochastically dominates the outcomes of c_2) have been observed in practice. In particular, Birnbaum and Navarrete (1998) presents a set of problems where such behavior is systematically observed. Regret theory predicts stochastic dominance violations (Loomes and Sugden, 1987). Quiggin (1990) showed that for a pairwise problem (i.e., $|C| = 2$) where S is finite and scenarios are equally probable, $c_1 \succeq c_2$ if and only if there exists ρ, a bijection of S onto itself, such that (i) $\rho^{-1} = \rho$ and (ii) $x_1^{(s)} \geq x_2^{(\rho(s))}$ for all s. Note that Condition (ii) implies that outcomes of choice 1 stochastically dominates the outcomes of choice 2. In addition, he also highlighted that Conditions (i) and (ii) are sufficient for problems involving multiple choices (i.e., $|C| > 2$) under additional conditions on the outcomes of the optimal choice under the bijection ρ (Quiggin 1994).

Using Quiggin's results, an analyst is able to construct a regret-theoretic efficient set by eliminating choices that are clearly inferior (i.e., suboptimal for all regret-theoretic utility functions). However, there are several limitations. First, the number of scenarios are assumed to be finite in Quiggin (1990, 1994). However, there are many problems where $|S|$ is infinite (for example, when outcomes are continuously distributed). Second, the conditions are only necessary and sufficient for pairwise problems. When $|C| > 2$, the conditions are sufficient, but not necessary (i.e., the efficient set may not be minimal in size). Third, the number of possible bijections grows exponentially with the number of scenarios and determining if there exists a ρ that satisfies conditions (i) and (ii) is generally challenging. In the next section, we present a set of conditions that addresses these limitations.

4.1 Multivariate Stochastic Dominance

In the previous section, we presented regret theory as a model where the outcome of each choice under each scenario is explicitly modeled. Here, we present an implicit form of the model where S is not explicitly defined. Rather the outcome of each choice is modeled as a random variable that can be dependent on the outcome of other choices.

Let X_c denote the reward of choice c where X_c is a random variable described by some known probability distribution function f_{X_c}. Let $\mathbf{X}_c$ denote a multivariate random variable defined as:

$$\mathbf{X}_c = (X_c, -X_1, -X_2, \dots, -X_{c-1}, -X_{c+1}, \dots, -X_{|C|}).$$

Regret-theoretic choice preferences for the implicit form are stated as follows:

$$c_1 \succeq c_2 \Leftrightarrow \mathbb{E}[u(\mathbf{X}_{c_1}) - u(\mathbf{X}_{c_2})] \geq 0.$$

L is a lower set if $(x_1, x_2, \dots, x_{|C|}) \in L$ implies $(y_1, y_2, \dots, y_{|C|}) \in L$ when $y_i \leq x_i$ for all i.

Consider two multivariate random variables, $\mathbf{X}_1$ and $\mathbf{X}_2$. We say that $\mathbf{X}_1$ stochastically dominates $\mathbf{X}_2$ if and only if:

$$P(\mathbf{X}_1 \in L) \leq P(\mathbf{X}_2 \in L) \text{ for all lower sets } L \subseteq \mathbb{R}^{|C|}.$$

We denote this relationship by $\mathbf{X}_1 \circledS \mathbf{X}_2$. It is well known that $\mathbf{X}_1 \circledS \mathbf{X}_2$ if and only if $\mathbb{E}[u(\mathbf{X}_1)] \geq \mathbb{E}[u(\mathbf{X}_2)]$ for all non-decreasing u (see, for example Shaked and Shanthikumar (2007)). Therefore, it follows from our definition of $\mathbf{X}_c$ that multivariate stochastic dominance is necessary and sufficient for choice preference over all non-increasing regret-theoretic utility functions. Stated formally:

$$\mathbf{X}_{c_1} \circledS \mathbf{X}_{c_2} \Leftrightarrow \mathbb{E}[u(\mathbf{X}_{c_1}) - u(\mathbf{X}_{c_2})] \geq 0 \text{ for all } u \in U, \tag{4–1}$$

where U denotes the set of non-decreasing regret-theoretic utility functions. Unlike the conditions in Quiggin (1994), Condition (4–1) is both necessary and sufficient. In addition, it is easier to check and applies to problems where $|S|$ is infinite.

Next, we present conditions where choice preferences in regret theory are consistent with that of expected utility theory (i.e., $c_1 \succeq c_2$ if $X_{c_1} \circledS X_{c_2}$). We begin with the following theorem. Let $\mathbf{X}_i(k)$ denote the k-th element of $\mathbf{X}_i$.

Theorem 4.1. $X_{c_1} \circledS X_{c_2} \Leftrightarrow \mathbf{X}_{c_1}(k) \circledS \mathbf{X}_{c_2}(k)$ for all k.

Proof. First, we show that:

$$X_{c_1} \circledS X_{c_2} \Rightarrow \mathbf{X}_{c_1}(k) \circledS \mathbf{X}_{c_2}(k) \text{ for all } k. \tag{4–2}$$

We prove this result by showing that (4–2) holds for each k. Without loss of generality, let choices $c_1 = 1$ and $c_2 = 2$. Since $\mathbf{X}_{c_1}(1) = X_{c_1}$ and $\mathbf{X}_{c_2}(1) = X_{c_2}$, (4–2) holds for $k = 1$. Since $\mathbf{X}_{c_1}(2) = -X_2$ and $\mathbf{X}_{c_2}(2) = -X_1$ and $-X_2 \circledS -X_1$, (4–2) holds for $k = 2$. Since $\mathbf{X}_{c_1}(k) = \mathbf{X}_{c_2}(k) = -X_k$, (4–2) holds for $k > 2$ as well. The proof in the reverse direction is straightforward. $\square$

Theorem 4.1 highlights that the outcomes of c_1 stochastically dominates the outcomes of c_2 if and only if each marginal distribution of $\mathbf{X}_{c_1}$ stochastically dominates that of $\mathbf{X}_{c_2}$.

Corollary 3. *When X_i are independent, $\mathbf{X}_{c_1} \circledS \mathbf{X}_{c_2}$ if and only if $X_{c_1} \circledS X_{c_2}$.*

Proof. When X_i are independent, stochastic dominance of each marginal distribution is necessary and sufficient. Therefore, it follows from Theorem 4.1 that:

$$\mathbf{X}_{c_1} \circledS \mathbf{X}_{c_2} \Leftrightarrow \mathbf{X}_{c_1}(k) \circledS \mathbf{X}_{c_2}(k) \text{ for all } k \Leftrightarrow X_{c_1} \circledS X_{c_2}.$$

□

Corollary 3 highlights that when rewards are independent, choices with outcomes that stochastically dominate are preferred by all regret-theoretic decision makers. Next, we show that choice preferences in regret theory are also consistent with that of expected utility theory when we limit ourselves to additive regret-theoretic utility functions. Let $U_1 \subseteq U$ denote the set of all non-decreasing additive regret-theoretic utility functions.

Corollary 4. $\mathbb{E}[u(\mathbf{X}_{c_1}) - u(\mathbf{X}_{c_2})] \geq 0$ for all $u \in U_1$ *if and only if $X_{c_1} \circledS X_{c_2}$.*

Proof. Stochastic dominance of each marginal distribution is necessary and sufficient for $\mathbb{E}[u(\mathbf{X}_{c_1}) - u(\mathbf{X}_{c_2})] \geq 0$ for all $u \in U_1$ (Levy and Paroush 1974). Therefore, it follows from Theorem 4.1 that:

$$\mathbb{E}[u(\mathbf{X}_{c_1}) - u(\mathbf{X}_{c_2})] \geq 0 \text{ for all } u \in U_1 \Leftrightarrow \mathbf{X}_{c_1}(k) \circledS \mathbf{X}_{c_2}(k) \text{ for all } k \Leftrightarrow X_{c_1} \circledS X_{c_2}.$$

□

Corollary 4 states that choices with outcomes that stochastically dominate are preferred by all regret-theoretic decision makers with additive utility functions, even if rewards are dependent.

4.2 Regret in Normative Decision Analysis

The concept of regret has been adopted by various researchers in normative decision analysis. In 1951, Savage proposed finding a solution with minimal maximum regret. It has been argued that these solutions tend to be less conservative than solutions that maximize minimum reward because the former considers missed opportunities over all possible scenarios rather than the reward in the worst case scenario alone. The min-max regret criteria has been used in a variety of problems, including linear programs with imprecise objective coefficients (Inuiguchi and Sakawa, 1995) and network optimization with interval data (Averbakh and Lebedev, 2004). The interested reader is referred to Kouvelis and Yu (1997) and Aissia et al. (2009) for a survey of min-max regret problems.

Despite a history of more than half a century, the representation of regret in the optimization literature has been restricted to the deviation and, to a lesser extent, the ratio of the optimal reward to the received reward. In practice, regret can depend on a subset of missed opportunities, rather than the outcome of the optimal alternative alone.

Let Y_c denote the regret associated with choice c, where Y_c is some function of the reward received and opportunities missed. Let φ denote the regret function:

$$Y_c = \varphi(-\mathbf{X}_c).$$

Since regret is non-increasing in the reward received and non-decreasing in the opportunities missed, we assume that φ is non-decreasing in $-\mathbf{X}_c$.

In the optimization literature, regret is commonly expressed as the difference between the optimal reward and the reward received:

$$Y_c = \varphi_a(-\mathbf{X}_c) = \max_{i \in S} X_i - X_c.$$

We refer to this as *absolute* regret and φ_a as the absolute regret function. In addition, researchers have also proposed expressing regret as a ratio of the optimal and received reward:

$$Y_c = \varphi_r(-\mathbf{X}_c) = \frac{\max_{i \in S} X_i - X_c}{X_c}.$$

We refer to this as *relative* regret and φ_r as the relative regret function. The interested reader is referred to Kovelis and Yu (1997) and Aissia et al. (2009) for a survey of problems that minimize absolute or relative regret.

Next, we generalize absolute regret by introducing the following two notions of regret. First, we express regret as the difference between the α-percentile reward and the received reward:

$$Y_c = \varphi_p(-\mathbf{X}_c) = X^{\alpha} - X_c,$$

where X^{α} is the reward at the α-percentile. We refer to this as *percentile* regret and φ_p as the percentile regret function. Note that percentile regret reduces to absolute regret when $\alpha = 1$.

Next, we express regret as the difference between the average of the k largest rewards and the received reward:

$$Y_c = \varphi_k(-\mathbf{X}_c) = \frac{1}{k}\sum_{i=1}^{k} X^{(i)} - X_c,$$

where $X^{(i)}$ denotes the i^{th} largest reward in S. We refer to this as k*-average* regret and φ_k as the k-average regret function. Note that k-average regret reduces to absolute regret when $k = 1$.

Let Φ denote the set of non-decreasing φ. Observe that $\varphi_a, \varphi_r, \varphi_p, \varphi_k \in \Phi$. The following theorem, Theorem 4.2, highlights that stochastic dominance relations are reversed under “reward-to-regret-transformation” (i.e., if outcomes of c_1 stochastically dominates outcomes of c_2. then regret of c_2 stochastically dominates regret of c_1) when the outcome of choices are independent. This result is general and holds for all forms of regret that are non-decreasing in $-\mathbf{X}_c$, including the four notions of regret presented in this section.

Theorem 4.2. *Suppose that $\varphi \in \Phi$ and X_i are independent. If $X_{c_1} \circledS X_{c_2}$, then $Y_{c_2} \circledS Y_{c_1}$.*

Proof. It follows from Corollary 3 that when X_i are independent:

$$\begin{aligned}
X_{c_1} \circledS X_{c_2} &\Rightarrow \mathbf{X}_{c_1} \circledS \mathbf{X}_{c_2} \\
&\Rightarrow \mathbb{E}[u(\mathbf{X}_{c_1}) - u(\mathbf{X}_{c_2})] \geq 0, \text{ for all non-decreasing } u \\
&\Rightarrow \mathbb{E}[u(-\mathbf{X}_{c_1}) - u(-\mathbf{X}_{c_2})] \leq 0, \text{ for all non-decreasing } u \\
&\Rightarrow \mathbb{E}[\rho(\varphi(-\mathbf{X}_{c_1})) - \rho(\varphi(-\mathbf{X}_{c_2}))] \leq 0, \text{ for all non-decreasing } \rho \\
&\Rightarrow \mathbb{E}[\rho(Y_{c_1}) - \rho(Y_{c_2})] \leq 0, \text{ for all non-decreasing } \rho \\
&\Rightarrow Y_{c_2} \circledS Y_{c_1}.
\end{aligned}$$

□

Since φ_r involves a max operator and the division of a dependent random variable, proving Theorem 4.2 in the absence of Corollary 3 is non-trivial, even for the special case of relative regret. Next, we illustrate how regret can be viewed as a measure of the risk and be adopted within the mean-risk framework. Let r_c denote the risk associated with choice c. The mean-risk framework seeks a choice c that maximizes the following objective:

$$g(c) = \mathbb{E}[\mathbf{X}_c] - \gamma r_c,$$

where γ is some constant. Conventional risk measures are based on the outcomes associated with a choice and the value of γ is defined as zero, positive and negative for a risk neutral, risk averse and risk seeking decision maker, respectively (Krokhmal et al., 2011).

We define a regret-based risk measure as follows:

$$r_c = \rho(Y_c),$$

where ρ is some non-decreasing function. Since the satisfaction of a decision is non-increasing in regret, the value of γ is non-negative in mean-regret models.

We say that φ is *separable* if there exists some random variable $\widetilde{X}$ such that $Y_c = \varphi(-\mathbf{X}_c) = \widetilde{X} - X_c$. Note that there is no independence restriction between $\widetilde{X}$ and X_c. Let Φ_1 denote the set of non-decreasing separable φ. Observe that $\varphi_a, \varphi_p, \varphi_k \in \Phi_1$. Next, we show that being stochastic dominance in regret is necessary and sufficient for mean-regret choice preference if $\varphi \in \Phi_1$.

Theorem 4.3. *If $\varphi \in \Phi_1$, ρ is non-decreasing and $\gamma \geq 0$,*

$$Y_{c_2} \circledS Y_{c_1} \Leftrightarrow \mathbb{E}[\mathbf{X}_{c_1}] - \gamma\rho(Y_{c_1}) \geq \mathbb{E}[\mathbf{X}_{c_2}] - \gamma\rho(Y_{c_2}).$$

Proof. First, we show that $\mathbb{E}[\mathbf{X}_{c_1}] \geq \mathbb{E}[\mathbf{X}_{c_2}]$ if $\varphi \in \Phi_1$ and $Y_{c_2} \circledS Y_{c_1}$. When $Y_{c_2} \circledS Y_{c_1}$, $\mathbb{E}[Y_{c_1}] \leq \mathbb{E}[Y_{c_2}]$. This implies that:

$$\begin{aligned}
\mathbb{E}[Y_{c_1}] &\leq \mathbb{E}[Y_{c_2}] \\
\mathbb{E}[\widetilde{X} - X_{c_1}] &\leq \mathbb{E}[\widetilde{X} - X_{c_2}] \\
\mathbb{E}[\widetilde{X}] - \mathbb{E}[X_{c_1}] &\leq \mathbb{E}[\widetilde{X}] - \mathbb{E}[X_{c_2}] \\
-\mathbb{E}[X_{c_1}] &\leq -\mathbb{E}[X_{c_2}] \\
\mathbb{E}[X_{c_1}] &\geq \mathbb{E}[X_{c_2}].
\end{aligned}$$

The second inequality follows from the fact that $\varphi \in \Phi_1$. In addition, it follows from standard results in stochastic dominance that $\rho(Y_{c_1}) \leq \rho(Y_{c_2})$ for all non-decreasing ρ. Therefore $Y_{c_2} \circledS Y_{c_1} \Rightarrow \mathbb{E}[\mathbf{X}_{c_1}] - \gamma\rho(Y_{c_1}) \geq \mathbb{E}[\mathbf{X}_{c_2}] - \gamma\rho(Y_{c_2})$, for all non-decreasing ρ and $\gamma \geq 0$.

Next, we provide the proof for the reverse direction. Since $\mathbb{E}[\mathbf{X}_{c_1}] - \gamma\rho(Y_{c_1}) \geq \mathbb{E}[\mathbf{X}_{c_2}] - \gamma\rho(Y_{c_2})$ for all $\gamma \geq 0$, $\rho(Y_{c_1}) \leq \rho(Y_{c_2})$ for all non-decreasing ρ, which implies $Y_{c_2} \circledS Y_{c_1}$. □

Theorem 4.3 highlights that regret dominance is necessary and sufficient for mean-regret preferences for all non-decreasing ρ when regret is separable. However, the results of Theorem 4.3 may not hold for non-separable regret. Example 5 highlights that regret dominance is not sufficient for mean-regret preferences under relative regret.

