Simultaneous Improvements in Manufacturing Systems and Effects on Investment Decisions

by

Timothy Joseph Nye

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 1997

0-612-22224-1

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

Manufacturing managers have the option of improving their operations through practices such as setup time reduction or quality improvement. Making improvements requires the investment of resources. For the purpose of determining the level of investments in each practice, projects are normally assumed to be independent, and are justified on the basis of the expected returns from the improvements. This research investigates the effects of this implicit assumption of independence between improvement practices on optimal investment decisions. To answer the question of whether improvement practices interact with each other, and if so, whether the effect of the interaction is significant to the investment decision, a model of a manufacturing system was developed. In the model, the independent variables are the levels of investment in each of two improvement practices and the performance measure of the system is relevant expected operating costs. For the purposes of this study, the two improvement practices implemented in the model are setup time reduction and quality improvement.

The model developed here draws primarily upon two previous models. The first, by Porteus [1985], adapted an Economic Order Quantity model to include investments in setup reduction, and was later extended [Porteus, 1986] to include investments in quality improvement. In these models, when the EOQ for a system was calculated, total costs were minimized as a function of order quantity, investment in setup reduction and investment in quality improvement. A limitation of Porteus' EOQ-based model is that it neglected WIP holding costs, which can be substantial in manufacturing systems. Karmarkar [1987] developed a model based on the M/M/1 queuing system which predicted WIP levels in a system, and with this model was able to calculate an order quantity which minimizes total costs of the system, although no work has been found on reducing setup times or improving quality with this type of model. In this research an M/G/1 queuing model was used to represent a manufacturing cell and estimate WIP levels. Stochastic service times include setup time and time for rework of defective units in each

batch processed, with these quantities being the independent variables of the model. By linking the levels of the independent variables to levels of investment necessary to achieve those values, a total expected relevant cost for the system can be estimated, and used as an objective function to optimize setup times and defect rates.

This model has been used to determine optimum investment strategies for cases where batch size is fixed or variable, and where the investment-improvement function for each decision variable is linear or strictly convex. Analytic and numerical results have been obtained.

By comparing optimal levels of the decision variables when each practice is explicitly assumed independent of the other to when the optimization is performed simultaneously, the question of whether interactions between practices can be answered. The research question has been answered in the affirmative: each case of this model shows that interactions between the improvement practices exist, and if ignored, these interactions can lead to significant levels of over-investment. The most significant factor in determining the potential levels of over-investment has been found to be the form of the investment-improvement function, which is also one of the empirically least understood elements of this model.

Acknowledgments

The author would like to express his sincere gratitude to his supervisors in this work, Dr. Elizabeth Jewkes and Dr. David Dilts. Without their advice, support, patience and perseverance, it is likely this work would not have been completed.

The author would also like to thank the other committee members, namely, Dr. Ruth Harris, Dr. Jan Huissoon, Dr. Ray Vickson and the external member, Dr. Jack Hayya for their interest and efforts. The quality of this research is a direct result of the substantial level of feedback and constructive support provided by the committee.

Of course, it also goes without saying that we all owe a special debt of gratitude to Mrs. Marion Reid, Graduate Secretary in Management Sciences, without whom work in this department would have been a far less pleasant experience.

Finally, the author wishes to acknowledge the Natural Sciences and Engineering Research Council, who funded the first half of the author's studies under an Employed Scientist and Engineer Scholarship.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Overview

This research is concerned with understanding optimal investment decisions when making simultaneous improvements, specifically in setup reduction and quality improvement, on a manufacturing system. It will be shown that these improvement practices can interact in a manufacturing system, and these interactions can lead to incorrect investment decisions if they are not taken into account. The consequence of these incorrect investment decisions is that more resources would be invested than are warranted, with potentially serious levels of over-investment occurring. Gaining insight into how simultaneous improvements might interact, the impact of such interaction on system performance, and the resulting effects on investment decisions is the primary focus of this research. This study models setup time reduction and quality improvement in the context of an existing, repetitive lot-based discrete manufacturing system.

To illustrate this idea, consider a manufacturing system being studied for potential improvements. A manufacturing engineer might propose a setup time reduction project and develop an investment proposal for the most appropriate level of investment in this particular system. Meanwhile, a quality engineer examines the same system and proposes an appropriate investment in a quality improvement project. Both investment proposals go to a decision maker who evaluates them against the firm's investment criteria as separate, independent projects. Suppose that both proposals are accepted and are implemented sequentially. The first project implemented starts with the manufacturing system in its initial state and significantly improves the system. The expected level of benefits, such as WIP reduction, should be obtained.

The second project is then implemented on a system much different than that assumed in the investment proposal. Since there are, for instance, less total WIP costs in the already improved system, there is less potential for savings and the expected level of benefits for the second project will not likely be obtained. Thus, because of the implicit assumption of independence between the two investment proposals, more investment may be made than was justified.

WIP reduction is mentioned here because it has been identified as one of the most significant cost savings attributable to improvement practices such as setup reduction and quality improvement [Primrose, 1992]. As well, Boucher [1984] has stated that "observation of large scale discrete parts production facilities will show that the largest component of inventory cost is work-in-process inventory". Similarly, analytic modeling by Karmarkar [1987] produced example results showing a holding cost of WIP inventory of eight times the cost of holding finished goods inventory. Since WIP is a function of server utilization, which itself is a function of setup and quality levels, investing in setup reduction and quality improvement would be expected to reduce server utilization and, ultimately, WIP level. Due to the relative magnitude of the cost of holding WIP, it should be included in any model for investment in setup and quality improvements, as is done in this research.

The previous example shows how erroneous decisions can occur when investing in multiple, simultaneous manufacturing improvement practices without taking into account how the improvements interact. The thrust of the remainder of this dissertation is to develop a model of a manufacturing system for use in studying costs and benefits of investing in manufacturing improvements. This model is then used to examine the sources and magnitudes of potential errors in the decision process resulting from neglecting interactive effects.

1.2 Manufacturing Improvement Practices

There have been a number of philosophies introduced in North America over the last decade or so which promise to improve the competitiveness of manufacturers implementing them. Some of the more common philosophies might be listed as:

- Lean Production [Womack et al., 1990]
- Toyota Production System [Shingo, 1992]
- Just In Time (JIT) [Monden, 1993]
- Continuous Improvement [Bakerjian, 1993]

Distinguishing among these philosophies tends to be difficult since they share important attributes. For instance, all require that setup time be minimized, that quality be improved and that material handling be minimized, among other common practices. While these philosophies have been widely publicized, there is evidence that there is significant opportunity left for implementation. For instance, Billesbach [1991] found that as recently as 1990, only 1 to 2% of American manufacturers had effectively implemented Just-In-Time production.

More recently, philosophies such as Fractal Manufacturing [Warnecke, 1993], Agile Manufacturing [Montgomery and Levine, 1996] and Next Generation Manufacturing [Agility Forum, 1997] have come into the fore. Each of these philosophies, however, demands that production operations have been made flexible (i.e., lot sizes effectively reduced to one through elimination of setup times) and predictable (i.e., quality improved to the point where each operation is done right the first time). For example, as Hilton and Gill [1994] point out, "Manufacturing process development is a critical element in achieving agility." Setup reduction and quality improvement still remain at the core of competitive manufacturing philosophies.