Example 5: Relative regret in mean-regret models. Consider two choices, c_1 and c_2, and two mutually exclusive events E_1 and E_2 such that $P(E_1) = P(X_{c_1} = 1, X_{c_2} = 2) = 0.5$ and $P(E_2) = P(X_{c_1} = 5, X_{c_2} = 3) = 0.5$. Computing the relative regret of c_1 and c_2 (i.e., $Y_c = \frac{\max_{i \in S} X_i}{X_c} - 1$), we get $P(Y_{c_1} = 0) = P(Y_{c_1} = 1) = P(Y_{c_2} = 0) = P(Y_{c_2} = 0.67) = 0.5$, which implies that $Y_{c_1} \circledS Y_{c_2}$. However, $\mathbb{E}[X_{c_1}] = 3 > \mathbb{E}[X_{c_2}] = 2.5$. Since $\mu_{c_1} > \mu_{c_2}$, a mean-regret decision maker with a sufficiently small γ will prefer c_1, even though it regret dominates c_2. $\square$

CHAPTER 5
EQUIPMENT REPLACEMENT ANALYSIS WITH AN UNCERTAIN FINITE HORIZON

Equipment replacement analysis is concerned with identifying the length of time to retain an asset and its replacements over some designated horizon of analysis such that capital and operating costs are minimized. The reasons for replacing an asset are generally economic, including deterioration of the asset itself, leading to increased operating and maintenance (O&M) costs and decreased salvage values, and technological advances in potential replacement assets on the market that may provide similar service at lower cost.

While much of the vast literature on equipment replacement analysis is concerned with the modeling of deterioration and technological change (see, for example, Oakford et al. (1984), Bean et al. (1994) and Regnier et al. (2004)), it should be clear that the selected horizon can have a drastic effect on the solution approach and resulting solutions (de Sousa and Guimaraes 1997). In an infinite horizon problem with stationary costs, the optimal policy is to repeatedly replace the asset at its economic life, the age which minimizes equivalent annual costs. If the horizon is finite, then the policy of replacing an asset at its economic life is only optimal if the horizon is a multiple of the economic life. Otherwise, a sequence of asset replacements must be found, generally through the use of dynamic programming (see Bellman (1955), Wagner (1975), or Hartman and Murphy (2006)).

In the case of nonstationary costs, constant age replacement policies are rarely optimal (see Fraser and Posey (1989)) for infinite horizon problems. Thus, most solutions focus on finding the optimal time zero decision. Generally, a forward dynamic programming algorithm is solved such that the initial replacement decision (keep or replace the asset at time zero) is optimal for any horizon greater than a designated solution or forecast (finite) horizon T (see Bean et al. (1985) and Chand and Sethi (1982)).

We are concerned with an equipment replacement problem that has a finite horizon, but the horizon length is unknown at time zero, and may not become known until reached. We term this problem the equipment replacement problem under an uncertain finite horizon. We have encountered a number of instances in practice in which this truly captures reality, including: (1) A production line is to be closed, but continues to operate while customers place orders; (2) The expected life of a mine is extended through the discovery of new ore deposits; (3) Makeshift solutions are provided while infrastructure is built, such as a company providing bus service for its workers while a tram system is constructed; and (4) A military base is slated for closing, but a date has not been set and most likely will be debated heavily by government officials.

In the above examples, assets are utilized in production or service to meet capacity needs and the required length of service is finite, but uncertain. It should be noted that the examples illustrate cases where the required service can extend beyond or fall short of the expected length of service. Furthermore, the length of time in which an asset is needed in these instances can be short (less than a year) or long (nearly a decade). In cases where the expected service requirements are short, it is conceivable that a company would merely utilize current assets over the remaining periods of service. However, when the required length of service is extensive or highly uncertain, then it should be clear that equipment replacement decisions should be considered over time in the interest of saving money. In addition, the company may explore options outside of traditional ownership, such as leasing.

The use of dynamic programming has become prevalent in solving equipment replacement problems because it easily accounts for the dynamic nature of the problem (keep or replace decision is made periodically) and it overcomes the traditional assumption that assets are repeatedly retained for the same length of time. While previously developed dynamic programming formulations differ, they can all be represented by acyclic graphs in which the states of the system are represented by nodes and the arcs represent decisions. As each decision carries a cost, the goal is to find the shortest (minimum cost) path through the network, beginning with an asset of known age at time zero and making optimal decisions through the horizon time T. In our problem, the value of T is uncertain. (Note that this is not an optimal stopping problem as the decision-maker cannot choose the finite horizon T. Rather, it is a product of the environment.)

We specifically define the equipment replacement problem with an uncertain finite horizon as follows: An asset is required to be in service for T periods where T is a random variable such that $T_s \leq T \leq T_l$. The values of T_s and T_l are known, along with the probability distribution of T. It is assumed that the asset is sold at the end of the realized horizon. At time zero, an asset of age n is owned and it, and its replacements, may be retained through age N, if not replaced sooner. It is assumed that the equipment can perform in its intended capacity in each period over the horizon.

As highlighted previously, the optimal replacement decision is dependent on the horizon and there may be a period before T_s where the optimal decision diverges for different horizon realizations. For example, it may be optimal to retain the equipment at the end of period 2 if $T = 7$ but if $T = 8$, the optimal decision may be to replace it at the end of period 2. We term period 2, t_c, the *critical decision period*. If the uncertainty of the horizon is not resolved by t_c, the risk of loss cannot be eliminated.

We assume that the costs for purchasing, operating and disposing of the asset and its potential replacements over time are known. Thus, we are not faced with a network of stochastic arc lengths, which is fairly common in the literature. Rather, we are faced with a network with an uncertain destination (state), which is not as common in the literature.

While we believe our problem is unique, there exists problems in the literature with similar characteristics. A significant number of studies consider the uncertain lifetime of equipment (see Sivazlian (1977), for example) that is replaced upon failure, such as an appliance. That is, economic replacement decisions are not made periodically over time as in capital equipment replacement analysis.

Using a mathematical approach similar to failure analysis, Pliska and Ye (2007) determined optimal insurance purchases and income consumption for a wage earner. Clearly, the lifetime of the person is uncertain and finite. However, these decisions are only made through the person's retirement age while we examine decisions through the uncertain horizon.

Trietsch (1985) studied an application in military target planning. The decisions include the trajectory and speed of the aircraft given potential targets upon departure until the ultimate destination is made known at some point on the route. As the problem is continuous, it takes on the characteristics of a dynamic facility location problem, as one must not head towards a possible destination, as in our discrete problem.

In this chapter, we study the equipment replacement problem under an uncertain horizon in detail. First we examine a traditional problem which minimizes cost. Next, we turn to finding a robust solution which minimizes maximum regret. In addition, we propose and design an option contract to lease an asset to hedge against the risks that result from the horizon uncertainty. While leasing options have been noted in the equipment replacement literature (see Hartman and Lohmann (1997)), to our knowledge, they have not been proposed to hedge risk.

In the next two sections, we illustrate how to obtain the optimal replacement policy that minimizes expected cost and also the optimal replacement policy that minimizes the maximum regret. After which, we illustrate how an optimal option contract, that generates revenue for the leasor and reduces the risk for the leasee, can be designed.

5.1 Minimizing Expected Cost

Define $f_t(i)$ as the minimum expected cost of making optimal replacement decisions for an i-period old asset at time t through the horizon $T \in [T_s, T_l]$. We assume that at the end of period $t \in [T_s, T_l - 1]$, it is made known whether the problem continues to time $t + 1$. The decision is whether to keep (K) or replace (R) the asset after each period. The asset is sold at the end of the realized horizon.

O&M costs for an $i-$period old asset at time t are defined as $m_t(i)$ while salvage values (assumed to be revenues) are defined as $s_t(i)$. The purchase cost for a new asset in period t is defined as c_t. There is only one available challenger (replacement asset) in each period. Asset purchases and sales are assumed to occur at the beginning of the period while O&M costs occur at the end of the period. All costs are discounted according to the periodic discount factor α.

The dynamic programming recursion is defined as stochastic in periods T_s through T_l and deterministic in periods $0 \leq t < T_s$. Define $P(T = t|T \geq t)$ as the probability that the horizon is equal to t (the problem ends) given that it is time t. The recursion is defined as follows for all values of $i = 1, 2, \ldots, N$, noting that if $i = N$, the asset must be replaced:

$$\begin{aligned} f_t(i) &= (1 - P(T = t|T \geq t)) \left(\min \begin{cases} \text{K:} & \alpha \left[m_{t+1}(i+1) + f_{t+1}(i+1) \right], \\ \text{R:} & c_t - s_t(i) + \alpha \left[m_{t+1}(1) + f_{t+1}(1) \right] \end{cases} \right) \\ &+ P(T = t|T \geq t)(-s_t(i)), \quad T_s \leq t < T_l \qquad \text{(5–1)} \\ f_{T_l}(i) &= -s_{T_l}(i) \qquad \text{(5–2)} \end{aligned}$$

Solution of the above defines values for $f_{T_s}(i)$ for all values of i. This is used as input to the deterministic part of the recursion, defined as follows:

$$f_t(i) = \min \begin{cases} \text{K:} & \alpha \left[m_{t+1}(i+1) + f_{t+1}(i+1) \right], \\ \text{R:} & c_t - s_t(i) + \alpha \left[m_{t+1}(1) + f_{t+1}(1) \right] \end{cases}, \quad 0 \leq t < T_s \qquad \text{(5–3)}$$

This formulation may be visualized in the network in Figure 5-1 with an initial asset of age $n = 2$ and a stochastic horizon between $T_s = 6$ and $T_l = 8$. The nodes, representing the state of the system, are labeled with the age of the asset and the arcs represent keep (connecting node i in period t to node $i + 1$ in period $t + 1$) and replace (connecting node n in period t to node 1 in period $t + 1$) decisions. At the end of periods

6, 7 and 8, arcs connect to node T, representing the termination of the problem, and, for periods 6 and 7, an intermediate node, signaling that the problem may continue for at least another period. Two arcs, representing keep and replace decisions, emanate from each intermediate node. Note that the arcs emanating from the state nodes at periods 6 and 7 are probabilistic.

Example 6: Minimizing Expected Cost. Consider a eight-year old asset owned at time zero. The asset must be in service for at least 6 years (T_s), but the service may last for a maximum of 8 years (T_l). A new asset can be purchased in any period for \$150,000. O&M costs for a new asset are \$15,000 in the first year, increasing at a compounding rate of 10% each year while the salvage values decline 20% per year. An asset can be retained for 20 years, at which time it must be replaced, such that it does not have to be replaced in the specified time frame. The annual interest rate is 10% and the probability distribution of the horizon is $P(T = 6) = P(T = 7) = 0.2$ and $P(T = 8) = 0.6$. Note that these probabilities are stationary and are assumed to not change with time, defining the conditional probabilities of $P(T = t|T \geq t) = 0.2, 0.25$ and 1, for $t = 6, 7$ and 8, respectively.

The associated network for this problem is illustrated in Figure 5-1. Note that in period 8, there are nine possible end conditions (ages 1 through 8 and 16). These ages define the boundary conditions (5–2) listed in Table 5-2.

Working through recursions (5–1) and (5–3) defines the optimal solution of replacing the asset at the end of the second period. Thus, for the various possible horizons of $T = 6,$ 7 and 8, the retired asset is age 4, 5, or 6, respectively. To provide some frame of reference, note that the economic life of this asset is 11 years.

The expected discounted cost of this policy is \$206,987. If it is known that $T = 6$, the optimal replacement policy is to retire the initial asset at the end of the sixth period. The discounted cost of this policy is \$171,661. When $T = 7$, it is optimal to replace the initial asset at the end of the third period and retire it at the end of the seventh period.

The discounted cost of this policy is \$200,161. When $T = 8$, replacing the initial asset at the end of the second period and retiring it at the end of the eighth period results in a minimum discounted cost of \$218,392. Thus, the policy which minimizes expected costs follows the deterministic solutions for $T = 8$, but diverges when $T = 6$ or $T = 7$. □

This leads us to investigate whether more robust policies are possible and whether the risk of loss from different horizon realizations can be mitigated. Also note that this example defines period 2 as the critical decision period, because the optimal paths of decisions diverge at this point for different realizations of the horizon.

5.2 Minimizing Maximum Regret

While minimizing expected costs is an appropriate objective function in many situations, it may not be appropriate for a risk-averse decision-maker. That is, the policy determined when minimizing costs may expose the possibility of great losses if certain realizations occur. Here, we explore more robust solutions by minimizing maximum regret. The regret for each horizon realization is the difference between the optimal solution, given that realization, and the cost obtained from the chosen replacement policy.

Obtaining the optimal solution for each possible horizon is straightforward as a deterministic dynamic program can be solved forwards from an initial condition, proceeding with keep and replace decisions through each possible horizon length. This is the method employed in many equipment replacement problems which guarantee an optimal time zero solution for the infinite horizon problem (see Bean et al. (1985)).

The forward dynamic programming recursion that obtains the policy which minimizes maximum regret across the uncertain horizon is defined with the following initial conditions at time 0:

$$g_0(i) = \begin{cases} -s_0(i), & i = n \\ \infty, & \text{otherwise} \end{cases} \tag{5–4}$$

and the recursion is defined as follows for the remaining periods t, where $1 \leq t \leq T$:

$$g_t(i) = \begin{cases} K : \alpha^{t-1} \left[s_{t-1}(i-1) + \alpha \left(m_t(i) - s_t(i) \right) \right] + g_{t-1}(i-1), & 2 \leq i \leq N \\ R : \alpha^{t-1} \left[c_t + \alpha \left(m_t(1) - s_t(1) \right) \right] + \min_{j=1,2,\ldots,N} \left\{ g_{t-1}(j) \right\}, & i = 1. \end{cases} \tag{5–5}$$

The recursion follows, as described in Park and Sharp (1990), in that the costing assumes the asset is sold at the end of each period. Modeled this way, the problem can be terminated after any period. If the problem continues, keeping an asset requires "purchasing" the asset at its salvage value, thus negating the previous sale for a net cash flow transaction of zero. If the asset is replaced, then a new asset is purchased with the salvage value for the old asset already received.

Note that the network (Figure 5-1) for the backward recursion is the same here. However, the recursion (5–5) differs in that the first expression relates to an asset that is kept (moves from age $i-1$ to i) while the second expression finds the minimum path resulting in a new asset ($i = 1$) at time t.

The advantage of solving the forward deterministic dynamic program in our setting is that it defines the optimal costs through any period $T \in \{T_s, T_l\}$ with an asset of age i, which is required for the computation of regret. To find the minimal maximum regret path, one could enumerate all possible paths from period 0 to period T_l. However, we can reduce the number of paths by noting that the minimum cost path is taken from period 0 to T_s, as the problem is deterministic in these periods. Therefore, it is only necessary to enumerate the paths from T_s to T_l. In addition, it may not be necessary to explicitly enumerate all possible paths. When two sub-paths end at the state, the dominated sub-path may be eliminated from further consideration. A dominated sub-path has a higher expected cost and a higher maximum regret. An algorithm for finding the minimal maximum regret is presented in Tan and Hartman (2010). Note that in the worst case, it is possible, although unlikely, that there is no dominated sub-path, defining the maximum number of paths to be explicitly evaluated as $2^{T_l - T_s}N$.

Example 7: Minimizing Maximum Regret. Consider Example 6 again. Solving Equation (5–5) with this data defines the discounted cost solutions through periods 6, 7 and 8 as shown in Table 5-4. The minimum cost solution in each period is in bold.

Suppose that $i = 1$ at the end of period 6. Keeping that asset for the next two periods would result in an asset of ages 2 and 3 at the end of periods 7 and 8, respectively. Examining the data in Table 5-4, the regret experienced would be \$3,241, \$1,681 and \$5,353 for periods 6, 7 and 8, respectively. These values are merely the differences between the functional values $g_6(1)$, $g_7(2)$ and $g_8(3)$ and the optimal solutions of $g_6(14)$, $g_7(4)$ and $g_8(6)$, respectively.