The importance of setup reduction and quality improvement has also been demonstrated empirically. In a study of automotive component and electronics manufacturing firms, Fader [1992] showed that manufacturing firms which utilized improvement practices such as setup reduction and quality improvement had significantly higher total factor productivities than firms in the same industries which did not. In turn, Porter [1991] has argued that the productivity of a country's manufacturing sector is a key determinant of the standard of living of a nation. As well, there is a substantial body of anecdotal evidence (for instance, see

[1986]) that these manufacturing improvement practices provide significant competitive benefits at the individual firm level.

The amount of improvement possible, and the economics of that improvement, in a given factory, will depend on the particular circumstances in that facility. However, Shingo [1989, p215] was able to draw the general conclusions that "no transition to a [JIT] system can occur without drastic reductions in setup times" and "when implementing the [JIT] system, therefore, we must challenge ourselves to achieve zero defects." Further, White [1993] in an empirical study of 1035 American manufacturing firms found that setup reduction and quality improvement were the two most frequently used practices.

Because setup reduction and quality improvement have been demonstrated to:

1. provide significant competitive advantage to firms implementing them,
2. are key to most recent manufacturing philosophies,
3. are the most commonly implemented improvement practices, and
4. appear to be not as widely implemented as might be expected.

they have been chosen to be the two illustrative practices used in this study. These practices applied to a manufacturing system provide the benefits of reduced inventory holding costs, reduced production lead times, reduced resource utilization and greater predictability of system performance. By examining two improvement practices together, potential interactions between the practices may be studied.

1.2.1 Setup Reduction

Often in manufacturing, a common machine is used to perform an operation on a number of different products by using adjustments or tooling specialized for each job. As each job reaches the operation, it requires the operation to be shut down for a period of time to perform a setup.

Because the server (usually the machine which performs the operation) is occupied while setup is taking place, it is not available to process jobs. For a given number of jobs, longer setup times means the server is busy a greater proportion of time. This proportion of busy time is known as the utilization of the server, and it can be seen that by reducing setup times, the utilization of the server will decrease.

While traditional manufacturing practice strives to maximize utilization of equipment [Shingo, 1992], higher utilization has a detrimental side effect. From basic queuing theory [Tijms, 1994] it is known that increased utilization of a server results in greater expected numbers of jobs waiting in the queue for the server. These jobs represent WIP inventories, which are expensive to maintain. By reducing setup times, server utilization decreases and WIP levels and costs decrease as well.

Another benefit of reduced setups predicted by queuing theory is that the throughput time of jobs through the system also decreases [Tijms, 1994]. This improves customer service since specially-ordered products can be delivered with shorter lead times.

Improving setup times typically requires a firm to invest resources in training, equipment modifications or layout changes. However, reduction of WIP inventories represents a quantifiable benefit which can be used to help justify the expenditure of those resources. The model developed in this research will be used to predict the costs and benefits of making these improvement investments, and will be used to explore optimal solutions.

1.2.2 Quality Improvement

The other practice of interest in this research is that of quality improvement. If we consider a production system which has imperfect quality, the operation must work longer to compensate for the scrap and rework required to produce a given output level as compared with a system with perfect quality. This extra work causes higher server utilization in the same way as longer setups. By improving quality, server utilization will be expected to drop and result in the same benefits (reduced WIP inventories and faster throughput) as are achieved by reducing setup times. Also as with setup reduction, resources are typically required to improve quality, but these resources can be balanced against the gains from reduced WIP costs.

1.3 Modeling Strategy

To understand the nature of improvements and to measure their impact, two aspects of a system need to be modeled: physical and economic. The physical system is characterized by the need to capture setup times, quality defect rework times and WIP levels, as well as including manufacturing system parameters such as arrival and service times of jobs. A schematic diagram of the model is shown in Figure 1.1. In its barest form, it is a system consisting of a single manufacturing operation. Production lots of material arrive into the system, queue-up for service, are served and leave the system. In turn, the servicing of each lot consists of a setup operation on the processor, processing of each of the units in the lot, inspection of each unit and a rework operation on each defective unit found.

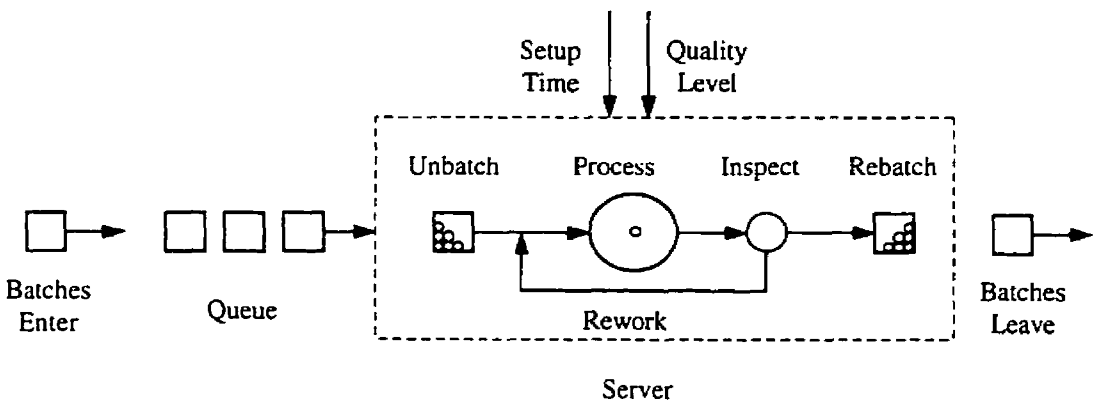


Figure 1.1: Model of Physical System

The economic system model is superimposed upon the physical system model in order to capture the costs of investments in improvement practices and the benefits derived in terms of reduced operating costs. Such a model is shown in Figure 1.2. Inputs are the amount of investment in each improvement practice, which are translated into setup time and quality level changes in the physical system model. The physical system model uses these variables to predict WIP level which, in turn, is translated into operating costs in the economic system model. By changing the levels of the two decision variables, changes to the total system operating cost can be estimated.

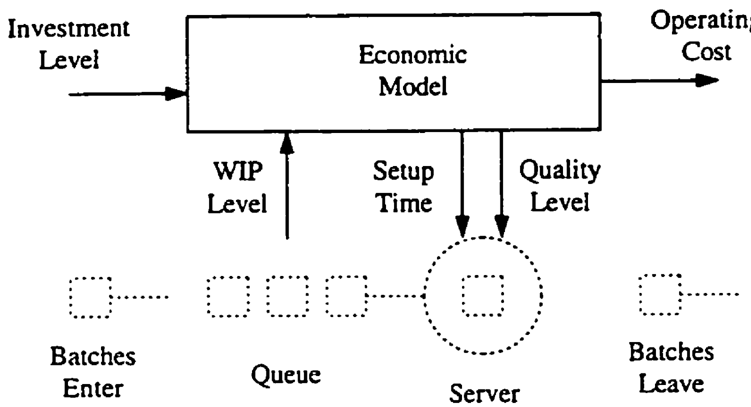


Figure 1.2: Model of Economic System

This type of compound model can then be used to determine investment decisions which result in minimum total costs for the manufacturing system. By comparing optimal investment levels for the cases where the improvement practices are assumed to be independent in terms of their effects on system performance and where they can potentially interact, the existence and significance of interactions can be examined and the magnitude of investment errors can be estimated.