Suppose that we want to reach state 1 of period 7 (i.e., replacement decision is made) by passing through state 1 of period 6. The cost of this path is \$202,450 + (\$174,902 - \$171,661) = \$205,691. The term in the parenthesis is the additional cost that has been incurred as a result of reaching state 1 of period 7 via a non-minimum cost path. The maximum regret of this path is $\max\{\$174,902 - \$171,661,\ \$205,691 - \$200,161\} = \$5,530$. Observe that the path ending at state 1 of period 7 which passes through state 14 of period 6 has a cost of \$202,450 and a maximum regret of \$2,289. This path dominates the path which passes through state 1 of period 6 and hence, the dominated path can be discarded.

Using the algorithm presented in Tan and Hartman (2010), the policy that minimizes maximum regret is to retain the initial asset for four periods and keep the replacement through the horizon. Thus, the asset age at the end of horizons of length 6, 7 and 8, is 2, 3, and 4, respectively. This results in a minimal maximum regret of \$5,112 from period 6.

Note that the policy which minimizes maximum regret carries an expected cost of \$208,084, which is greater than the minimum possible expected cost. However, the optimal policy for minimizing expected costs exposes the owner to a potential loss of

\$7,855 at $T = 6$. Thus, the policy which minimizes expected costs may not always be the most attractive to a risk-averse decision-maker. Note that no replacement policy can eliminate all potential losses in this example. □

It is clear that replacement decisions require extensive forecasting of data, including future expected costs. We re-examine the previous examples with sensitivity analysis to illustrate how the solutions change with different data forecasts.

Example 8: Sensitivity Analysis. To gain a better understanding of the problem, the replacement policies that minimize expected cost and maximum regret are recomputed for different rates of increasing O&M costs and rates of decreasing salvage values. In addition, we consider two different initial asset ages and two different horizon realization distributions.

The results are illustrated in Tables 5-7 through 5-10, where the optimal replacement policy is described by listing the (end of) periods in which replacements occur. A "-" represents no replacements over the horizon. Columns E[Cost] and MR list the expected cost and maximum regret of the policy, respectively. Tables 5-7 and 5-8 consider an initial asset of age 4 while Tables 5-9 and 5-10 consider an initial asset of age 8.

Of the 36 scenarios examined, there were 14 scenarios in which the regret could not be eliminated. These tend to occur when the rates of the O&M cost increases and salvage value decreases are high. This is reasonable as retaining an asset beyond its economic life can be quite costly in these situations. Note that regret is also more likely when starting with an older asset, even when the rates are not as high.

The different objectives of minimizing expected costs and minimizing maximum regret led to eight different replacement policies in the 36 scenarios analyzed. The differences in maximum regrets in those eight cases varied from a from a mere 5.6% to over 140%. Clearly, minimizing expected cost alone may not be appropriate for a risk-averse decision-maker.

It is also interesting to note the variance in the age of the asset during the uncertain periods for each of the scenarios. For the scenarios in which a regret must be incurred, the tendency is to replace the asset at some point before T_s, such that the asset is relatively young when entering the uncertain periods. These are also the cases in which the equivalent annual cost curve defining the economic life of the asset is steep, defining greater risk of regret. For the flatter curves and longer economic lives, the tendency is to retain the assets longer, thus entering the uncertain periods with older assets. Note that the initial age of the asset does play a strong role in this outcome. □

5.3 Leasing Options to Further Mitigate Risk

The previous section highlights a critical issue: risk can be reduced by minimizing maximum regret, but unless the associated policy aligns with the optimal policy for each horizon realization, the risk of loss cannot be eliminated. However, it may be possible to lower the risk if a vendor is willing to provide the decision-maker with the option to lease an asset at a favorable rate if the required horizon extends beyond T_s. The vendor will, of course, charge a premium for the flexibility that it provides to the decision-maker.

Here, we design a leasing option contract that can further mitigate the owner's risk exposure. Specifically, (i) we design a leasing contract according to length and price; (ii) price an option to make the lease available; and (iii) determine whether this contract is beneficial to both the vendor (leasor) and decision-maker (leasee). Note that we design the option contract such that a decision must be made at the critical time period (t_c), the first period where optimal policies for different horizon realizations diverge.

Consider the network representation of the dynamic programming formulation in Figure 5-1. To include a lease option, the decisions at the critical time period become twofold: (i) should the option contract be purchased and (ii) should the asset be kept or replaced. If the option is not purchased, then the network is defined as before. However, if the lease option is purchased, then another (parallel) network is traversed in which leasing becomes a viable option after period T_s. Note that the purchase of the option contract gives the owner the right to exercise it at a later date and lease an asset; however, the owner is not required to do so.

Suppose that the decision-maker purchases and exercises an option to lease at the end of period T_s through the horizon. Let x denote the price of the option (discounted to period 0) and y_t denote the leasing rate for period t. Our goal is to minimize the maximum regret z for the decision-maker. Assuming that the leasor (vendor) requires some minimum expected profit (net present value) ω, we can solve for the leasing contract variables as follows:

$$\text{Minimize} \quad z \tag{5–6}$$

$$\text{Subject to} \quad z \geq \left(\min_i g_{T_s}(i) + x + \sum_{i=T_s+1}^{t} \alpha^i y_t \right) - \min_j \; g_t(j) \qquad t = T_s, \ldots, T_l \tag{5–7}$$

$$\omega \leq x + \sum_{t=T_s+1}^{T_l} P(T \geq t)\alpha^t (y_t - \lambda) \tag{5–8}$$

$$x, y_t \geq 0 \qquad t = T_s + 1, \ldots, T_l \tag{5–9}$$

$$z \in R \tag{5–10}$$

The objective function (5–6) minimizes the maximum regret, which is defined in Constraints (5–7). The term in the parentheses defines the total cost incurred through leasing for each realized period of the lease, $T_s + 1$ through the horizon, while $\min_j g_t(j)$ defines the minimum cost incurred when owning the asset through the horizon. Constraint (5–8) defines the profit for the vendor, according to the decision variables x and y_t, as being greater than ω. The value of λ captures the periodic costs of the vendor.

This contract design can lower the decision-maker's risk. The following theorem quantifies the decision-maker's exposure.

Theorem 5.1. *If:*

$$g_t(j) \quad \textit{is non-decreasing in t,} \tag{5–11}$$

and

$$\omega \geq \sum_{t=T_s+1}^{T_l} P(T \geq t)\alpha^t(\lambda - y_t), \tag{5–12}$$

there exists an optimal leasing contract (minimizes maximum regret) where $z^* = x^*$.

Proof. Consider the problem where we relax the non-negativity constraints on x and y_t. Observe that the relaxed problem is a linear program with $T_l - T_s + 2$ decision variables with the same number of independent constraints. This implies that at the optimal solution, all constraints are binding. Solving the system of linear equations yields the following solution:

$$y_t^* = \alpha^{-t}\left(\min_i g_t(i) - \min_i g_{t-1}(i)\right) \qquad t = T_s + 1, ..., T_l \tag{5–13}$$

$$z^* = x^* = \omega - \sum_{t=T_s+1}^{T_l} P(T \geq t)\alpha^t(y_t^* - \lambda) \tag{5–14}$$

If Condition (5–11) holds, the cost of the optimal replacement policy without the leasing option is non-decreasing in the horizon and hence all y_t^* are nonnegative (from Equation (5–13)). If Condition (5–12) holds, it follows from Constraint (5–8) that x^* is nonnegative. Therefore under Conditions (5–11) and (5–12), the x^*, y_t^* and z^*, as defined by Equations (5–13) and (5–14) are optimal (since they also satisfy the relaxed constraints) and $z^* = x^*$. □

The set of conditions described in Theorem 5.1 are reasonable. The first condition holds if the O&M costs are nonnegative and the salvage value of the asset is non-increasing with age (Tan and Hartman, 2010), which is commonly assumed in replacement analysis. The second condition simply implies that the expected profit of the vendor cannot be too low.

The implication of Theorem 5.1 is that a leasing contract can be constructed such that all possible regret is eliminated from the problem, with the exception of the cost of the contract, which must be incurred. However, it is unlikely that the lease rates, y_t, will vary each period – unless systematically. In addition, the decision-maker may consider initiating a lease at some period later than T_s. Here we assume that the lease rates are fixed once a lease is initiated, but the rate can differ for contracts of different lengths. Let y denote the constant lease rate and k denote the period in which the asset is first leased (at the end of the period). It is assumed that once the asset is leased, it is leased through the end of the horizon.

As the vendor's profit is defined by the difference in revenues received minus costs paid, observe that the following relationship between x and y must hold:

$$\underbrace{x}_{\text{option price}} = \underbrace{-\left(\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t\right) y}_{\text{expected lease revenue}} + \underbrace{\left(\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t\right) \lambda}_{\text{expected vendor costs}} + \underbrace{\omega}_{\text{expected vendor profit}} . \tag{5–15}$$

The option price can be viewed as the combination of three separate components: expected lease revenue; expected vendor costs; and expected vendor profit. Equation (5–15) implies that if λ and ω are constant, a unit increase in x decreases y by exactly $\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t$.

Suppose that the decision-maker follows a replacement policy π and $r_j^\pi, j \leq k$, is the regret (price of the option not included) incurred if $T = j$. Recall that y_i^* is the optimal "variable lease rate" for period i and the leasing option is not exercised until the end of period k. Let a denote the period with the highest regret, r_a^π, (minus the option price) and b the period with the lowest variable lease rate, $y_b^* = \min_t y_t^*$ where y_t^* is defined in Theorem 5.1. Note that $a \leq k$ and $b > k$. λ and ω are assumed to be non-negative. The following theorem, Theorem 5.2, provides the optimal contract parameters under certain conditions, as presented in the theorem statement below:

Theorem 5.2. *For a replacement policy π through period $k \geq T_s$, if:*

$$0 < \sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t \leq 1, \tag{5–16}$$

and

$$\omega > \left(\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t \right) (y_b^* + r_a^\pi - \lambda), \tag{5–17}$$

there exists an optimal leasing contract where

$$x^* = \omega + (\lambda - y^*) \left(\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t \right), \quad y^* = y_b^* + r_a^\pi \quad \text{and} \quad z^* = x^* + r_a^\pi.$$

Proof. Note that Condition (5–17), with Equation (5–15), guarantees that $x > 0$. By definition, r_a^π is the maximum regret experienced in all periods through k. As y_b^* is the minimum lease rate, it guarantees that no regret is experienced in periods greater than k. (The option price is ignored in these two statements.) Therefore, the minimum, maximum regret z^* that is achievable is $x^* + r_a^\pi$. Note that increasing y to $y_b^* + r_a^\pi$ decreases x while not increasing the regret experienced, as r_a^π is already experienced through period k. Further increasing y decreases x, but at a slower rate than increasing the regret in periods after k, from Condition (5–16). Therefore, $y^* = y_b^* + r_a^\pi$ and $z^* = x^* + r_a^\pi$. The definition of x^* follows from Equation (5–15). □

If the conditions stated in Theorem 5.2 do not hold, the option price is driven to zero, as in the following corollary.

Corollary 5. *For a replacement policy π through period $k \geq T_s$, if:*

$$\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t > 1, \tag{5–18}$$

or

$$\omega \leq \left(\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t \right) (y_b^* + r_a^\pi - \lambda), \tag{5–19}$$

the optimal option price $x^* = 0$.

Proof. When $y \leq y_b^* + r_a^\pi$, the maximum regret occurs at $T = a$. When $y \geq y_b^* + r_a^\pi$, the maximum regret occurs at $T = b$. Hence it follows that the maximum regret will always occur at either period a or period b.

Given Condition (5–18), the regret at period a and b is increasing in x. Therefore, it is optimal to set x at its lowest possible value, i.e., $x^* = 0$.

Given Condition (5–19), y must be less than or equal to $y_b^* + r_a^\pi$ for $x \geq 0$ to hold. This implies that the maximum regret occurs at $T = a$. Since the maximum regret occurs at $T = a$, the maximum regret is minimized by setting x at the lowest possible value, i.e., $x^* = 0$. □

Under the stated conditions, Theorem 5.2 provides the optimal option price and the corresponding lease rate for a given k-period replacement policy. In addition, it also follows from the theorem that the replacement decisions for periods less than k should be those that minimize the maximum regret across periods of ownership.

Figures 5-2 and 5-3 examine the tradeoff in option price and other parameters more closely. When α is high, there is little discounting on the cost of leasing in the future and hence the vendor will need to charge a high option price. This is illustrated in Figure 5-2. Similarly, if k is small, the vendor will need to charge a high option price to account for the possibility of a long lease. This is illustrated in Figure 5-3. Therefore in these situations, the high option price that is required may not be justified and the decision-maker may be better off either leasing without an option or to not lease at all.

Corollary 5 illustrates that the optimal option price is zero when the stated conditions do not hold. In this situation, it may still be beneficial for the decision-maker to lease the equipment, but the decision-maker is not able to further reduce his or her risk by paying an upfront premium to reduce the lease rate.

Two issues remain. First, we do not know the optimal k. However, we can easily find the optimal solution for each k value with Theorem 5.2 and identify the best answer from this set. Second, we need to determine appropriate ω and λ values. We assume that the option contract is attractive to the vendor if its expected profit is non-negative. Under this assumption, we set $\omega = 0$, although in practise some margin may be required. The value of λ is usually unknown to the decision-maker, however the decision-maker should be able to reasonably estimate λ with its own cost data. In the following example, we examine the option contract design.

Example 9: Option contract. Again consider Example 6. There are two possible option lease contracts, (i) leasing for periods 7 and 8 ($k = 6$) and (ii) leasing for period 8 only ($k = 7$). For this example, ω is assumed to be 0.

For (i), the policy which minimizes the maximum regret for a 6-period problem is to salvage the initial asset at the end of period 6 such that $r_6 = \$0$. The optimal variable lease rates are $y_7^* = \$55,540$ and $y_8^* = \$39,080$. Since $\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t = (0.8)(1.1)^{-7} + (0.6)(1.1)^{-8} = 0.690 < 1$, Condition (5–16) holds. If $\lambda > \min\{y_7^*, y_8^*\} + r_6 = \$39,080 + \$0 = \$39,080$, Condition (5–17) holds and the optimal option price and lease rate are $0.690\lambda - \$26,982$ and $\$39,080$ respectively. Since $r_6 = \$0$, the maximum regret is also $0.690\lambda - \$26,982$. If $\lambda \leq \$39,080$, Condition (5–19) holds and the optimal option price, lease rate and maximum regret are $\$0$, λ and $\$0$ respectively. These are summarized in Table 5-5.

For (ii), the policy which minimizes the maximum regret for a 7-period problem is to salvage the initial asset at the end of the horizon (period 7) such that $r_6 = \$0$ and $r_7 = \$1,746$. The optimal variable lease rate for period 8 is $\$39,080$. Since $\sum_{t=k+1}^{T_l} P(T \geq t)\alpha^t = (0.6)(1.1)^{-8} = .280 < 1$, Condition (5–16) holds. If $\lambda > y_8^* + \max\{r_6, r_7\} = \$39,080 + \$1,746 = \$40,825$, Condition (5–17) holds and the optimal option price and lease rate are $0.280\lambda - \$11,427$ and $\$40,825$ respectively. Since $\max\{r_6, r_7\} = \$1,746$, the maximum regret is also $0.280\lambda - \$9,681$. If $\lambda \leq \$40,825$, Condition (5–19) holds and the optimal option price, lease rate and maximum regret are $\$0$, λ and $\$1,746$ respectively. These are summarized in Table 5-6.

When $\lambda > \$42,142$, policy (ii) results in a smaller maximum regret than policy (i). Note that the economic life of the asset was 11 years which was defined from an annual equivalent cost of \$45,494. Hence it is likely that λ is going to be greater than \$42,142 and policy (ii) will be the optimal policy. □

Table 5-1. Conditional probabilities of the various scenarios.

t	$P(T=t \mid T \geq t)$
6	0.40
7	0.67
8	1.00

Table 5-2. Boundary conditions at $t=8$.

i	$f_8(i)$	i	$f_8(i)$	i	$f_8(i)$
1	-\$120,000	4	-\$61,440	7	-\$31,457
2	-\$96,000	5	-\$49,152	8	-\$25,166
3	-\$76,800	6	-\$39,322	16	-\$4,222

Table 5-3. Boundary conditions at $t=6$.

i	$f_6(i)$	i	$f_6(i)$	i	$f_6(i)$
1	-\$40,404	4	\$103	7	\$18,098
2	-\$24,662	5	\$8,754	8	\$21,101
3	-\$11,446	6	\$14,094	10	\$25,043

Table 5-4. Discounted costs through periods 6, 7 and 8 periods with an asset of age i.

i	$g_6(i)$	$g_7(i)$	$g_8(i)$
1	\$174,902	\$202,450	\$228,152
2	\$176,773	\$201,843	\$226,942
3	\$178,093	\$200,866	\$223,745
4	\$179,515	**\$200,161**	\$220,928
5	\$181,530	\$200,243	\$219,005
6	\$184,456	\$201,494	**\$218,392**
7		\$204,146	\$219,394
8			\$222,185
14	**\$171,661**		
15		\$201,907	
16			\$231,876

Table 5-5. x^*, y^* and z^* values for policy (i).

λ	x^*	y^*	z^*
$\leq 39,080$	0	λ	0
$> 39,080$	$0.690\lambda - 26,982$	$39,080$	$0.690\lambda - 26,982$

Table 5-6. x^*, y^* and z^* values for policy (ii).

λ	x^*	y^*	z^*
$\leq 40,825$	0	λ	$1,746$
$> 40,825$	$0.280\lambda - 11,427$	$40,825$	$0.280\lambda - 9,681$

Table 5-7. Optimal policies: $P(T=6)=0.6$, $P(T=7)=0.2$, $P(T=8)=0.2$ and $n=4$.

Rates		**Econ.**	**Min E[Cost] Policy**			**Min Max Regret Policy**		
O&M	Salvage	**Life**	Policy	E[Cost]	MR	Policy	E[Cost]	MR
0.3	0.3	5	(1,7)	\$247,668	\$2,889	(1,7)	\$247,668	\$2,889
0.3	0.2	4	(0,3,7)	\$202,757	\$1,994	(0,4)	\$202,825	\$821
0.3	0.1	1	(1,2, . . . ,7)	\$111,007	\$0	(1,2, . . . ,7)	\$111,007	\$0
0.2	0.3	7	(2)	\$213,548	\$0	(2)	\$213,548	\$0
0.2	0.2	6	(1)	\$182,504	\$2,425	(2)	\$182,545	\$982
0.2	0.1	1	(1,2, . . . ,7)	\$111,007	\$0	(1,2, . . . ,7)	\$111,007	\$0
0.1	0.3	12	(-)	\$129,836	\$0	(-)	\$129,836	\$0
0.1	0.2	11	(-)	\$124,030	\$0	(-)	\$124,030	\$0
0.1	0.1	7	(2)	\$99,997	\$0	(2)	\$99,997	\$0

Table 5-8. Optimal policies: $P(T = 6) = 0.2$, $P(T = 7) = 0.2$, $P(T = 8) = 0.6$ and $n = 4$.

Rates		**Econ.**	**Min E[Cost] Policy**			**Min Max Regret Policy**		
O&M	Salvage	**Life**	Policy	E[Cost]	MR	Policy	E[Cost]	MR
0.3	0.3	5	(2,7)	$274,872	$5,791	(1,7)	$275,165	$2,889
0.3	0.2	4	(0,4)	$224,389	$821	(0,4)	$224,389	$821
0.3	0.1	1	(1,2, . . . ,7)	$128,641	$0	(1,2, . . . ,7)	$128,641	$0
0.2	0.3	7	(2)	$231,736	$0	(2)	$231,736	$0
0.2	0.2	6	(2)	$202,429	$982	(2)	$202,429	$982
0.2	0.1	1	(1,2, . . . ,7)	$128,641	$0	(1,2, . . . ,7)	$128,641	$0
0.1	0.3	12	(-)	$146,377	$0	(-)	$146,377	$0
0.1	0.2	11	(-)	$141,715	$0	(-)	$141,715	$0
0.1	0.1	7	(2)	$116,358	$0	(2)	$116,358	$0

Table 5-9. Optimal policies: $P(T = 6) = 0.6$, $P(T = 7) = 0.2$, $P(T = 8) = 0.2$ and $n = 8$.

Rates		**Econ.**	**Min E[Cost] Policy**			**Min Max Regret Policy**		
O&M	Salvage	**Life**	Policy	E[Cost]	MR	Policy	E[Cost]	MR
0.3	0.3	5	(0,6)	$281,870	$5,499	(0,5)	$283,674	$5,205
0.3	0.2	4	(0,3,7)	$239,031	$1,994	(0,4)	$239,099	$821
0.3	0.1	1	(1,2, ... ,7)	$144,852	$0	(1,2, ... ,7)	$144,852	$0
0.2	0.3	7	(0)	$250,336	$0	(0)	$250,336	$0
0.2	0.2	6	(0,7)	$222,678	$1,330	(0,7)	$222,678	$1,330
0.2	0.1	1	(1,2, ... ,7)	$144,852	$0	(1,2, ... ,7)	$144,852	$0
0.1	0.3	12	(-)	$192,459	$0	(-)	$192,459	$0
0.1	0.2	11	(6)	$188,875	$8,550	(4)	$190,422	$5,112
0.1	0.1	7	(0)	$135,044	$0	(0)	$135,044	$0

Table 5-10. Optimal policies: $P(T=6)=0.2$, $P(T=7)=0.2$, $P(T=8)=0.6$ and $n=8$.

Rates		**Econ.**	**Min E[Cost] Policy**			**Min Max Regret Policy**		
O&M	Salvage	**Life**	Policy	E[Cost]	MR	Policy	E[Cost]	MR
0.3	0.3	5	(0,5)	$306,515	$5,205	(0,5)	$306,515	$5,205
0.3	0.2	4	(0,4)	$260,663	$821	(0,4)	$260,663	$821
0.3	0.1	1	(1,2, . . . ,7)	$162,486	$0	(1,2, . . . ,7)	$162,486	$0
0.2	0.3	7	(0)	$271,930	$0	(0)	$271,930	$0
0.2	0.2	6	(0,6)	$245,211	$1,752	(0,7)	$245,489	$1,330
0.2	0.1	1	(1,2, . . . ,7)	$162,486	$0	(1,2, . . . ,7)	$162,486	$0
0.1	0.3	12	(-)	$215,980	$0	(-)	$215,980	$0
0.1	0.2	11	(2)	$206,987	$7,855	(4)	$208,084	$5,112
0.1	0.1	7	(0)	$153,841	$0	(0)	$153,841	$0

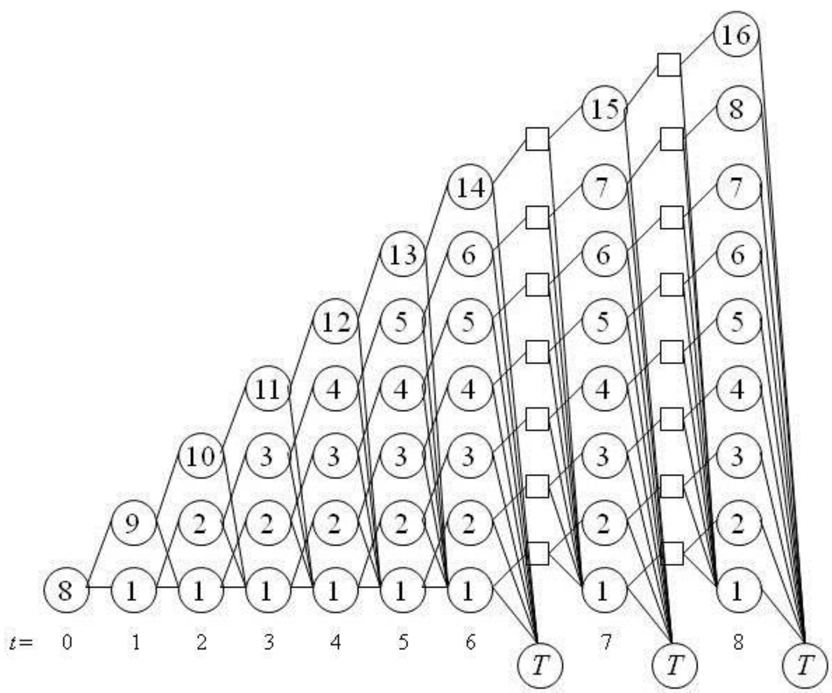

Figure 5-1. Visualization of dynamic programming formulation with $n = 8$, $T_s = 6$ and $T_l = 8$.

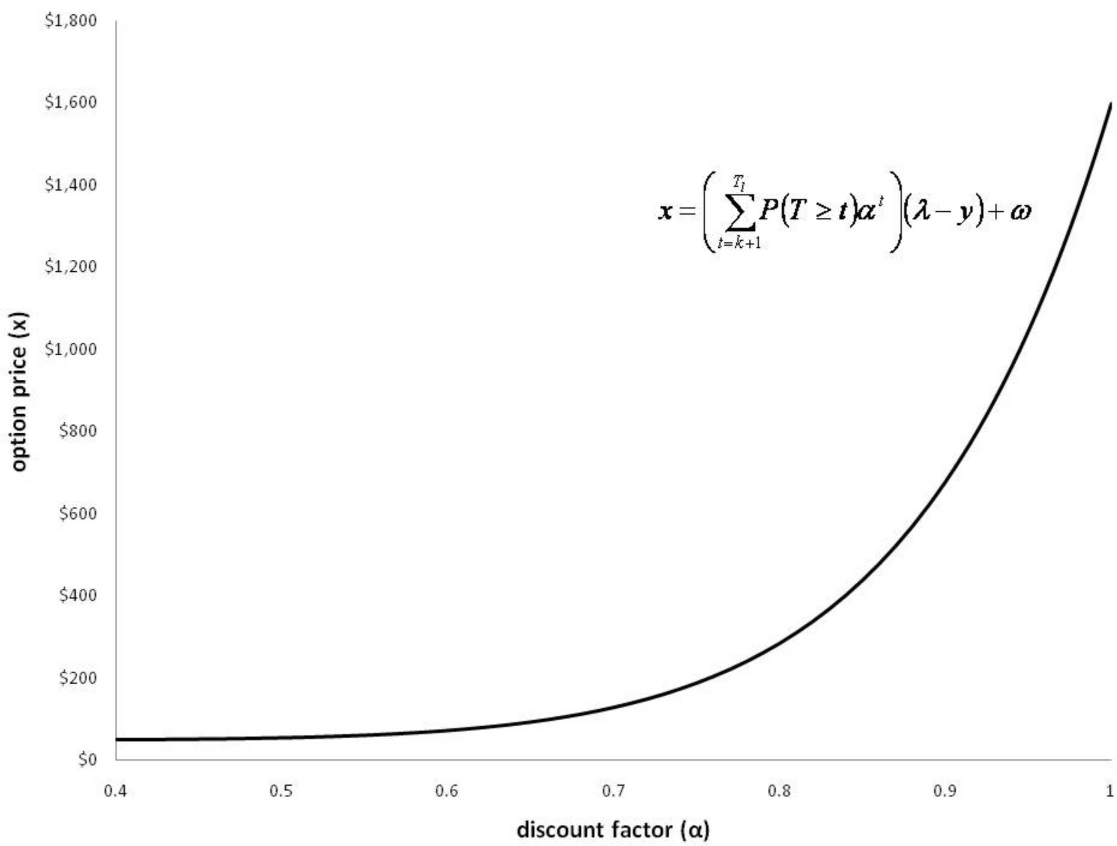

Figure 5-2. Effect of the discount rate on the option price.

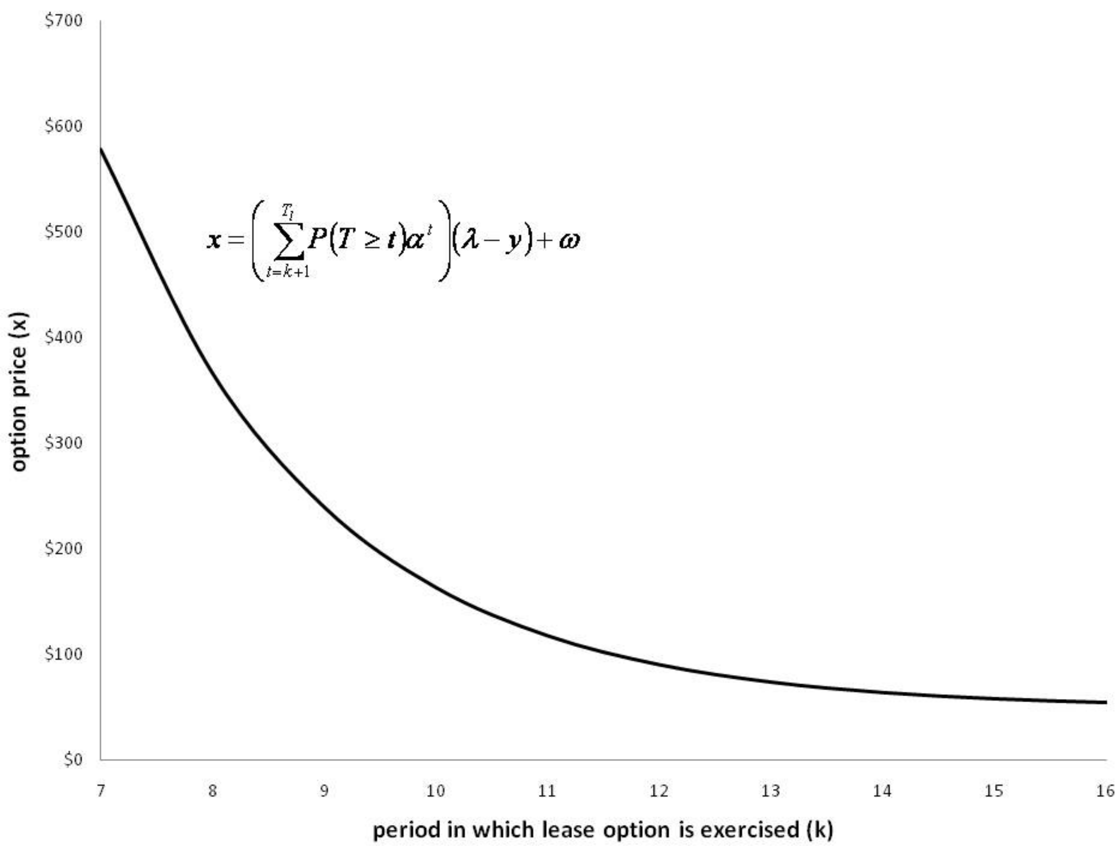

Figure 5-3. Effect of the number of lease periods on the option price.

CHAPTER 6
OPTIMAL PRODUCT LINE FOR CONSUMERS THAT EXPERIENCE REGRET

Individuals have different preferences and many will agree that it is better to have more choices than less. For example, Nike allows its customers to design their own shoes, Dell sells custom built laptops and Subway lets its customers decide what goes into a sandwich. Empirical studies on the relationship between product variety and firm profits suggest that greater product variety can lead to higher profits (see, for example Kotler and Keller (2006) and Nestessine and Taylor (2007)), especially when production, supply chain and marketing strategies are aligned (see Kekre and Srinivasan (1990), Safizadeh et al. (1996), Berry and Cooper (1999) and Randall and Ulrich (2001)). For-profit companies are not the only organizations that recognize the need to offer variety. The United States government allows employers to automatically enroll employees in 401(k) pension plans, but employees can select different funds, or even opt out of the plan completely. The Singapore government requires every citizen and permanent resident to be enrolled in a savings plan, but allows individuals to determine how their savings are invested.

The problem of determining an optimal product line has been studied extensively in the optimization literature. An optimal product line is obtained by solving an optimization problem where the benefits and costs of introducing variety are modeled analytically. The benefits are often measured in terms of welfare or revenue, which is based on the utility and price of the product, respectively. It is generally assumed that the utility of a product is known or follows some known probability distribution. The price of a product is either fixed (see, for example, Van Ryzin and Mahajan (1999)) or obtained endogenously (see, for example, Aydin and Porteus (2008)). Production and inventory holding costs are two cost components that are commonly considered in these models. Dobson and Kalish (1993) considered a model with fixed and variable production costs. Desai et al. (2001) studied the tradeoffs between lower manufacturing costs and lower selling prices that result from commonalities within a product line. Nestessine and Taylor (2007) found that manufacturers with inefficient production technologies can have lower production costs than those with more efficient production systems as it is optimal for the former to offer lower quality products. Alptekinoglu and Corbett (2008) analyzed the competition between a mass customizer and a mass producer within a game-theoretic context and found that a mass producer can, under certain scenarios, coexist with a mass customizer that has a more efficient production system. Kok et al. (2008) provides an excellent review of the literature that considers the inventory aspect of the problem.