For simplicity, the model will only consider a single production cell. Likewise, only a single product is considered. Since implementation time of setup and quality projects is typically much shorter than investment amortization periods used in industry, transient periods in the production system are not considered.

1.4 Organization of Remainder of Dissertation

In summary, the issues to be addressed in this study are:

- It is hypothesized that simultaneous improvements in setup time and quality level will interact when considering the performance of a manufacturing system. What impact may this interaction have on economic system performance?

- Interactions between improvements will also affect the decision to implement these practices. How might implementation decisions (i.e., investment levels) change when interactions are taken into account? What are the potential magnitudes of errors if they are not?

- Which factors affect the magnitudes of these potential errors?

Developing a model to capture this potential interaction and gaining insight into these questions is the objective of this research.

Chapter 2 presents a review of the literature relevant to the issues outlined in this introduction. It also shows how components of previous models can be adapted for use in examining the research issues in this study. Chapter 3 presents the development of an analytic model of the system shown schematically in Figures 1.1 and 1.2 for a specific case. With this model, a solution is obtained which provides insight into the interactive nature of the improvement practices and the consequences for decision makers. Assumptions in the basic model of Chapter 3 regarding the form of the investment-improvement function used in the economic model and whether production batch size is treated as a parameter or variable are then sequentially relaxed. The resulting variations of the basic model are then analyzed in Chapters 4, 5 and 6 and implications for the investment decision making process are examined. Overall results are summarized in Chapter 7, conclusions are drawn and recommendations are made for future research.

Chapter 2. Literature Review

In order to examine the behaviour of simultaneous setup reduction and quality improvement projects on a production system, a model must be developed. This chapter examines previous models of the setup and quality improvement process in manufacturing systems, explains why these models are not directly suitable for the current research problem, and discusses aspects of other models which will be utilized in this research.

2.1 EOQ-Based Setup and Quality Improvement Models

Setup reduction practices gained popularity in North American manufacturing in the early 1980's [Hall, 1983]. Arguments justifying investments in setup reduction at that time were intuitive and empirical in nature [e.g., Hall, 1983; Schonberger, 1982]. This changed, however, when Porteus [1985] introduced an Economic Order Quantity (EOQ) model which included setup reduction as a decision variable. A short explanation of his model is given here as much subsequent work relies on extension of this basic model.

In the classic EOQ model, two costs are considered: the cost of placing orders (setup cost) and the cost of holding finished goods cycle stock inventory. A manufacturing system is assumed to have a fixed production demand per period, D, which is produced in lots of size Q. The processing of each lot involves a setup operation, each of which costs K. The period cost of performing setups then becomes $K \cdot D/Q$. Average finished goods cycle stock inventory is assumed to be $Q/2$ in systems with constant demand. Given a period holding cost rate h, cycle inventory holding cost is then $h\ Q/2$. These two costs, as well as total cost, are plotted as a function of lot size in Figure 2.1.

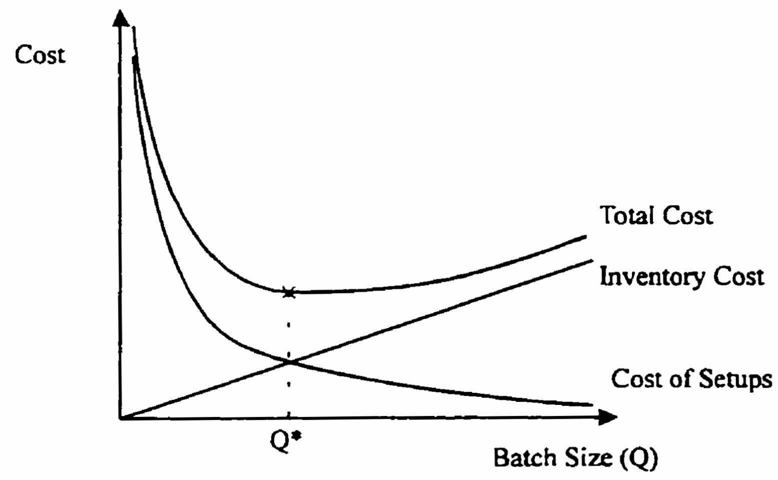


Figure 2.1: Setup, Inventory and Total Costs in EOQ Model

Total cost in the EOQ model is the sum of the setup and inventory costs and becomes:

$$TC(Q) = \frac{KD}{Q} + \frac{hQ}{2} \qquad (2.1.1)$$

By setting the first derivative of the total cost with respect to lot size to zero, the optimal lot size, or economic order quantity, Q*, is:

$$Q^* = \sqrt{\frac{2KD}{h}} \qquad (2.1.2)$$

So far in this model setup cost has been assumed to be a fixed parameter. In light of industrial practice, however, Porteus conjectured that setup cost could be reduced by investing in setup reduction projects. Accordingly, he introduced an investment function to relate the cost of performing a setup (K) to the amount invested in setup reduction, a_K. One case modeled by Porteus was the logarithmic investment function:

$$a_K = a - b\ln(K) \tag{2.1.3}$$

For $0 < K \leq K_0$

where $K_0 \equiv$ unreduced setup cost

a, b are positive constants

The one-time investment cost of making improvements was amortized into an undiscounted period cost through the use of a capitalized cost of capital, i. The new total cost for the system including setup reduction then became:

$$TC(Q, K) = \frac{KD}{Q} + \frac{hQ}{2} + ia_K(K) \tag{2.1.4}$$

By minimizing TC with respect to both lot size, Q, and setup cost, K, Porteus found the optimal values for these two variables to be:

$$Q^* = \min\left(\sqrt{\frac{2KD}{h}}, \frac{2bi}{h}\right) \tag{2.1.5a}$$

$$K^* = \min\left(K_0, \frac{2b^2i^2}{Dh}\right) \tag{2.1.5b}$$

In the event that no setup reduction takes place, the optimal lot size is the same as that predicted by the EOQ formula in Equation 2.1.2. However, if setup reduction is indicated, the optimal lot size will decrease.

This basic model has been adapted and expanded by Porteus and a number of other researchers. A summary of these adapted models (including Economic Production Quantity (EPQ) based models) is presented in Table 2.1.

Table 2.1: Adaptations of Porteus' 1985 Model

Author	Model and Extensions	Decision Variable
Porteus [1986a]	Discounted EOQ, Optimal sales rate	Setup
Porteus [1986b]	EOQ	Setup and Quality
Billington [1987]	EPQ with & w/o Discounting	Setup
Spence and Porteus [1987]	Multiproduct Capacitated EOQ	Setup
Zangwill [1987]	Dynamic Lot Sizing	Setup
Keller and Noori [1988]	(Q,r) Inventory Model	Setup
Rogers [1989]	General Convex and Concave Investment Functions	Setup
Cheng [1989]	EPQ with Geometric Programming Solution	Setup and Quality
Freeland et al. [1990]	Multiproduct EOQ	Setup
Goyal and Gunasekaran [1990]	Multistage EOQ	Setup
Nasri et al. [1990]	EOQ with Stochastic Lead Times	Setup
Cheng [1991a]	EPQ with Geometric Programming Solution	Quality
Cheng [1991b]	EOQ with Geometric Programming Solution	Quality
Kim et al. [1992]	EPQ with several Investment Functions	Setup
Kim and Arinze [1992]	EPQ results put into Expert System	Setup
Trevino et al. [1993]	EPQ model based on Empirical Data	Setup
Hwang et al. [1993]	Multiproduct Capacitated EOQ	Setup and Quality
Hong et al. [1993]	Dynamic Lot Sizing	Setup and Quality
Mekler [1993]	Dynamic Lot Sizing	Setup
Moon [1994]	Multiproduct Capacitated EOQ	Setup and Quality
Diaby [1995]	Dynamic Lot Sizing	Setup
Hong and Hayya [1995]	EOQ with Budget Constraint	Setup and Quality
Min and Chen [1995]	EOQ	Setup and Holding Costs
Banerjee et al. [1996]	Multiproduct Capacitated EOQ	Setup
Lee et al. [1996]	EOQ with Geometric Programming Solution	Quality
Soesilo and Min [1996]	EOQ	Quality
Sarker and Coates [1997]	EPQ with Variable Leadtimes and Geometric Programming Solution	Setup