Over the last two decades, researchers have proposed various consumer choice models, including the multinomial logit (MLN) model, exogenous demand model and locational choice model. In the MLN model, the utility of a product to a specific consumer is decomposed into deterministic and stochastic components. Van Ryzin and Mahajan (1999) showed that the optimal assortment consists of the most popular products under static choice assumptions (i.e., the consumer will only purchase the most attractive product from the assortment if it is available) and considered a model where the consumer purchases the best available alternative in Mahajan and van Ryzin (2001). Cachon et al. (2005) considered a model where the consumer might choose to explore other stores, even when an acceptable product is available. When similar products are available in different retail stores, the value of exploring decreases, resulting in wider product assortments when consumer search considerations are included. Smith and Agrawal (2000) considered an exogenous demand model where the consumer purchases an alternative product according to some substitution probability distribution if their best choice product is not available. Kok and Fisher (2007) illustrated how the probabilities that a consumer will purchase a substitute product can be estimated from sales data. In the locational choice model, consumers select products that are closest to his or her preference, as defined on an attribute space (e.g., percentage of fat content in milk). Researchers that have considered locational choice models include de Groote (1994), Chen et al. (1998), Gaur and Honhon (2006), Alptekinoglu and Corbett (2008) and Kuksov and Villas-Boas (2010). These models are discussed in greater detail in Kok et al. (2008). Recently, Honhon et al. (2010) proposed a choice model which generalizes the MLN and locational choice models.

However, production and inventory costs are not the only incentives for limiting the number of products offered. Researchers have highlighted that individuals may be happier when choices are limited. For example, Tversky and Shafir (1992) found that having a large choice set causes individuals to defer their decisions or choose default options. In two separate experiments, Iyengar and Lepper (2000) found that people were more likely to purchase gourmet jams and chocolates when a limited array of choices is presented. In a third experiment, they found that students were more likely to undertake optional class essay assignments when they were given 6, rather than 30, potential essay topics from which to choose. Furthermore, individuals who were presented with a limited array of choices reported greater satisfaction with their choices. It has also been observed that people are unhappy, indecisive or may even not choose at all when faced with a choice among several alternatives (see Tversky (1972), Shafir et al. (1993), Redelmeier and Shafir (1995) and Brenner et al. (1999)). For further research on product variety and consumer satisfaction, the interested reader is referred to Chernev (2003), Schwartz (2004) and Scheibehenne (2008).

Research in psychology and economics suggests that the satisfaction associated with a choice is influenced by feelings of regret (Connolly and Zeelenberg, 2002). In addition, choices may also be influenced by the regret that one anticipates (see Zeelenberg (1999) and Engelbrecht-Wiggans and Katok (2009)). The effect of regret on consumer decisions has been studied in a variety of settings. Braun and Muermann (2004) examined the optimal insurance purchases of regret-averse consumers. Irons and Hepburn (2007) were interested in the optimal choice set for a specific individual with regret considerations under three search scenarios: (i) a predetermined number of choices are examined; (ii) the decision-maker dynamically decides whether to continue the search or select from the set of examined choices; and (iii) after a choice has been examined, the decision-maker either selects that choice and stops or rejects it and proceeds to examine the next choice (i.e., no recall). Syam et al. (2008) considered a problem where a consumer has to decide between purchasing a standard or custom product. Engelbrecht-Wiggans and Katok (2008) looked at how feedback information affects bidding behavior. Recently, Nasiry and Popescu (2011) studied the effects of anticipated regret within an advance selling context.

We solve for the optimal product line using a model that accounts for the regret that is anticipated and experienced. Unlike previous research, the satisfaction associated with a choice is dependent on the entire product line, rather than the selected product alone. In our model, consumers are uncertain about the true utility of a product. Villas-Boas (2009) considered a similar problem where the uncertainties can be resolved by performing a search or evaluation and found that the optimal number of products to offer is decreasing in evaluation costs. Kuksov and Villas-Boas (2010) highlighted that search and evaluation costs can cause consumers to refrain from searching or evaluating when the number of alternatives are high or low. We consider a problem where the uncertainties in the utility of the products are unresolved at the time of purchase. This is true for many products, ranging from consumer goods like cars and laptops to services like medical treatment and financial planning. To the best of our knowledge, the problem of finding the optimal product line under regret considerations has never been explored. In practice, the anticipated regret and the regret experienced may differ. For example, an optimistic investor may underestimate the regret associated with a bad investment. For simplicity, we assume that the difference between the two is negligible. This assumption is reasonable since individuals are often able to give reasonably accurate predictions of their feelings (Loewnstein and Schkade, 1999).

This chapter proceeds as follows. First, we present our model. Next, we study the optimal number of products to offer under various limiting conditions. In addition, we run various numerical experiments and derive a tight upper bound on the optimal number of products to offer.

6.1 Model Description

Consider a market with N segments, where each segment is defined by a unique customer type. Let α_j denote the market share of segment j such that $\alpha_j = \frac{1}{N}$ for $j = 1, 2, ..., N$ if all market segments are equally large. We consider a problem where the utility of a product is uncertain and only resolved after purchase according to some scenario s. Let S denote the set of possible scenarios, p_s denote the probability of scenario s occurring and u_{ijs} denote the utility of choice i for segment j under scenario s. Without loss of generality, we assume that u_{ijs} is positive. In addition, we assume that u_{ijs} is finite.

Let x and X denote a product and the set of products that the firm can offer, respectively. We assume that a non-negative finite fixed cost c_i is incurred if product i is made available. The firm decides on a set of products, $\pi \subseteq X$, to be made available (i.e., product line) and Π denotes the set of possible product lines. Let r^{π}_{ijs} denote the regret experienced by segment j under scenario s when choice $i \in \pi$ is chosen:

$$r^{\pi}_{ijs} = \max_{x \in \pi} u_{xjs} - u_{ijs}, \tag{6–1}$$

Note that r^{π}_{ijs} is, by definition, non-negative. Define "net satisfaction" as the weighted combination of the utility received and the regret experienced:

$$v^{\pi}_{ijs} = u_{ijs} - \gamma_j r^{\pi}_{ijs}, \tag{6–2}$$

where γ_j is some non-negative constant that represents the regret aversion of segment j and v^{π}_{ijs} is the net satisfaction of segment j with choice i under scenario s and product line π. The weighted combination of the utility received and the regret experienced has been referred to as the agent's utility and net utility in Irons and Hepburn (2007) and Syam et al. (2008), respectively. We believe that the term net satisfaction is more appropriate in our context.

6.1.1 Consumer Choice Model

Clearly, the choices of consumers are dependent on the products that are available in the product line. Let $\pi(j)$ denote the product that is selected by segment j when product line π is offered. We assume that consumers are rational and select products that maximize their individual expected net satisfaction:

$$\pi(j) = \arg\max_{i\in\pi} \sum_s p_s \ v_{ijs}^{\pi}. \tag{6–3}$$

Syam et al. (2008) proposed a similar choice model in studying consumer preferences for standardized and customized products. A similar choice model was also adopted by Nasiry and Popescu (2011) in studying the purchasing behavior of rational consumers that experience regret under advance selling. Finding $\pi(j)$ with Equation (6–3) is challenging because it involves the regret that is experienced by consumers.

Theorem 6.1. *Given Equation (6–3),* $\pi(j) = \arg\max_{i\in\pi} \sum_s p_s \ u_{ijs}$.

Proof.

$$\begin{aligned}
\pi(j) &= \arg\max_{i\in\pi} \sum_s p_s \ v_{ijs}^{\pi} \\
&= \arg\max_{i\in\pi} \sum_s \left[p_s \ u_{ijs} - \gamma_j r_{ijs}^{\pi}\right] \\
&= \arg\max_{i\in\pi} \sum_s \left[p_s \ u_{ijs} - \gamma_j \left(\max_{x\in\pi} u_{xjs} - u_{ijs}\right)\right] \\
&= \arg\max_{i\in\pi} \sum_s \left[p_s \ (1+\gamma_j) u_{ijs} - \gamma_j \max_{x\in\pi} u_{xjs}\right] \\
&= \arg\max_{i\in\pi} \sum_s p_s \ u_{ijs}
\end{aligned}$$

The last equality follows from the fact that $\arg\max_{i\in\pi} \sum_s p_s \ (1+\gamma_j) u_{ijs} = \arg\max_{i\in\pi} \sum_s p_s \ u_{ijs}$ and $\gamma_j \max_{x\in\pi} u_{xjs}$ is independent of i. □

Theorem 6.1 states that rational consumers that maximize expected net satisfaction select products that result in the highest expected utility from the product line. Although anticipated regret can complicate the decision process (i.e., decision maker accounting for the regret that is anticipated), it does not affect the final decision of a rational consumer that is described by our choice model (i.e., Equation (6–3)).

In our model, we assume that all consumers will purchase a product. This is true if all products result in positive expected utility for each consumer. If that is not true, a consumer may choose to not make any purchase when none of the products are satisfactory (i.e., all products result in negative expected utility). To allow for this, we include a dummy product with zero fixed costs and zero utility for all consumers across all scenarios in our product set X when necessary. Note that the problem can always be scaled by a positive factor. Hence, it does not violate our assumption that u_{ijs} is positive.

6.1.2 Problem Formulation

We are interested in finding a product line π^* that maximizes the expected net satisfaction of the market, while accounting for the fixed costs of introducing products:

$$\pi^* = \arg\max_{\pi \in \Pi} f(\pi),$$

where $f(\pi)$ is the expected performance of product line π defined as:

$$f(\pi) = \sum_j \sum_s \alpha_j \, p_s \, \left(u_{\pi(j)js} - \gamma_j r^{\pi}_{\pi(j)js}\right) - \sum_{i \in \pi} c_i x_i.$$

When $\gamma_j = 0$ for all j, the problem reduces to an uncapacitated facility location problem, which is NP-hard (Cornuejols et al. 1990). Hence, the general problem is also NP-hard.

We solve this problem by formulating it as a binary linear program. First, we define a couple of notations. Let w^{π}_{js} denote the maximum utility of segment j under scenario s and product line π:

$$w^{\pi}_{js} = \max_{i \in \pi} u_{ijs}.$$

The net satisfaction of segment j under scenario s and product line π (i.e., Equation (6–2)) can be re-expressed as follows:

$$v_{ijs}^{\pi} = (1+\gamma)u_{ijs} - \gamma_j w_{js}^{\pi}.$$

Note that w_{js}^{π} is independent of the consumer's choice. The optimal product line can be obtained by solving the following binary linear programming formulation for the product line design problem with regret considerations, **PLD-R**:

$$\max \sum_j \sum_s \alpha_j \, p_s \left[(1+\gamma_j)\left(\sum_i u_{ijs} y_{ij}\right) - \gamma_j w_{js}\right] - \sum_i c_i x_i \tag{6–4}$$

$$\text{s.t. } w_{js} \geq u_{ijs} x_i \qquad \forall i, j, s \tag{6–5}$$

$$y_{ij} \leq x_i \qquad \forall i, j \tag{6–6}$$

$$\sum_i y_{ij} \leq 1 \qquad \forall j \tag{6–7}$$

$$y_{ij} \geq 0 \qquad \forall i, j \tag{6–8}$$

$$x_i \in \{0, 1\} \qquad \forall i \tag{6–9}$$

$$w_{js} \in \mathbb{R} \qquad \forall j, s \tag{6–10}$$

where the decision variables x_i and y_{ij} denote if choice i is available and if choice i is selected by segment j, respectively. The unrestricted decision variables w_{js}, defined by Constraint (6–5), denote the maximum utility of segment j under scenario s. Constraint (6–6) ensures that only available choices are selected. Constraint (6–9) ensures that x_i are binary. Since y_{ij} are variables that indicate if a choice is selected by a segment, they should also be binary. Theorem 6.2 highlights that Constraints (6–7) and (6–8) collectively ensure that there is an optimal solution where each consumer makes exactly one selection.

Theorem 6.2. *There exists an optimal solution to **PLD-R** where $y_{ij}^{*} \in \{0, 1\}$ for all i and j.*

Proof. We prove this by contradiction. Assume there does not exist an optimal solution to **PLD-R** where $y^*_{ij} \in \{0, 1\}$ for all i and j. First, we note that it follows from Constraints (6–7) and (6–8) that $0 \leq y_{ij} \leq 1$. It follows from our assumption that there exists i' and j' such that $0 < y^*_{i'j'} < 1$. Since $\sum_s \alpha_j p_s (1 + \gamma_j) u_{ijs}$ (i.e., the coefficient of y_{ij}) is positive and $y^*_{i'j'}$ is optimal, $\sum_s \alpha_j p_s (1 + \gamma_j) u_{i'js} = \sum_s \alpha_{j'} p_s (1 + \gamma_{j'}) u_{i'j's}$ for all j where $y^*_{i'j} > 0$. Setting $y^*_{i'j'} = 1$ and $y^*_{i'j} = 0$ for all $j \neq j'$ maintains optimality (since all positive y variables have the same coefficient) and feasibility (i.e., does not violate any constraints in **PLD-R**). This argument applies to all i, contradicting the assumption that there does not exist an optimal solution where $y^*_{ij} \in \{0, 1\}$ for all i and j. □

Theorem 6.2 highlights that there exists an optimal solution where y^*_{ij} is binary for all i and j. In addition, it follows from the proof of Theorem 6.2 that an optimal solution where y^*_{ij} is binary for all i and j can be easily obtained by arbitrarily setting a non-zero y variable to take the value 1 and the other non-zero y variables associated with the same choice to take the value 0.

6.2 Optimal Number of Products to Offer

In this section, we study the relationships between various problem parameters and the optimal number of products to offer. Let $|\pi^*|$ denote the number of products in an optimal product line. When multiple optimal solutions exist, let $|\pi^*|$ denote the number of products in the smallest optimal product line. Since a non-negative fixed cost is incurred when a product is introduced into the product line, $|\pi^*|$ is likely to be smaller when fixed costs are high. We begin this section by showing that $|\pi^*|$ is non-increasing when the fixed cost of introducing products increases uniformly.

Theorem 6.3. *$|\pi^*|$ is non-increasing in Δ, where Δ is a uniform increase in the fixed cost of introducing each product.*

Proof. Let $\Pi_{>|\pi^*|}$ denote the set of product lines that are strictly larger than $|\pi^*|$ (i.e., $\Pi_{>|\pi^*|} = \{\pi : |\pi| > |\pi^*|\}$) . We prove this theorem by showing that $f(\pi^*) > f(\pi)$ for all

$\pi \in \Pi_{>|\pi^*|}$ when each c_i is increased by Δ. Since π^* is optimal when fixed costs are c_i:

$$\begin{aligned}
\sum_j \sum_s \alpha_j p_s v^{\pi^*}_{\pi^*(j)js} - \sum_{i \in \pi^*} c_i &\geq \sum_j \sum_s \alpha_j p_s v^{\pi}_{\pi(j)js} - \sum_{i \in \pi} c_i, \forall \pi \in \Pi \\
\sum_j \sum_s \alpha_j p_s v^{\pi^*}_{\pi^*(j)js} - \sum_{i \in \pi^*} c_i &\geq \sum_j \sum_s \alpha_j p_s v^{\pi}_{\pi(j)js} - \sum_{i \in \pi} c_i, \forall \pi \in \Pi_{>|\pi^*|} \\
\sum_j \sum_s \alpha_j p_s v^{\pi^*}_{\pi^*(j)js} - \sum_{i \in \pi^*} c_i - |\pi^*|\Delta &> \sum_j \sum_s \alpha_j p_s v^{\pi}_{\pi(j)js} - \sum_{i \in \pi} c_i - |\pi|\Delta, \forall \pi \in \Pi_{>|\pi^*|} \\
f(\pi^*) &> f(\pi), \forall \pi \in \Pi_{>|\pi^*|},
\end{aligned}$$

The third inequality follows from the fact that $|\pi^*|\Delta < |\pi|\Delta$ for all $\pi \in \Pi_{>|\pi^*|}$. □

Theorem 6.3 states that $|\pi^*|$ is non-increasing when the fixed cost of introducing a product increases uniformly, which suggests that $|\pi^*|$ is likely to be small when fixed costs are high. We note that the result of Theorem 6.3 is general, holding for all problem instances. Intuitively, one would expect a similar relationship between $|\pi^*|$ and the regret aversion of consumers (i.e., smaller $|\pi^*|$ when γ_j are high). However, that is not true, even when all consumers are described by the same regret aversion parameter γ. This is illustrated in Example 10.

Example 10: $|\pi^*|$ increases with γ. Consider a problem where $N = 3$, $|X| = 5$, $|S| = 3$, $c_1 = c_2 = c_3 = c_4 = c_5 = 0$, $p_1 = p_2 = p_3 = \frac{1}{3}$ and $\alpha_1 = \alpha_2 = \alpha_3 = \frac{1}{3}$. The respective u_{ijs} values are listed in Table 6-1. Given this information, we solve **PLD-R**.