A separate line of EOQ based models which study improvement to setup and/or quality via learning has also arisen from Porteus' 1985 model. Karwan et al. [1988] introduced an EOQ model in which setup costs decreased according to a learning curve function as the number of setups performed accumulated. Replogle [1988] developed a similar model, followed by Cheng [1991c], each using a somewhat different learning curve function than Karwan et al. Cheng followed his model with an EPQ model with setup learning [1994]. An EOQ model with 'myopic' setup learning, that is, where optimal lot size is only estimated one lot into the future, was recently provided by Rachamadugu and Tan [1997]. Improvement of quality through learning in an EOQ model has also been investigated, by Ocana and Zemel [1996]. While these variants of Porteus' original model study the effects of setup and quality improvement on a production system, they implicitly assume that improvements occur without capital expenditure. Similarly, Fine and Porteus [1989] had developed an EOQ model with learning in setup and quality improvement, but with the assumption that learning was induced through small non-capital stochastically-timed expenditures. As the objective of this research is to examine possible interactions between investment projects in setup and quality, these learning-based models have little applicability.

Another evolutionary path of Porteus' original model are for those models which have been adapted into multi-item inventory models, for instance, Hong and Hayya [1992] (with varying demands), Hong et al. [1992] (with setup investment) and Hong et al. [1996] (with one time vs. dynamic investment in setup). Gallego and Moon [1992] and [1995] have also considered multi-item economic lot scheduling problems with investment in setup time reduction. Similarly, Sung and Lee [1995] studied setup time reduction using a dynamic lot sizing model. While perhaps capturing a greater degree of the complexity of real-world production systems, these multi-item models have also been found to be too mathematically complex to obtain analytic solutions for optimum investment levels, so that their utility in this research is limited.

A feature common to all the EOQ-based models is that they are based on cost. An implication of using a cost basis is that utilization of the server is not modeled as a function of the decision variables. For instance, as setup times are reduced, service time of each batch should drop and server utilization is thus reduced. In the EOQ-based models, however, setup time is not considered explicitly, and server utilization is not predicted.

Knowledge of server utilization is important in this study because, as is known from basic queuing theory (e.g., Tijms [1994]), utilization is strongly correlated with queue lengths and WIP inventories. WIP inventory levels are arguably more important to manufacturing system costs than the finished goods cycle stock inventories used in EOQ based models [Boucher, 1984; Primrose, 1992]. For example, Karmarkar [1987] modeled WIP levels in a lot-based manufacturing system and found that for typical parameter values, the holding cost of WIP inventory was eight times as great as the cost of holding finished goods inventory. Because of this significant difference in the magnitude of costs, a model should take into account the cost of holding WIP, which EOQ based models do not. Furthermore, even if an EOQ model were to be adapted to consider the cost of WIP, being cost-based it would still fail to capture the changes to server utilization, and ultimately WIP level, expected from changes to the setup and quality decision variables.

For these reasons EOQ based models are not appropriate as the basis of a model for this study. However, elements of these models, such as treatment of setup and quality, investment functions and amortization of investment costs will be utilized in this research and are discussed in greater detail in subsequent sections of this chapter.

2.2 Inclusion of WIP Inventories

WIP reduction can be one of the larger sources of benefits for setup reduction projects [Boucher, 1984; Karmarkar, 1987; Primrose, 1992]. Surprisingly, only one setup reduction model has been found which links improvements in setup times to changes in WIP holding costs [Yang and Deane, 1993], and none which link quality improvement to WIP.

Two approaches to including WIP have been found in the literature: modeling WIP as a parameter, and estimating it as a function of the decision variables. The first approach, treating WIP as a constant parameter, was used by Goyal and Gunasekaran [1990] who recognized the significance of WIP costs and included them in their multistage economic production quantity (EPQ) model. Similarly, Hong [1994] developed a multi-product EPQ model which included both WIP costs and setup reduction. However, neither of these models treated WIP as a function of server utilization. As such, the expected linkage between changes to the decision variables and changes to WIP levels are not captured, so no benefits are obtained from WIP reduction.

A second approach to including WIP is to treat it as a function of the decision variables of the model. Karmarkar [1987] created a model with this property by modeling a manufacturing system as a queuing system. In his model he assumed that a stream of jobs arrived to the system and were batched into lots before being advanced to the server for processing. The server required a new setup before each lot was processed. With this model he was able to show that for small lot sizes, system WIP level increased due to the large numbers of setups performed while large lot sizes also caused high WIP levels due to the amount of material waiting in the batching operation and in the server. There was found to be a lot size which minimizes WIP level in the system. This behaviour is shown in Figure 2.2.

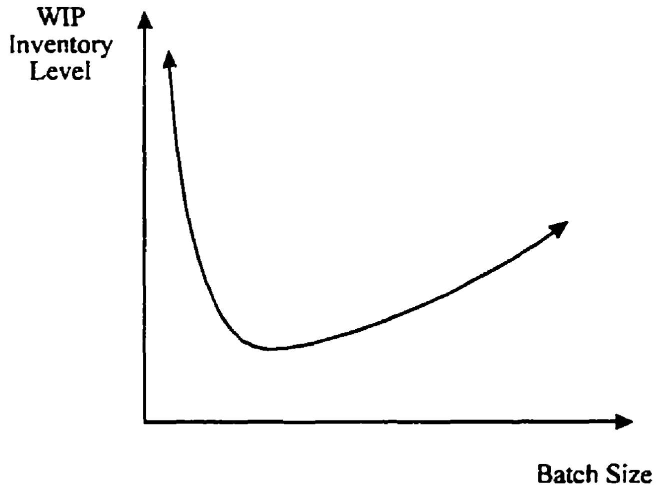


Figure 2.2: WIP Inventory as a function of Batch Size (following Karmarkar [1987])

A limitation of this work is that an M/M/1 type queue (exponentially-distributed interarrival and processing times, single server) was assumed. This assumption may not be reasonable since the mean service time changes, as the decision variables are changed, and the service time distribution would not be expected to remain exponential. Fortunately, this difficulty can be overcome by employing the M/G/1 type queue (general distribution of service times), for which analytic results are available [Buzacott and Shanthikumar 1993]. In fact, this is the approach taken by Yang and Deane [1993] who used an M/G/1 queuing model as the basis of a system to optimize batch size and setup reduction investment. With this model, they were able to show in general that there are diminishing returns from investment in setup reduction, but solutions for batch sizes and investment levels were not obtained, nor was any specific investment functional form used.