When $\gamma = 0$, product line $\{4, 5\}$ is optimal and $f(\{4, 5\}) = 4.73$. When $\gamma = 0.2$, product line $\{1, 2, 3\}$ is optimal and $f(\{1, 2, 3\}) = 4.6$. □

Example 10 highlights that $|\pi^*|$ can increase when consumers become more regret averse. This occurs when a small set of risky products (i.e., highly variable outcomes) is appealing to the market. However, these products result in high regret and a larger set of less risky products (for example, a standard product with minor modifications for different customer types) can result in greater market satisfaction when regret aversion increases. This is discussed in greater detail in Section 7.2.

Although $|\pi^*|$ can increase when regret aversion increases, one would expect $|\pi^*|$ to be non-increasing in γ for most problems. Next, we look at how the expected optimal number of products to offer, $\mathbb{E}[|\pi^*|]$, varies with γ, N, S and X by generating a series of different problems in the following manner:

- Each u_{ijs} follows a probability distribution U with finite mean and upper support U_{max}. In addition, each u_{ijs} is independent and identically distributed.
- $p_s = \frac{q_s}{\sum_s q_s}$, where each q_s is positive and follows a probability distribution Q with finite mean. In addition, each q_s is independent and identically distributed.
- $\alpha_j = \frac{a_j}{\sum_s a_j}$, where each a_j is positive and follows a probability distribution A with finite mean. In addition, each a_j is independent and identically distributed.
- u_{ijs}, q_s and α_j are independently distributed for all i, j and s.

First, we show that offering a single product can be optimal under various limiting cases. Next, we present our experimental results for various sets of problem parameters. Third, we compute a tight upper bound on $\mathbb{E}[|\pi^*|]$.

6.2.1 Limiting Cases

In this section, we study various limiting cases. In particular, we show that offering a single product is likely to be optimal when γ, $|S|$ or $|X|$ approaches infinity. Furthermore, we highlight that the expected optimal number of products to offer under these conditions is 1.

Theorem 6.4. $\lim_{\gamma\to\infty} |\pi^*| = 1$.

Proof. This is a proof by contradiction. Assume that $\lim_{\gamma\to\infty} |\pi^*| > 1$. This implies that $\lim_{\gamma\to\infty} f(\pi^*) > \lim_{\gamma\to\infty} f(\pi)$ for all $\pi \in \Pi_{|1|}$, where $\Pi_{|1|} = \{\pi : |\pi| = 1\}$. Note that $\lim_{\gamma\to\infty} f(\pi)$ is finite for all $\pi \in \Pi_{|1|}$. This follows from the fact that $f(\pi)$ is independent of γ for all $\pi \in \Pi_{|1|}$ and u_{ijs}, α_j and c_i are finite. Since $\lim_{\gamma\to\infty} f(\pi^*) > \lim_{\gamma\to\infty} f(\pi)$ for all $\pi \in \Pi_{|1|}$ and $\lim_{\gamma\to\infty} f(\pi)$ is finite for all $\pi \in \Pi_{|1|}$, $\lim_{\gamma\to\infty} f(\pi^*)$ is bounded from below by a finite constant. This implies that $r^{\pi^*}_{ijs} = 0$ for all i, j and s because $\lim_{\gamma\to\infty} f(\pi^*) = -\infty$ if some $r^{\pi^*}_{ijs} > 0$. It follows from Equation (6–1) that $r^{\pi^*}_{ijs} = 0$ for all i, j and s if and only if $u_{ijs} = u_{kjs}$ for all $j \in N$, $s \in S$ and $i, k \in \pi^*$. Since all choices in π^* result in the same utility across all scenarios for all segments and c_i is non-negative, there exists a $\pi_1 \in \Pi_{|1|}$ such that $f(\pi_1) \geq f(\pi_2)$ for all $\pi_2 \in \Pi_{r=0} = \{\pi : r^{\pi}_{ijs} = 0, \ \forall i, j, s\}$, contradicting the assumption that $|\pi^*| > 1$. $\square$

Theorem 6.4 states that it is optimal to offer a single product when γ approaches infinity. This implies that the firm should offer a single product when consumers are heavily influenced by feelings of regret. This is because the likelihood of experiencing regret increases with the number of products offered. It follows from Theorem 6.4 that $\mathbb{E}[|\pi^*|] = 1$ as γ approaches ∞.

Corollary 6. $\lim_{\gamma \to \infty} \mathbb{E}[|\pi^*|] = 1$.

Proof. Follows immediately from Theorem 6.4. $\square$

Next, we show that there almost surely exists an optimal product line that consists of a single product when $|S|$ approaches ∞.

Theorem 6.5. $\lim_{|S| \to \infty} P(|\pi^*| = 1) = 1$.

Proof. The expected utility of segment j when choice i is chosen is $\sum_s p_s u_{ijs}$:

$$
\begin{aligned}
\lim_{|S| \to \infty} \sum_s p_s u_{ijs} &= \lim_{|S| \to \infty} \frac{\sum_s q_s u_{ijs}}{\sum_s q_s} \\
&= \frac{\lim_{|S| \to \infty} \sum_s q_s u_{ijs}}{\lim_{|S| \to \infty} \sum_s q_s} \\
&\overset{p}{=} \frac{\lim_{|S| \to \infty} \frac{\mathbb{E}[QU]}{|S|}}{\lim_{|S| \to \infty} \frac{\mathbb{E}[Q]}{|S|}} && \text{(by the law of large numbers)} \\
&= \frac{\mathbb{E}[QU]}{\mathbb{E}[Q]} \\
&= \frac{\mathbb{E}[Q] \cdot \mathbb{E}[U]}{\mathbb{E}[Q]} && \text{(by the independence of } Q \text{ and } U\text{)} \\
&= \mathbb{E}[U],
\end{aligned}
$$

where $=^p$ denotes "equality with probability 1".

Since each p_s and each u_{ijs} are independent and identically distributed, all choices result in the same expected utility (i.e., $\mathbb{E}[U]$) for each segment and a product line that only consists of the product with the lowest fixed cost is optimal with probability 1 when $|S| \to \infty$. $\square$

Theorem 6.5 states that there almost surely exists an optimal product line that consists of a single product when $|S| \to \infty$. When $|S|$ is high, there is high uncertainty in the true utility of each product. Under this condition, decision makers are likely to be indifferent between the choices and hence, it is optimal to offer a single product. It follows from Theorem 6.5 that $\mathbb{E}[|\pi^*|] = 1$ as $|S|$ approaches ∞.

Corollary 7. $\lim_{|S| \to \infty} \mathbb{E}[|\pi^*|] = 1$.

Proof. Follows immediately from Theorem 6.5. $\square$

Theorem 6.6 states that there also almost surely exists an optimal product line that consists of a single product when $|X| \to \infty$.

Theorem 6.6. *If $P(U = U_{max}) > 0$, $\lim_{|X| \to \infty} P(|\pi^*| = 1) = 1$.*

Proof. Consider a choice i' where $u_{i'js} = U_{max}$ for all j and s. Since $P(U = U_{max}) > 0$ and each u_{ijs} is independent and identically distributed, there is a positive probability that $i' \in X$. Hence, $\lim_{|X| \to \infty} P(i' \in X) = 1$. Since $\{i'\}$ is optimal and $\lim_{|X| \to \infty} P(i' \in X) = 1$, $\lim_{|X| \to \infty} P(|\pi^*| = 1) = 1$. $\square$

A product line consisting of a single product that achieves maximum utility for all consumers across all scenarios is clearly optimal. When the number of products that a firm can offer is infinitely large, there almost surely exists such a product within the set of products that the firm can offer. Therefore, there almost surely exists an optimal product line that consists of a single product when $|X| \to \infty$. The result of Theorem 6.6 is based on the assumption that each product has a positive probability of achieving maximum utility, which is reasonable for computer-based simulations. It follows from Theorem 6.6 that $\mathbb{E}[|\pi^*|] = 1$ as $|X|$ approaches ∞.

Corollary 8. *If* $P(U = U_{max}) > 0$, $\lim_{|X|\to\infty} \mathbb{E}[|\pi^*|] = 1$.

Proof. Follows immediately from Theorem 6.6. □

6.2.2 Experimental Results

We perform our experiments by generating each u_{ijs} from a uniform distribution with minimum and maximum values, 5000 and 10000, respectively. In order to obtain a larger range of $|\pi^*|$ values, we study problems where fixed costs are low (since $|\pi^*|$ is likely to be 1 when fixed costs are high). However, multiple solutions are likely to exist when $c_i = 0$ for all i. In our experiments, we set $c_i = 0.01$ for all i to reduce the number of problem instances with multiple optimal solutions. Q and A follow a discrete uniform distribution with a range of [1, 32767] and [10,20], respectively. The large range of values that q_s can take allows for the modeling of improbable scenarios while a_j takes a small range of values since the effects of very small market segments can usually be ignored. We consider 3 levels for each parameter that we study (i.e., γ, $|S|$, N and $|X|$) and solve 20 instances with each problem parameter configuration. The problems were solved using CPLEX 12.1 on a 3.40GHz Intel Pentium CPU with 2GB RAM. The solution time was the greatest when $N = |X| = |S| = 30$ and $\gamma = 0.3$, averaging 51.6 minutes per problem instance. Estimates of the expected optimal number of products to offer, $\hat{\mathbb{E}}[|\pi^*|]$, are obtained by computing the average $|\pi^*|$ of 20 problem instances and are listed in Table 6-2.

The experiments highlight that $|\pi^*|$ is generally non-increasing with γ and $|S|$. As γ increases, the effect of regret is amplified and a smaller product line is more attractive. As $|S|$ increases, there is greater uncertainty in the true utility of a product, which generally results in an increase in the likelihood of experiencing regret. Therefore a small product line is desirable when $|S|$ is large.

The experiments also indicate that $|\pi^*|$ is generally non-decreasing with N when γ and $|S|$ are small. For example, $\hat{\mathbb{E}}[|\pi^*|] = 8.45, 13.60$ and 16.96 under the problem settings ($\gamma = 0.1, |S| = 10, N = 10, |X| = 30$), ($\gamma = 0.1, |S| = 10, N = 20, |X| = 30$) and ($\gamma = 0.1, |S| = 10, N = 30, |X| = 30$), respectively. This result is intuitive. When there is little uncertainty in the utility of products and consumers are not influenced by regret, it is optimal to offer a large product line to a diverse market. This supports the conventional belief that consumers are better off with more choices. However, the relationship between $|\pi^*|$ and N is reversed when γ and $|S|$ are large. For example, $\hat{\mathbb{E}}[|\pi^*|] = 2.10, 1.10$ and 1.00 under the problem settings ($\gamma = 0.5, |S| = 20, N = 10, |X| = 30$), ($\gamma = 0.5, |S| = 20, N = 20, |X| = 30$) and ($\gamma = 0.5, |S| = 20, N = 30, |X| = 30$), respectively. When γ and $|S|$ are small, the effects of regret are small and hence the benefits (i.e., higher utility) of having a large product line outweigh the costs (i.e., higher regret). However, the effects of regret are pronounced when γ and $|S|$ are high. Under these conditions, having a large product line increases consumer regret and it is better for the firm to offer a small number of products by focusing on selected market segments.

The relationship between $|\pi^*|$ and $|X|$ is driven by two conflicting forces. When $|X|$ is high, there is a higher chance that there will be a product that is "customized" for a particular segment. However, there is also a higher chance that there will be a product that is appealing to a wide portion of the market. The former suggests that $|\pi^*|$ is non-decreasing with $|X|$, while the latter suggests that $|\pi^*|$ is non-increasing with $|X|$. Our experimental results indicate that $|\pi^*|$ is generally non-decreasing with $|X|$ for small problems. However, Theorem 6.6 highlights that $|\pi^*|$ is almost surely 1 when $|X|$ is very large.

6.2.3 Upper Bound on $\mathbb{E}[|\pi^*|]$

In this section, we provide a tight upper bound for $\mathbb{E}[|\pi^*|]$. We begin by deriving a recursive expression for $\mathbb{E}[|\pi^*|]$ when there are no fixed costs (i.e., $c_i = 0$ for all i) and regret considerations (i.e., $\gamma_j = 0$ for all j).

Theorem 6.7. *If $c_i = 0$ for all i and $\gamma_j = 0$ for all j:*

$$\mathbb{E}[|\pi^*|] = \sum_{n=1}^{\min\{|X|,N\}} nP(|\pi^*| = n),$$

where:

$$P(|\pi^*| = n) = \begin{cases} \frac{1}{|X|^{N-1}} & n = 1 \\ \binom{|X|}{n}\left(\frac{n}{|X|}\right)^N - \sum_{k=1}^{n-1} \binom{|X|-k}{n-k} P(|\pi^*| = k) & n > 1. \end{cases}$$

Proof. It follows from Theorem 6.1 that $|\pi^*|$ is equal to the number of choices that yields the highest expected utility for at least one segment. Since each u_{ijs} is independent and identically distributed, each choice has an equal chance of being most preferred (i.e., yielding the highest expected utility) by a segment. Specifically, the probability that choice i is most preferred by segment j is $\frac{1}{|X|}$.

First, we consider the $n = 1$ case (i.e., offer single product). The probability that choice i is most preferred by all segments i is $\frac{1}{|X|^N}$. Since there are $|X|$ choices:

$$P(|\pi^*| = 1) = |X| \cdot \frac{1}{|X|^N} = \frac{1}{|X|^{N-1}}.$$

Next, we consider the general n case (i.e., offer n products). Let $P(\pi^* \subseteq \pi)$ denote the probability that π contains choices that are most preferred by all segments. The probability that a particular product line of size n contains a choice that is most preferred by segment j is $\frac{n}{|X|}$. Furthermore, $P(\pi^* \subseteq \pi \mid |\pi| = n) = \left(\frac{n}{|X|}\right)^N$. In addition, note that there are $\binom{|X|}{n}$ product lines of size n. For example, consider a problem where $|X| = 3$. There are $\binom{3}{2} = 3$ product lines of size 2 (i.e., $\{0, 1\}$, $\{0, 2\}$ and $\{1, 2\}$). However, $P(\pi^* = \{0\})$ is accounted for in both $P(\pi^* \subseteq \{0, 1\})$ and $P(\pi^* \subseteq \{0, 2\})$. In particular, a specific product line of size k is a subset of $\binom{|X|-k}{n-k}$ product lines of size n, $n > k$. This implies that $P(|\pi^*| = k)$ is accounted for $\binom{|X|-k}{n-k}$ times in $P(\pi^* \subseteq \pi \mid |\pi| = n)$. Hence, we obtain the following recursive equation:

$$P(|\pi^*| = n) = \binom{|X|}{n}\left(\frac{n}{|X|}\right)^N - \sum_{k=1}^{n-1}\binom{|X|-k}{n-k}P(|\pi^*| = k).$$

Since $1 \leq |\pi^*| \leq \min\{|X|, N\}$:

$$\mathbb{E}[|\pi^*|] = \sum_{n=1}^{\min\{|X|,N\}} nP(|\pi^*| = n).$$

$\square$

Theorem 6.7 can be used to compute the expected optimal number of products to offer when there are no fixed costs and regret considerations. It follows from Theorem 6.3 that the value obtained by Theorem 6.7 is also an upper bound on the expected optimal product line size for problems with uniform non-negative fixed costs and no regret considerations. This result is expressed in Corollary 9.

Corollary 9. *If there exists some $c \geq 0$ where $c_i = c$ for all i and $\gamma_j = 0$ for all j:*

$$\mathbb{E}[|\pi^*|] \leq \sum_{n=1}^{\min\{|X|,N\}} nP(|\pi^*| = n),$$

where:

$$P(|\pi^*| = n) = \begin{cases} \frac{1}{|X|^{N-1}} & n = 1 \\ \binom{|X|}{n}\left(\frac{n}{|X|}\right)^N - \sum_{k=1}^{n-1}\binom{|X|-k}{n-k}P(|\pi^*| = k) & n > 1. \end{cases}$$

Proof. Follows directly from Theorems 6.3 and 6.7. $\square$

We note that the results of Theorem 6.7 and Corollary 9 are general and applicable to all sets of problems that satisfy the problem generation assumptions listed in Section 6.2. Furthermore, experimental results in Section 6.2.2 indicate that $\mathbb{E}[\pi^*]$ is unlikely to increase if γ were to increase uniformly across segments, suggesting that Corollary 9 also applies to problems with uniform non-negative fixed costs and γ. However, in the absence of a formal proof, the validity of this claim remains an open question.