Two other recent adaptations to Karmarkar's 1987 model were found which examined optimal batch size in a system with WIP. Lambrecht et al. [1996] modeled a system in which jobs arrived individually and were batched before processing. They were able to develop solutions for the two cases of deterministic and Poisson arrivals with deterministic processing times. Tielemans and Kurk [1996] developed an approximate solution for a system with Erlang interarrival times of batches and general processing times. Neither of these models attempted to optimize setup time or quality level in addition to batch size.

Liu and Yang [1996] developed a renewal-theory model of a single server system in which a fraction of jobs were processed imperfectly, of which some could be reworked. This led to a system in which new jobs queued for service, as well as rework jobs. Service times included random processing times of units, which depended upon whether the unit was new or being reworked, and a random setup time for the server. This model was used to show that there is an optimal batch size which minimizes operating costs, although an analytic solution was not provided for this optimum batch size, nor was the possibility of investing in quality or setup improvement examined.

While no model has been found in the literature which determined optimum levels of investment in setup and quality improvement, the components necessary to create such model exist. Yang and Deane [1993] demonstrated that the use of an M/G/1 queue to capture WIP costs while postulating the existence of an optimal setup time. Liu and Yang [1996] showed that the service time of a queuing model can include setup time and time for dealing with quality problems. Thus, a queuing model can form the basis of the physical system model shown in Figure 1.1. This model will estimate WIP levels as a function of setup times and quality levels.

2.3 Additional Aspects of the Model

In addition to modelling the basic manufacturing system as a queue, the following aspects also need to be considered:

- Treatment of setup and quality decision variables.
- Relationship of investments in improvements to levels of the decision variables.
- Comparing capital costs to period costs.

The remaining subsections of this chapter will address previous work in each of these model characteristics.

2.3.1 Modelling Setup and Quality Improvements

Almost all the models listed in Table 2.1 dealt with the setup variable in terms of setup cost rather than setup time. An example of the inclusion of setup time in a model was given by Spence and Porteus [1987] who examined how setup time reduction increases system capacity by recognizing that setup times influence the utilization of the server. They included setup time as one component of service time for each batch processed, where service time of the batch was the sum of setup time and the time to process each of the units making up the batch. Karmarkar [1987] dealt with setup time in a similar manner. His model of a queuing system with variable lot sizes assumed that a setup was performed at the beginning of service of each batch.

Banerjee et al. [1996] produced a model in which the primary decision variable was setup cost, but the corresponding setup time was used as an optimization constraint on system utilization. As his model was based on the EOQ model, the objective function was expressed in terms of cost, which explains why the primary decision variable was setup cost. The need for a constraint on system utilization (that is, average utilization of the system must be less than 100%) expressed in terms of setup time, however, is puzzling. Intuitively, one would expect that in its initial state, the production system will be utilized less than 100%. Since setup time reduction will only decrease system utilization, it is not known when this constraint would ever become binding.

Spence and Porteus [1987], Karmarkar [1987] and others such as Yang and Deane [1993] and Liu and Yang [1996], used models which assumed that each batch of work entering processing has setup time added to processing time for the batch. Ultimately, this treatment of setup times is intuitively reasonable, and will be used in this study.

Quality, on the other hand, has been modelled in a variety of ways. Porteus [1986b], followed by Rosenblatt and Lee [1986] and Hong et al. [1993], assumed that a manufacturing process is brought 'in-control' (that is, makes acceptable product) after each setup. When processing individual parts in each batch, the process is assumed to go 'out of control' (that is, produces defectives) with a given probability for each part processed. Once out of control, the process stays out of control and continues to make defective units. The larger the batch, the greater the chance that the process will go out of control, and the greater the expected number of defective units produced.

One difficulty with this approach is that the effect of poor quality becomes a function of two variables: the probability of the process going out of control with each unit processed, and the batch size. This issue can cloud the distinction between changes in system performance brought about by changes to quality level and changes to setup times. For instance, reduced setup times can lead to smaller batch sizes, which in turn will result in fewer defects even though the quality decision variable (probability of the process going out of control with each unit produced) was not changed. This behaviour in a model makes the task of discovering interactive behaviour between the two decision variables very difficult due to this built-in interaction between setup time and quality performance.

Along lines opposite to that of Porteus [1986b] is Soesilo and Min [1996], who developed an EOQ-based model in which quality was the decision variable, and affected setup cost through a power-function relationship. This model was developed to represent situations such as the manufacture of microprocessors where the quality of the product is measured by the rated speed of the microprocessor, which is determined to a significant extent by the care taken in, and consequently the cost of, the setup operations when producing these components. Just as in Porteus' model, this representation explicitly links setup and quality levels.

Another approach to modelling quality, used by Cheng [1989, 1991a,b], Hwang et al. [1993] and Moon [1994], was to assume that a fixed fraction of production was defective. This might come about in a production system, for instance, if a fixed fraction of production is found defective and sent back to the server for rework regardless of whether a unit had been previously reworked. Given this assumption, the server would need to process more units in order to meet a target of demand of good units. By improving quality, the number of defective units decreases which reduces service time for each batch. Reduced service times result in lower server utilization and less expected WIP, providing a benefit to offset the cost of improving quality.

Liu and Yang [1996] extended this idea by assuming that defective items can be divided according to whether they are reworked or discarded. Each outcome was given a fixed probability of occurring. Furthermore, reworked items are assumed to be good, or discarded after rework, with a fixed probability. In order to develop solutions, Liu and Yang resorted to renewal theory with which they were able to show an optimum batch size exists which maximizes profit, although the value of the batch size must be calculated numerically. Investment in quality improvement was not considered in this model.

A variation of Cheng's [1989, 1991a,b], Hwang et al.'s [1993] and Moon's [1994] approach is to add the assumption that defective units can only be reworked once. Such an approach was used by Goyal and Gunasekaran [1990]. A consequence of this modified assumption is that the mathematical complexity of the resulting model is reduced. For the sake of simplicity, this approach will be used to model quality level in this study.

2.3.2 Relating Investments to Improvements

In order to make improvements to setup and quality levels in a manufacturing operation, investment is assumed necessary. In practice, there will be a finite number of improvement projects which will be considered for implementation, each of which will have an expected benefit for a given cost. Figure 2.3 shows an example of the relationship between the expected benefit and required investment for a number of projects under study.

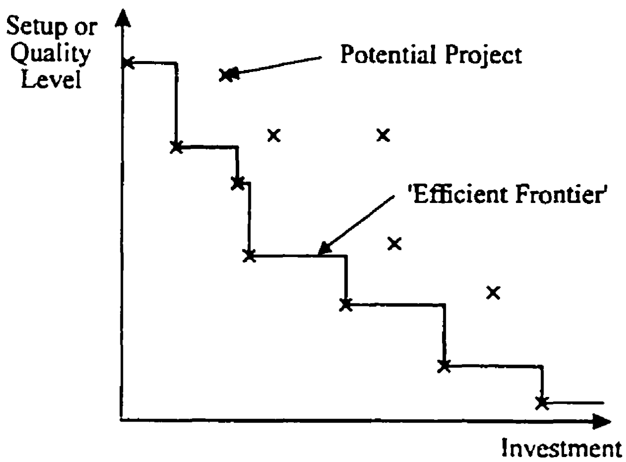


Figure 2.3: Investment vs. Improvement, Discrete Project Case

For instance, Leschke [1996] in a series of case studies was able to develop a classification scheme for setup reduction projects, dividing projects into "product level," "process level" and "policy level," depending upon whether the project affects the setup time for a single product (e.g., such as adapting a single die for a single product to a standardized shut-height), all the products in a single production cell (e.g., adding automatic die clamping to a stamping press which produces many products) or many production cells (e.g., training teams to reduce setup time in all production cells in a factory). Even under this classification system, as all potential projects regardless of category are aggregated, the cost-benefit structure will approach that shown in Figure 2.3.

As shown in Figure 2.3, some projects will provide a greater degree of improvement for a given investment than others. Decision makers will choose from among those projects giving the greatest benefit at each level of investment. The set of these choices will form the 'efficient frontier', which represents the investment-improvement function for the discrete projects case.

As the number of improvement projects considered is increased, the number of points on the 'efficient frontier' will also tend to increase. Through this process the investment function might be expected to eventually assume the shape of a smooth, continuous function such as that shown in Figure 2.4.

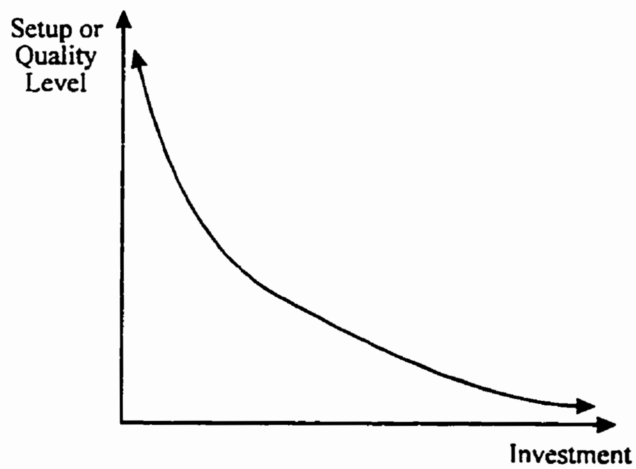


Figure 2.4: Continuous Case Investment-Improvement Function

Given this assumption that the investment-improvement function can be approximated by a smooth, continuous curve, it can be conveniently represented by analytic functions for use in mathematical models. For example, Porteus' model [1985] approximated setup cost alternatively with logarithmic and power functions of the investment in setup reduction. These forms were not only relatively simple mathematically, but also captured the idea of declining marginal returns as increased levels of investment are reached.

As Porteus' work was expanded by others, a number of other functional forms have been proposed by various authors. Table 2.2 presents a summary of the types of functions found in the literature, four of which are shown schematically in Figure 2.5.

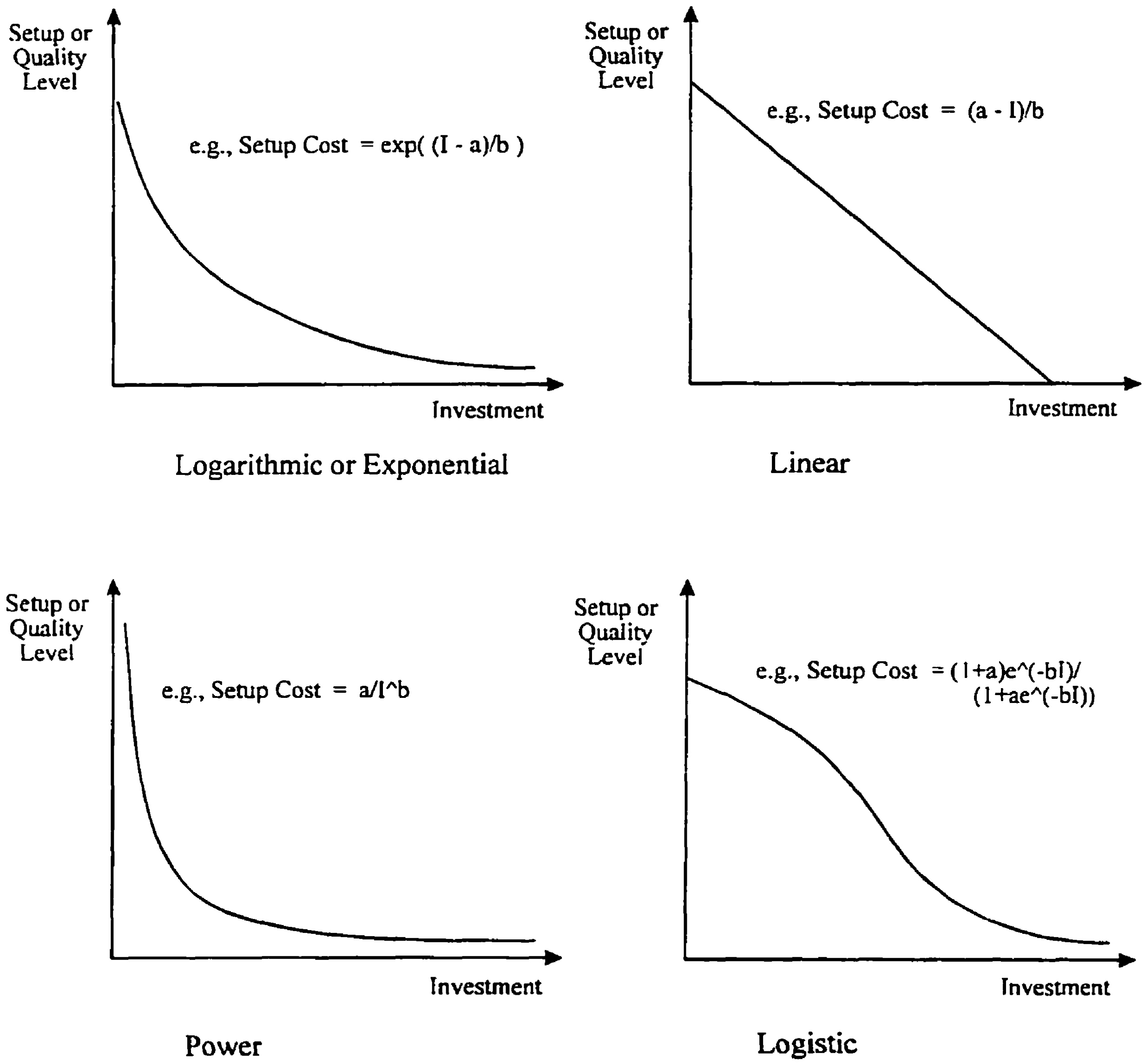


Figure 2.5: Shapes of Representative Investment Functions

Table 2.2: Investment-Improvement Functions Found in the Literature

Author	Functional Form
Chakravarty and Shtub [1985]	Arbitrary Discrete and Continuous
Porteus [1985]	Logarithmic*, Power
Porteus [1986a]	General Convex, Piecewise Linear
Porteus [1986b]	Logarithmic
Billington [1987]	Exponential*, Linear (both with upper and lower bounds)
Spence and Porteus [1987]	Logarithmic
Keller and Noori [1988]	Logarithmic, Power
Rogers [1989]	Linear, Arbitrary Convex and Concave
Cheng [1989]	Power
Nasri et al. [1990]	Logarithmic
Cheng [1991a], [1991b]	Power
Kim et al. [1992]	Linear, Parabolic, Logarithmic*, Exponential*, Logistic, Step
Kim and Arinze [1992]	Linear, Exponential, Logistic
Trevino et al. [1993]	Empirical Step, Exponential
Hwang et al. [1993]	Power
Hong et al. [1993]	Exponential and Power
Moon [1994]	Power
Diaby [1995]	Logarithmic, Power
Hong and Hayya [1995]	Arbitrary Convex and Concave
Min and Chen [1995]	Arbitrary Convex
Banerjee et al. [1996]	Exponential
Lee et al. [1996]	Power
Soesilo and Min [1996]	Power
Sarker and Coates [1997]	Exponential

* Logarithmic and exponential functions represent the same relationship expressed in terms of different decision variables (setup cost vs. investment). Kim et al. [1992] presents separate logarithmic and exponential functional forms due to his use of different upper and lower bounds built into the functions.