Table 6-1. Respective u_{ijs} values.

	Segment 1			Segment 2			Segment 3		
Product	$S=1$	$S=2$	$S=3$	$S=1$	$S=2$	$S=3$	$S=1$	$S=2$	$S=3$
1	5.0	5.0	4.0	5.0	4.0	4.0	5.0	4.0	4.0
2	5.0	4.0	4.0	5.0	5.0	4.0	4.0	4.0	5.0
3	4.0	4.0	5.0	4.0	4.0	5.0	5.0	5.0	4.0
4	7.1	0.0	7.1	7.1	0.0	7.1	0.0	12.0	0.0
5	0.0	12.0	0.0	0.0	12.0	0.0	7.1	0.0	7.1

Table 6-2. Estimated expected optimal number of products to offer, $\hat{\mathbb{E}}[|\pi^*|]$.

		$N=10$			$N=20$			$N=30$		
γ	$\lvert S\rvert$	$\lvert X\rvert=10$	$\lvert X\rvert=20$	$\lvert X\rvert=30$	$\lvert X\rvert=10$	$\lvert X\rvert=20$	$\lvert X\rvert=30$	$\lvert X\rvert=10$	$\lvert X\rvert=20$	$\lvert X\rvert=30$
0.1	10	6.35	7.40	8.45	7.70	11.60	13.60	8.65	14.60	16.96
0.1	20	5.70	6.80	7.65	7.75	10.90	13.30	8.35	13.80	15.88
0.1	30	5.05	6.20	7.25	6.55	10.80	12.65	7.95	12.35	12.92
0.3	10	4.50	6.00	7.25	5.60	9.60	12.05	6.90	12.00	16.20
0.3	20	2.60	4.40	5.95	3.25	6.35	9.10	4.10	9.25	12.65
0.3	30	1.65	2.95	2.65	1.80	2.05	3.95	1.25	2.55	3.65
0.5	10	2.55	5.00	5.95	2.90	6.60	9.85	4.10	9.65	13.35
0.5	20	1.15	1.45	2.10	1.00	1.05	1.10	1.00	1.00	1.00
0.5	30	1.05	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	UB	6.51	8.03	8.63	8.78	12.83	14.77	9.58	15.71	19.15

CHAPTER 7
SUMMARY AND FUTURE RESEARCH

7.1 Sensitivity Analysis

Dynamic programming (DP) is a versatile optimization methodology that can be used to solve a variety of problems. Conventional solution approaches assume that the model parameters are known. However, the parameters are often uncertain in practice and hence the stability of the solution is of interest to the decision maker. The typical approach is to solve the problem for different realizations of parameter values. However, there are both theoretical and practical reasons for developing more efficient sensitivity analysis approaches for dynamic programs. Sensitivity analysis is a well-established topic in LP and can be found in most LP textbooks (for example, Bazaraa et al. (2005)). However, this topic is rarely mentioned, much less discussed, in most DP textbooks (see, White (1993), Puterman (1994), Bather (2000) and Powell (2007)).

In Chapter 2, we highlight the relationship between DP and LP and demonstrate how approaches and results from the LP literature can be applied to discrete dynamic programs. However, dynamic programs are rarely solved as linear programs in practice and there is a need to consider sensitivity analysis approaches that exploit the structure of DP problems (for example, Bellman equations).

In Chapter 3, we examine how sensitivity analysis can be performed directly for a MDP with uncertain rewards. For the single parameter problem, we illustrate how the optimal region of a policy can be obtained by considering the region in which the current policy is optimal with respect to each possible action (Proposition 3.1). When the uncertain parameters are allowed to vary simultaneously, we compute the maximum allowable error in the estimated values such that the current solution remains optimal (Proposition 3.2) and illustrate how the maximum allowable tolerance can be computed when the uncertain parameters are non-stationary (Proposition 3.3) by showing that it is sufficient to consider a subset of possible estimation errors (Theorem 3.2). In addition, we highlight that the maximum allowable tolerance of the stationary problem is at least as great as that of the non-stationary problem (Theorem 3.3) and derive the conditions where the tolerances of the stationary and non-stationary problems are the same (Corollary 2) and the conditions where they differ (Theorem 3.4).

This work is motivated by the fact that rewards are often estimated and uncertain in practice. We illustrate the applicability of this work through a capacitated stochastic lot-sizing problem where the ordering costs and backlog penalties are uncertain. Other sequential problems that involve uncertain rewards include equipment replacement (for example, uncertain salvage value), medical decision making (for example, value of a human life) and dynamic assignment (for example, value of a task).

The sensitivity analysis approaches proposed in Chapter 3 assume that rewards can be expressed as affine functions of uncertain parameters. One extension is to consider rewards involving more general functions. We also highlight the conditions where stationary uncertain parameter assumptions lead to overly optimistic tolerance levels for a general lot-sizing problem under mild assumptions (Theorem 3.5). Another area of further research is to identify conditions where this is true for other sequential decision problems.

The monotonicity and continuity of V with respect to the model parameters are discussed in Section 2.2. In particular, we highlight that V is not necessarily monotone in T, but it is possible to derive sufficient conditions that guarantee monotonicity. A potential area of research is to explore if more of these sufficient conditions can be derived for specific dynamic programming problems (i.e., lot-sizing problem, knapsack problem, etc.). Another interesting area of research is to identify additional properties of V and π^* with respect to the model parameters. For example, White and El-Deib (1986) highlighted that V is piecewise affine and concave in $\boldsymbol{\rho}$ if r^a and $V_T(s)$ are affine in $\boldsymbol{\rho}$.

A number of problems that are encountered in practice are large and it is not possible to compute the exact solutions for these dynamic programs. These problems are often solved by approximating the value function and/or aggregating the state space. Topaloglu and Powell (2007) were able to exploit the structure of their value function approximation to efficiently revise the solution when the model parameters change. It would be interesting to explore how sensitivity analysis can be performed for different approximate dynamic programs. We also note that researchers have typically explored problems where the parameter uncertainties are restricted to one group of parameters, such as uncertainties solely in the rewards or transition probabilities. However, parameter uncertainty is rarely confined to a single class of parameters in practice. Another promising area of research will be to explore problems where different types of parameters are allowed to vary simultaneously.

7.2 Regret

Research in psychology suggests that the satisfaction resulting from a decision under uncertainty is often influenced by feelings of regret. However, the definition of regret in normative decision theory is limited. In Chapter 4, we review regret theory and illustrate how regret-theoretic choice preferences are described by multivariate stochastic dominance. In addition, we present regret-based risk measures and illustrate how they can be adopted within the mean-risk framework.

In Chapter 5, we examine an equipment replacement problem under a finite horizon, but the actual horizon length is uncertain. In our experience, this occurs often in practice, from the uncertainty of when a production line is to be shut down to the closing of military bases. Stochastic dynamic programming formulations are presented in order to minimize expected costs and maximum regret. As the replacement policies for these objectives are often different, we design an option contract to lease the asset after period T_s, the earliest horizon time, in order to mitigate the risk of loss due to the uncertain horizon. The contract is made available for a given price at t_c, the time period when optimal decisions for the different objectives (minimizing expected costs and minimizing maximum regret) first diverge.

The optimal contract parameters (option price and periodic lease rates) are dependent on a number of factors, including the probable length of the lease, discount rate, and the leasor's expected profit margin. The option price, which must be incurred by the decision maker whether he or she exercises the lease or not and represents significant income to the leasor, increases with the discount rate and lease length. We provide the optimal parameters in terms of the costs and expected profit of the leasor (Theorem 5.2 and Corollary 5).

This work illustrates the need to look at equipment replacement problems with contractual obligations more closely. We examine the parameters for a leasing option contract to be pursued in lieu of ownership. There are numerous situations with regards to asset ownership that require sophisticated contracts. For example, many owners, such as airlines, outsource the service of their assets to third parties or the original equipment manufacturer. Additionally, equipment sellers are now under greater scrutiny to generate stated benefits under contract or face penalties. That is, equipment must meet the expectations of advertised technological improvements. These situations open tremendous research opportunities into contract design in light of equipment replacement decisions.

In Chapter 6, we investigate the relationship between anticipated regret and choices. Individuals experience regret in a variety of settings. However, choices can be independent of anticipated regret, even when satisfaction is influenced by feelings of regret. In practice, individuals may make computational mistakes when identifying the optimal choice. Hence, an objective that is easy to compute (i.e., independent of regret) is preferred over one that is more complicated (i.e., influenced by regret). Theorem 6.1 states that choice preference is independent of anticipated regret when the decision maker maximizes expected net satisfaction defined by Equation (6–2). It is not hard to see that this result extends for all definitions of net satisfaction that are linear in regret.

In addition, we obtain theoretical and experimental results on the optimal number of products to offer under various conditions in a heterogeneous market. When consumer preferences vary greatly, a company that wishes to capture a broad market share needs to provide a rich product line. However, this is suboptimal when the cost of allowing for a wide selection is high (Theorem 6.3). Therefore, a company seeking a broad market share should focus on reducing the cost of introducing variety into their product line. Dell is an excellent example of how a firm can support a wide consumer base when it is able to offer a wide variety at competitive prices. However, the product line should be narrow when outcomes are uncertain and consumers are affected by regret (Section 6.2.2). This is because a wide array of products increases the likelihood and magnitude of regret experienced. For example, financial advisors and insurance agents generally do not present a multitude of products to their customers.

One interesting observation is that the optimal number of products to offer can increase when regret aversion increases (Example 10). This happens when a small set of risky products results in high expected market satisfaction when regret aversion is low. When regret aversion increases, expected market satisfaction decreases as a result of increased regret and it may be better to offer a larger set of less risky products. For example, although people often regret food choices, many fast-food restaurants offer a variety of products in their menus. One observation is that the choices offered by fast-food restaurants are usually relatively safe in that the regret of making a wrong choice (for example, purchasing a hamburger when a chicken sandwich results in the greatest satisfaction) is generally low.

Our research also suggests that consumers are likely to be indifferent between choices when the associated outcomes are extremely variable (see, proof of Theorem 6.5). When the outcome of choices are highly uncertain, it is almost surely optimal for the firm to offer a single product (Theorem 6.5). This prediction is consistent with studies that highlight that patients are often uneasy and undecided when given a wide selection of treatments to choose from, relying strongly on medical professionals to decide the treatment plan for them when the associated outcomes are highly uncertain (Schneider 1998).

In our optimal product line analysis, we use a choice model that is based on absolute regret. One extension is to consider other notions of regret discussed in Section 4.2. In addition, recent studies suggest that individuals can be disproportionately adverse to large regret (Bleichrodt et al. 2010). Another extension of the problem is to consider consumer satisfaction that is non-linear in regret. Third, consumers may not always be rational and select choices that maximize expected net satisfaction (Simon et al. 1995). Another interesting extension is to consider product line design with regret considerations under bounded rationality.

REFERENCES

[1] Ahuja, R.K., Magnanti, T.L. and Orlin, J.B. (1993). *Network Flows*, Prentice Hall, NJ.

[2] Aissia, H., Bazgana, C. and Vanderpooten, D. (2009). Minmax and minmax regret versions of combinatorial optimization problems: A survey. *European Journal of Operational Research*, **197**, 427–438.

[3] Altman, E. and Shwartz, A. (1991). Sensitivity of constrained Markov decision processes. *Annals of Operations Research*, **32**, 1–22.

[4] Alptekinoglu, A. and Corbett, C.J. (2008). Mass customization versus mass production: Variety and price competition. *Production and Operations Management* , **17**, 587–598.

[5] Averbakh, I. and Lebedev, V. (2004). Interval data minmax regret network optimization problems. *Discrete Applied Mathematics*, **138**, 289–301.

[6] Aydin, G. and Porteus, E.L. (2008). Joint inventory and pricing decisions for an assortment. *Operations Research*, **56**, 1247–1255.

[7] Bazaraa, M.S., Jarvis, J.J. and Sherali, H.D. (2005). *Linear Programming and Network Flows*, John Wiley & Sons, NJ.

[8] Bean, J.C., Hopp, W.J. and Duenyas, I. (1992). A stopping rule for forecast horizons in nonhomogeneous Markov decision processes. *Operations Research*, **40**, 1188–1199.

[9] Bean, J.C., Lohmann, J.R. and Smith, R.L. (1985). A dynamic infinite horizon replacement economy decision model. *The Engineering Economist*, **30**, 99–120.

[10] Bean, J.C., Lohmann, J.R. and Smith, R.L. (1994). Equipment replacement under technological change. *Naval Research Logistics*, **41**, 117–128.

[11] Bean, J.C. and Smith, R.L. (1984) Conditions for the existence of planning horizons. *Mathematics of Operations Research*, **9**, 391–401.

[12] Bell, D. E. (1987). Regret in decision making under uncertainty. *Operations Research*, **30**, 961–981.

[13] Bellman, R. (1952). The theory of dynamic programming. *Proceedings of the National Academy of Sciences*, **38**, 716-719.

[14] Bellman, R.E. (1955). Equipment replacement policy. *Journal of the Society for the Industrial Applications of Mathematics*, **3**, 133–136.

[15] Berry, W.L. and Cooper, M.C. (1999). Manufacturing flexibility: Methods for measuring the impact of product variety on performance in process industries. *Journal of Operations Management*, **17**, 163–178.

[16] Bertsimas, D., Brown, D.B. and Caramanis C. (2011). Theory and applications of robust optimization. *SIAM Review*, **53**, 464–501.

[17] Beyer, H. and Sendhoff, B. (2007). Robust optimization - A comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, **196**, 3190–3218.

[18] Birnbaum, M.H. and Navarrete, J.B. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty*, **17**, 49–79.

[19] Bleichrodt, H., Cillo, A. and Diecidue, E. (2010). A quantitative measurement of regret theory. *Management Science*, **56**, 161–175.

[20] Bradley, S.P., Hax, A.C. and Magnanti, T.L. (1977) *Applied Mathematical Programming*, Addison-Wesley.

[21] Braun, M. and Muermann, A. (2004). The impact of regret on the demand for insurance. *The Journal of Risk and Insurance*, **71**, 737–767.

[22] Brenner, L., Rottenstreich, Y. and Sood, S. (1999). Comparison, grouping, and preference. *Psychological Science*, **10**, 225–229.

[23] Briggs, A., Sculpher, M. and Buxton, M. (1994) Uncertainty in the economic evaluation of health care technologies: The role of sensitivity analysis. *Health Economics*, **3**, 95–104.

[24] Cachon, G.P., Terwiesch, C. and Xu, Y. (2005). Retail assortment planning in the presence of consumer search. *Manufacturing & Service Operations Management* **7** 330–346.

[25] Chand, S. and Sethi, S. (1982). Planning horizon procedures for machine replacement models with several possible replacement alternatives. *Naval Research Logistics Quarterly*, **29**, 483–493.

[26] Charalambous, C. and Gittins, J.C. (2008). Optimal selection policies for a sequence of candidate drugs. *Advances in Applied Probability*, **40**, 359–376.

[27] Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **158**, 419–466.

[28] Cheevaprawatdomrong, T. and Smith, R.L. (2003) A paradox in equipment replacement under technological improvement. *Operations Research Letters*, **31**, 77–82.

[29] Chen, F., Eliashberg, J. and Zipkin, P. (1998). Customer preferences, supply-chain costs, and product-line design. In T. H. Ho and C. S. Tang (Eds.), *Product Variety Management: Research Advances*, Kluwer Academic Publishers, Norwell, MA.

[30] Chernev, A. (2003). When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *Journal of Consumer Research*, **30**, 170–183.

[31] Connolly, T. and Zeelenberg, M. (2002). Regret in decision making. *Current Directions in Psychological Science*, **11**, 212–216.

[32] Cornuejols, G., Nemhauser, G.L. and Wolsey, L.A. (1990). The uncapacitated facility location problem. In P. B. Mirchandani and R. L. Francis (Eds.), *Discrete Location Theory*, Wiley-Interscience, New York.

[33] de Groote, X. (1994). Flexibility and marketing/manufacturing coordination. *International Journal of Production Economics*, **36**, 153–167.

[34] de Sousa, J.F. and Guimaraes, R.C. (1997). Setting the length of the planning horizon in the vehicle replacement problem. *European Journal of Operational Research*, **101**, 550–559.

[35] Derman, C. (1962). On sequential decisions and Markov chains. *Management Science*, **9**, 16–24.

[36] Desai, P., Kekre, S., Radhakrishnan, S. and Srinivasan, K. (2001). Product differentiation and commonality in design: Balancing revenue and cost drivers. *Management Science*, **47**, 37–51.

[37] Dobson, G. and Kalish, S. (1993). Heuristics for pricing and positioning a product-line using conjoint and cost data. *Management Science*, **39**, 160–175.

[38] Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57**, 45–97.

[39] Elton, E.J. and Gruber, M.J. (1971) Dynamic programming applications in finance. *The Journal of Finance*, **26**, 473–506.

[40] Engelbrecht-Wiggans, R. and Katok, E. (2008). Regret and feedback information in first-price sealed-bid auctions. *Management Science*, **54**, 808–819.

[41] Engelbrecht-Wiggans, R. and Katok, E. (2009). A direct test of risk aversion and regret in first price sealed-bid auctions. *Decision Analysis*, **6**, 75–86.

[42] Eppstein, D. (2003). Setting parameters by example. *SIAM Journal on Computing* , **32**, 643–653.

[43] Erkin, Z., Bailey, M.D., Maillart, L.M., Schaefer, A.J. and Roberts, M.S. (2010). Eliciting patients' revealed preferences: An inverse Markov decision process approach. *Decision Analysis*, **7**, 358–365.

[44] Fernandez-Baca D. and Venkatachalam, B. (2007) Sensitivity analysis in combinatorial optimization. *In Handbook of Approximation Algorithms and Metaheuristics*, **18**, 673–696. Chapman & Hall/CRC

[45] Filippi, C. (2005). A fresh view on the tolerance approach to sensitivity analysis in linear programming. *European Journal of Operational Research*, **167**, 1–19.

[46] Fraser, J.M. and Posey, J.W. (1989). A framework for replacement decisions. *European Journal of Operational Research*, **40**, 43–57.

[47] Gal, T. and Greenberg, H.J. (1997). *Advances in Sensitivity Analysis and Parametric Programming*, Kluwer Academic, Boston, MA.

[48] Gaur, V. and Honhon, D. (2006). Assortment planning and inventory decisions under a locational choice model. *Management Science*, **52**, 1528–1543.

[49] Glazebrook, K.D., Ansell, P.S., Dunn, R.T. and Lumley R.R. (2004). Eliciting patients' revealed preferences: An inverse Markov decision process approach. *Journal of Applied Probability*, **41**, 51–72.

[50] Given, R., Leach, S. and Dean, T. (2000). Bounded-parameter Markov decision processes. *Artificial Intelligence*, **122**, 71–109.

[51] Goldman, A.J. and Tucker A.W. (1956). Theory of linear programming. In Kuhn H.W. and Tucker A.W., editors, *Linear Inequalities and Related Systems*, 53–97. Princeton, NJ.

[52] Hansen, S., Labbe, M. and Wendell, R.E. (1989). Sensitivity analysis in multiple objective linear programming: The tolerance approach. *European Journal of Operational Research*, **38**, 63–69.

[53] Harmanec, D. (2002). Generalizing Markov decision processes to imprecise probabilities. *Journal of Statistical Planning and Inference*, **105**, 199–213.

[54] Hartman, J.C. and Lohmann, J. R. (1997). Multiple options in parallel replacement analysis: Buy, lease or rebuild. *The Engineering Economist*, **42**, 223–248.

[55] Hartman, J.C. and Murphy, A. (2006). Finite-horizon equipment replacement analysis. *IIE Transactions*, **38**, 409–419.

[56] Held, M. and Karp, R.M. (1962). A dynamic programming approach to sequencing problems . *Journal of the Society for Industrial and Applied Mathematics*, **10**, 196-210.

[57] Honhon, D., Gaur, V. and Seshadri, S. (2010). Assortment planning and inventory decisions under stockout-based substitution. *Operations Research*, **58**, 1364–1379.

[58] Hopp, W.J. (1988). Sensitivity analysis in discrete dynamic programming. *Journal of Optimization Theory and Applications*, **56**, 257–269.

[59] Hopp, W.J. (1989). Identifying forecast horizons in nonhomogeneous Markov decision processes. *Operations Research*, **37**, 339–343.

[60] Inuiguchi, M. and Sakawa, M. (1995). Minimax regret solution to linear programming problems with an interval objective function. *European Journal of Operational Research*, **86**, 526–536.

[61] Irons, B. and Hepburn, C. (2007). Regret theory and the tyranny of choice. *Economic Record*, **83**, 191–203.

[62] Iyengar, G.N. (2005). Robust dynamic programming. *Mathematics of Operations Research*, **30**, 257–280.

[63] Iyengar, S.S. and Lepper, M.R. (2000). When choice is demotivating: Can one desire too much of a good thing?. *Journal of Personality and Social Psychology*, **79**, 995–1006.

[64] Kalyanasundaram, S., Chong, E. and Shroff, N. (2002). Markov decision processes with uncertain transition rates: Sensitivity and robust control. *Proceedings of the 41st IEEE Conference on Decision and Control*, **4**, 3799–3804.

[65] Kekre, S. and Srinivasa, K. (1990). Broader product line: A necessity to achieve success? *Management Science*, **36**, 1216–1232.

[66] Kim, E. and Kececioglu, J. (2007). Inverse sequence alignment from partial examples. *Algorithms in Bioinformatics*, **4645**, 359–370.

[67] Kok, A.G. and Fisher, M.L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, **55**, 1001–1021.

[68] Kok, A.G., Fisher, M. and Vaidyanathan, R. 2008. Assortment planning: Review of literature and industry practice. In N. Agrawal and S.A. Smith (Eds.), *Retail Supply Chain Management*, Springer.

[69] Kotler, P. and Keller, K. (2006). *Marketing Management (12th edn)*, Pearson Education, Upper Saddle River, NJ.

[70] Kouvelis, P. and Yu, G. (1997). *Robust discrete optimization and its applications*, Kluwer Academic Publishers, Boston, MA.

[71] Krokhmal, P., Zabarankin, M. and Uryasev, S. (2011). Modeling and optimization of risk. *Surveys in Operations Research and Management Science*, **49**, 49–66 .

[72] Kuksov, D. and Villas-Boas, J.M. (2010). When more alternatives lead to less choice *Marketing Science*, **29**, 507–524.

[73] Leland, J. (2010). The hunt for a descriptive theory of choice under risk–A view from the road not taken. *Journal of Socio-Economics*, **39**, 568–577.

[74] Levy, H. 2006. *Stochastic Dominance*, Springer, New York.

[75] Levy, H. and Paroush, G. (1992). Multi-period stochastic dominance. *Management Science*, **21**, 428–435.

[76] Lim, C., Bearden, J.N. and Smith, J.C. (2006). Sequential search with multiattribute options. *Decision Analysis*, **3**, 3–15.

[77] Loewenstein, G.F. and Schkade, D. (1999). Wouldnt it be nice? Predicting future feelings. In E. Diener, N. Schwartz and D. Kahneman (Eds.), *Well-being: The Foundations of Hedonic Psychology*, New York: Russell Sage Foundation.

[78] Loomes, G. and Sugden, R. (1982). Regret theory: an alternative theory of rational choice under uncertainty. *The Economic Journal*, **92**, 805–824.

[79] Loomes, G. and Sugden, R. (1987). Some implications of a more general form of regret theory. *Journal of Economic Theory*, **41**, 270–287.

[80] Mahajan, S. and van Ryzin, G. (2001). Stocking retail assortments under dynamic consumer substitution. *Operations Research*, **49**, 334–351.

[81] Manne, A. S. (1960). Linear programming and sequential decisions. *Management Science*, **6**, 259–267.

[82] McKay, M.D. (1979). Sensitivity analysis. *In Proceedings of the Workshop on Validation of Computer-based Mathematical Models in Energy Related Research and Development*, Fort Worth, TX.

[83] Mitrophanov A.Y., Lomsadze A., Borodovsky M. (2005). Sensitivity of hidden Markov models. *Journal of Applied Probability*, **42**, 632–642.

[84] Muckstadt, J.A. and Sapra A. (2010). *Principles of Inventory Management*, Springer, NY.

[85] Muller, A. (1997). How does the value function of a Markov decision process depend on the transition probabilities? *Mathematics of Operations Research*, **22**, 872–885.

[86] Nasiry, J. and Popescu, I. (2011). Advance selling when consumers regret. *INSEAD Working Paper No, 2011/03/DS*.

[87] Netessine, S. and Taylor, T.A. (2007). Product line design and production technology. *Marketing Science*, **26**, 101–117.

[88] Nilim, A. and El Ghaoui, L. (2005). Sensitivity of hidden Markov models. *Journal of Applied Probability*, **42**, 632–642.

[89] Oakford, R.V., Lohmann, J.R. and Salazar, A. (1984). A dynamic replacement economy decision model. *IIE Transactions*, **16**, 65–72.

[90] Park, C.S. and Sharp-Bette, G.P. (1990). *Advanced Engineering Economics*, John Wiley and Sons, New York.

[91] Pliska, S.R. and Ye, J. (2007). Optimal life insurance purchase and consumption/investment under uncertain lifetime. *Journal of Banking and Finance*, **31**, 1307–1319.

[92] Powell, W.B. (2007). *Approximate Dynamic Programming*, John Wiley & Sons, New Jersey.

[93] Pruzan, P.M. and Jackson, J.T.R. (1967). A dynamic programming application in production line inspection. *Technometrics*, **9**, 73–81.

[94] Puterman, M.L. (1994). *Markov Decision Processes*. John Wiley & Sons, New York.

[95] Puumalainenab, J. (1998). Optimal cross-cutting and sensitivity analysis for various log dimension constraints by using dynamic programming approach. *Scandinavian Journal of Forest Research*, **13**, 74–82.

[96] Quiggin, J. (1990). Stochastic dominance in regret theory. *Review of Economic Studies*, **57**, 503–511.

[97] Quiggin, J. (1994). Regret theory with general choice sets. *Journal of Risk and Uncertainty*, **8**, 153–165.

[98] Randall, T. and K. Ulrich. 2001. Product variety, supply chain structure, and firm performance: Analysis of the U.S. bicycle industry. *Management Science* **47** 1588–1604.

[99] Ravi, N. and Wendell, R.E. (1989). The tolerance approach to sensitivity analysis of matrix coefficients in linear programming. *Management Science*, **35**, 1106–1119.

[100] Redelmeier, D. A. and E. Shafir (1995). Medical decision making in situations that offer multiple alternatives. *Journal of the American Medical Association* **273** 302–205.

[101] Regnier, E., Sharp, G. and Tovey, C. (2004). Replacement under ongoing technological change. *IIE Transactions*, **36**, 497–508.

[102] Safizadeh, M.H., Ritzman, L.P., Sharma, D. and Wood, C. (1996). An empirical analysis of the product-process matrix. *Management Science*, **42**, 1576–1591.

[103] Sandvik, B. and Thorlund-Petersen, L. (2010). Sensitivity analysis of risk tolerance. *Decision Analysis*, **7**, 313–321.

[104] Satia, J.K. and Lave, R.E. (1973). Markovian decision processes with uncertain transition probabilities. *Operations Research*, **21**, 728–740.

[105] Savage, L.J. (1951). The theory of statistical decision. *American Statistical Association Journal*, **46**, 55–67.

[106] Scheibehenne, B. (2008). *The Effect of Having Too Much Choice*, Unpublished Dissertation, Humboldt University, Berlin, Germany.

[107] Schneider, C.E. (1998). *The Practice of Autonomy: Patients, Doctors, and Medical Decisions*, Oxford University Press, New York.

[108] Schwartz, B. (2004). *The Paradox of Choice: Why Less Is More*, Ecco, New York.

[109] Shafira, E., Simonsonb, I. and Tversky, A. (1993). Reason-based choice. *Cognition*, **49**, 11–36.

[110] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*, Springer, New York.

[111] Sivazlian, B. D. (1977). Life cycle cost analysis: Under uncertainty in systems lifetime. *The Engineering Economist*, **25**, 91–105.

[112] Smith, S.A. and Agrawal, N. (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research*, **48**, 50–64.

[113] Sotskov, Y.N., Leontev, V.K. and Gordeev E.N. (1995). Some concepts of stability analysis in combinatorial optimization. *Discrete Applied Mathematics*, **58**, 169–190.

[114] Sugden, R. (1993). An axiomatic foundation for regret theory. *Journal of Economic Theory*, **60**, 159–180.

[115] Sun, F., Fernandez-Baca, D. and Yu, W. (2004). Inverse parametric sequence alignment. *Journal of Algorithms*, **58**, 36–54.

[116] Syam, N., Krishnamurthy, P. and Hess, J.D. (2008). That's what I thought I wanted? Miswanting and regret for a standard Good in a mass-customized world. *Marketing Science*, **27**, 379–397.

[117] Tan, C. and Hartman, J.C. (2010). Equipment replacement analysis with an uncertain finite horizon. *IIE Transcations*, **42**, 342–353.

[118] Tan, C. and Hartman, J.C. (2011). Sensitivity analysis and dynamic programming. In Cochran, J. J. (Ed.), *Wiley Encyclopedia of Operations Research and Management Science*, John Wiley & Sons, NJ.

[119] Tilahun, K. and Raes, D. (2002). Sensitivity analysis of optimal irrigation scheduling using a dynamic programming model. *Australian Journal of Agricultural Research*, **53**, 339–346.

[120] Topaloglu, H. and Powell, W.B. (2007). Sensitivity analysis of a dynamic fleet management model using approximate dynamic programming. *Operations Research*, **55**, 319–331.

[121] Trietsch, D. (1985). Optimal motion towards a stochastic destination. *Discussion Paper No. 654, Northwestern University, Center for Mathematical Studies in Economics and Management Science, May 1985*.

[122] Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, **79**, 281–299.

[123] Tversky, A. and Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological Science*, **3**, 358–361.

[124] Veinott, A.F. and Wagner, H.M. (1965). Computing optimal (s, S) inventory policies. *Management Science*, **11**, 525–552.

[125] Van de Ven, N. and Zeelenberg, M. (2011). Regret aversion and the reluctance to exchange lottery tickets. *Journal of Economic Psychology*, **32**, 194–200.

[126] van Ryzin, G. and Mahajan, S. (1999). On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, **45**, 1496–1509.

[127] Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

[128] Wagner, H.M. (1975). *Principles of Operations Research*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

[129] Wagner, H.M. (1995). Global Sensitivity Analysis. *Operations Research*, **43**, 948–969.

[130] Walker D. and Fox-Rushby J.A. (2001). Allowing for uncertainty in economic evaluations: Qualitative sensitivity analysis. *Health Policy and Planning*, **16**, 435–443.

[131] Wallace, S.W. (2000). Decision making under uncertainty: Is sensitivity analysis of any use? *Operations Research*, **48**, 20–25.

[132] Ward, J.E. and Wendell, R.E. (1990). Approaches to sensitivity analysis in linear programming. *Annals of Operations Research*, **27**, 3–38.

[133] Wendell, R.E. (1985). The tolerance approach to sensitivity analysis in linear programming. *Management Science*, **31**, 564–578.

[134] Wendell, R.E. (2004). Tolerance sensitivity and optimality bounds in linear programming. *Management Science*, **50**, 797–803.

[135] White, C.C. and El-Deib, H.K. (1986). Parameter imprecision in finite state, finite action dynamic programs. *Operations Research*, **34**, 120–129.

[136] White, C.C. and Eldeib, H.K. (1994). Markov decision processes with imprecise transition probabilities. *Operations Research*, **42**, 739–749.

[137] White, D.J. (1993). *Markov Decision Processes*, John Wiley & Sons, UK.

[138] Williams A.C. (1970). Complementary theorems for linear programming. *SIAM Review*, **12**, 135–137.

[139] Wondolowski, F.R. (1991). Generalization of Wendell's tolerance approach to sensitivity analysis in linear programming. *Decision Sciences*, **22**, 792–811.

[140] Yakowitz, S. (1982). Dynamic programming applications in water resources. *Water Resources Research*, **18**, 673–696.

[141] Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, **12**, 93–106.

BIOGRAPHICAL SKETCH

Chin Hon Tan is a Ph.D. candidate in the Department of Industrial and Systems Engineering at the University of Florida. He received his B.Eng. (2005) and M.S. (2009) in Industrial and Systems Engineering from the National University of Singapore and the University of Florida, respectively. Prior to his graduate studies, he worked as a reliability engineer with Shell Eastern Petroleum and also taught applied mathematics at Republic Polytechnic in Singapore. His research interests include decision analysis, stochastic modeling and dynamic programming. He is a member of IIE and INFORMS.