The choice of the functional form used appears to be primarily for reasons of mathematical tractability in the analysis. For instance, Cheng [1989], [1991a] and [1991b], Hwang et al. [1993], Moon [1994], Lee et al. [1996] and Soesilo and Min [1996] analyzed their models with geometric programming methods, and consequently used a power function to represent the investment-improvement function as power functions are amenable to geometric programming formulations.

To generalize, the majority of investment functions used by previous authors can be classed into two categories: constant returns to scale (i.e., linear) or decreasing returns to scale (i.e., convex, such as exponential or power functions). These two types of investment functions will be employed in this study, although the particular form of a convex investment function will be left until the analysis of Chapter 5.

2.3.3 Comparing Capital To Period Costs

An assumption, used implicitly or explicitly, by the authors of the models listed in Table 2.1 is that a one-time capital investment is made at the beginning of the project to bring about the improvement in the production system. In the model developed in this research, this cost is balanced against the benefits obtained from reduced WIP levels. An issue needing to be addressed is that the investment in improvements is a one-time cost while the benefits accrue over time. The magnitude of a one-time cost is not directly comparable to the periodic economic benefits obtained from the improvement(s).

Two approaches may be taken to make the costs and benefits comparable. The first approach is to discount the stream of periodic benefits to a present value and determine the net present value (NPV) of the improvement project. This approach has been used by Porteus [1986a], Billington [1987] and Rogers [1989]. Due to the mathematical complexity of the NPV calculation, these authors were only able to offer algorithms for the problem solution and present numeric examples.

The other approach is to convert the investment into an amortized period cost, and work with the total period operating costs of the system under study. This approach has been used by all the researchers listed in Table 2.1 with the exception of Porteus [1986a]. (Billington [1987] and Rogers [1989] used both the NPV and amortization approaches and found the results to be very similar.)

Authors who used the amortization approach did so for an infinite number of study periods, except for Trevino et *al.* [1993] who used a 5 year amortization period (although conversion from finite to infinite number of periods is easily accomplished, e.g., see Fraser et *al.* [1997]). With this approach, most authors were able to obtain analytic solutions to the problem of finding optimal levels of the decision variables.

Following the experience of the majority of previous researchers, costs of capital investments are dealt with by assuming the one-time investment in improvements is amortized over an infinite time span while savings also accrue over an infinite time span. Use of this approach increases the likelihood of obtaining analytic results while, as Billington [1987] and Rogers [1989] found, provide optimal solutions not significantly different from those obtained numerically through a NPV approach.

2.4 Discussion

Based on the findings of the review of the literature, the research issues of this study can be examined by building a model with the following characteristics:

- Use of an M/G/1 queuing model to estimate WIP inventory level as a function of the decision variables.
- Setup time modelled as a given amount of time added to the service time of each batch of material.
- Quality level modelled as a fixed probability of production found defective after processing and returned to the server for a single rework.
- An investment-improvement function to relate an investment in improvements to the level of each decision variable.
- Amortization of the one-time improvement investments over an infinite time span to convert investments to period costs.

Models incorporating some, but not all, of these characteristics were found in the literature. Of these models, none have been used to examine how setup and quality might interact when considering the performance of the manufacturing system and how such interactions might affect the decision to make improvements. Creating such a model and exploring these questions will represent the contribution of this research.

In the next chapter, these components will be assembled into a model of a manufacturing system which will be used to explore the issues outlined in this study.

Chapter 3: Modeling

In this chapter a mathematical model of a production system is developed. The purpose of this model is to study how changes to both setup time and quality level might interact to affect the economic performance of the system, and to gain insight into how this potential interaction may affect management decisions to invest in those two improvement practices.

In an effort to manage the complexity of the solutions, and to gain additional insight into the behaviour of the optimal investment decisions, four variants of the basic model will be studied. Two types of investment function, linear and convex, will be used. Batch size will also have two treatments, as a fixed parameter and as a decision variable. This leads to the following combinations which are studied in this dissertation:

- Linear Investment Functions, Fixed Batch Size (this chapter)
- Linear Investment Functions, Variable Batch Size (Chapter 4)
- Convex Investment Functions, Fixed Batch Size (Chapter 5)
- Convex Investment Functions, Variable Batch Size (Chapter 6)

The basic model is developed in this chapter, and results are obtained for the Linear Investment Function, Fixed Batch Size case.

The schematic model of the system from Figure 1.2 is repeated in Figure 3.1. In this system, work arrives in batches, waits in the queue, is processed, and leaves the system. Upstream operations are aggregated as an arrival process to this cell, and downstream operations can be ignored by assuming that once work is completed it leaves the cell and has no more effect upon it.

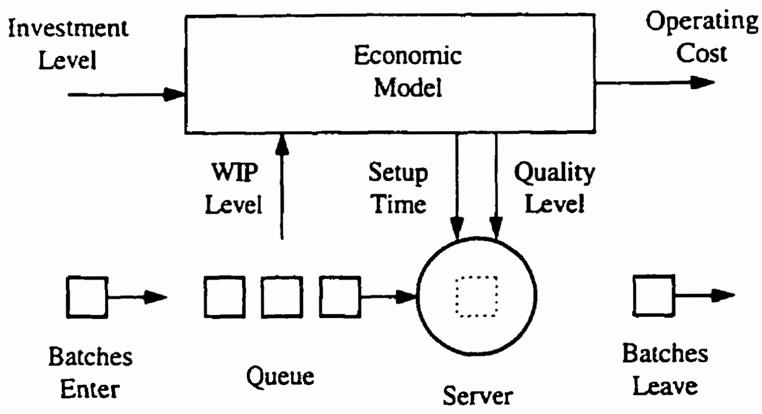


Figure 3.1: Schematic of Manufacturing System

By modeling the manufacturing system as a queuing system, WIP inventories can be estimated as a function of the decision variables: setup time and quality level. The economic model superimposed upon this physical model converts the levels of the decision variables into costs of investments in improvements and the WIP level into an inventory holding cost. Selecting levels of the decision variables which balance the cost of investing against the benefit obtained from WIP reduction provides the minimum operating cost for the system and the optimal investment decision. The existence of interactions can then be determined by comparing optimal investment decisions when decision variables are optimized individually or simultaneously.

The remaining sections of this chapter will describe the model assumptions, develop the physical and economic models for this system, and determine optimum investment decisions. By comparing the results for a case where investments in each practice are assumed independent to a case where potential interactions are considered, the effect of interactions on investment decisions will be determined.

3.1 Overview and Assumptions used in the Model

Figure 3.2 expands the server portion of the model shown in Figure 3.1. Batches of material completed by an upstream operation are transported into the cell and join the server queue. If the server is empty, a batch is moved into it. The batch is processed within the server, and when complete, leaves the system. The cell is considered to be the server and server queue, and material in these locations represents the WIP inventory of the system.

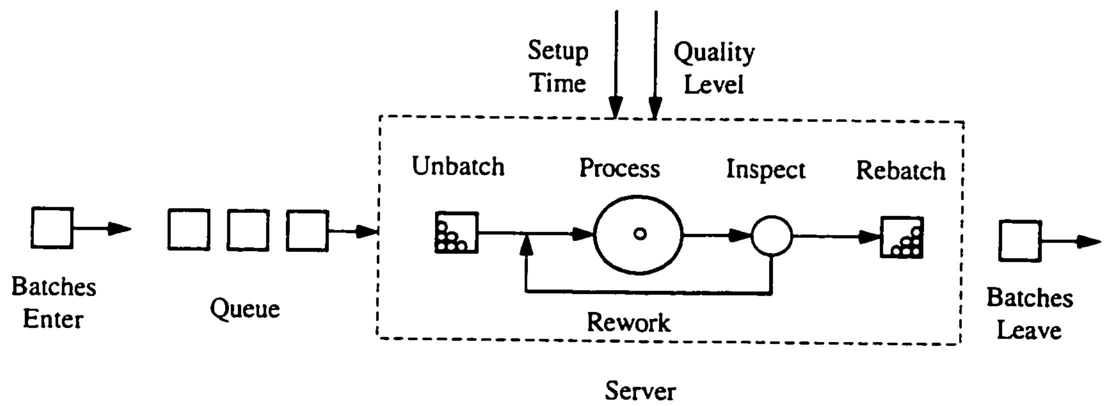


Figure 3.2: Detailed View of Physical System Model.

As each batch enters the processor, a new setup of the processor must take place. Setups will be assumed to require a fixed amount of time for each batch, but that time can be reduced by investing in setup reduction.

Once the setup is complete, processing of the batch begins. Each unit of the batch is processed sequentially by the server. As each unit is completed, it is assumed to be immediately inspected. Defective items are immediately reworked. (This is an increasingly common strategy in manufacturing [Bakerjian, 1993].) Rework is assumed to require the same amount of time as the initial processing, require no additional raw material and result in a good unit after a single reworking. Thus, only full batches enter and leave the cell. It is further assumed that the likelihood of a processed unit being defective is constant over time with a fixed probability of being defective. This probability can only be reduced by investing in quality improvement.

Previous models (e.g., see Table 2.1) have included batch size as a decision variable. While treating batch size as a variable will yield a richer model, it also results in a substantial increase in mathematical complexity of the solution process and a corresponding decrease in the number of closed-form results which can be obtained. For these reasons, this research considers models with both fixed batch sizes (Chapters 3 and 5) and variable batch sizes (Chapters 4 and 6).

Tijms [1994] has described the Poisson process to be a reasonable representation of the arrival of jobs in a manufacturing system, and this assumption will be used here. Processing times will be assumed stochastic and are described with a general service time distribution. With these two assumptions, this system can be modelled on the physical level as a M/G/1 type queue, which will be used to predict steady-state flow time for material in the cell. By invoking Little's Law [Tijms, 1994], the flow time is related to the WIP level of the system.

Only steady-state conditions will be studied in this model. Intervention in the system (i.e., implementation of improvement practices) is assumed to require insignificant time, and full benefits are accrued immediately.

An economic model is superimposed upon the physical model by attaching costs to investment in setup and quality improvements and to the cost of holding WIP inventories. Processing and material costs are assumed to be independent of changes in setup and quality levels, so these costs are not included in the economic model. Of the functions found in the literature to relate investment to improvement, the linear function will be used in this chapter in order to facilitate examination of the behaviour of the system while limiting the mathematical complexity of the model at this stage. A more complex investment function, a convex function, is studied in Chapters 5 and 6.

It is assumed that WIP holding costs are proportional to the number of batches in the queue plus any batch undergoing processing, since only complete batches are transported into and out of the cell. While there will be some completed units within the batch undergoing processing which would be expected to have a higher economic value, it is assumed that the value of the batch will only increase once all its units are completed, at which time the batch leaves the cell.

The remaining sections of this chapter develop the mathematical models of this system and add further assumptions as needed.

3.2 Model of Physical System

3.2.1 Nomenclature

The following set of symbols will be used in the mathematical models of this section. The first group are the basic parameters used in the model:

D	arrival rate of job orders	[units/period]
Q	lot size of batches	[units/batch]
τ	time to do each setup†	[period/batch]
r	probability of producing defective unit†	[%]
X	processing time for each unit	[period/unit]

The subsequent group represent derived quantities used in the modeling:

S	processing time for each batch	[period]
W	time a unit spends in the system	[period]
λ	arrival rate of batches of jobs (= D/Q)	[batches/period]
ρ	utilization rate of server (= D E[S]/Q)	[%]

(The † notation indicates the decision variables in this model, namely, setup time and quality level.)

The processing time of each unit, X, is a random variable with a positive, general distribution. The variables S and W are also positive random variables which are derived as functions of X.

3.2.2 Development of Model

The physical system is modelled as an M/G/1 queue. It is assumed that complete batches arrive to the system according to a Poisson arrival process. A first-come, first-serve discipline is used in the queue. As well, batches are kept intact (i.e., no splitting or combining batches) during their time in the system.

As each batch enters the server for processing it causes a new setup of the server. Setup is followed by processing each of the individual units and then by rework of the fraction of items found to be defective. It is assumed that defective units are only reworked once and that the reworking time and initial processing time have the same distribution. The setup operation occupies the server for a set time, τ, during which no processing can take place. The batch service time, S, is then the sum of these times:

$$S = \tau + \left(X_1 + X_2 + \ldots + X_Q\right) + \left(X_{Q+1} + X_{Q+2} + \ldots + X_{Q+rQ}\right)$$

By assuming processing times are independent and identically distributed, and that they are also independent of setup times, the expected service time, E[S], can be written as:

$$\begin{aligned} E[S] &= E\left[\tau + \left(X_1 + X_2 + \ldots + X_Q\right) + \left(X_{Q+1} + X_{Q+2} + \ldots + X_{Q+rQ}\right)\right] \\ &= E[\tau] + QE[X] + rQE[X] \qquad (3.2.1) \\ &= E[\tau] + (1 + r)QE[X] \end{aligned}$$

(The notation E[•] and E[•²] is used to represent the first and second moments, respectively, of the variable • and will be used in subsequent equations.)

For the second moment of service times, consider the variance of S:

$$\mathrm{Var}[S] = \mathrm{Var}\left[\tau + \left(X_1 + X_2 + \ldots + X_Q\right) + \left(X_{Q+1} + X_{Q+2} + \ldots + X_{Q+rQ}\right)\right]$$

Which, due to independence of setup and processing times, can be written as:

$$\begin{aligned}\mathrm{Var}[S] &= \mathrm{Var}[\tau] + Q\mathrm{Var}[X] + rQ\mathrm{Var}[X] \\ &= \mathrm{Var}[\tau] + (1 + r)Q\mathrm{Var}[X] \\ &= E[\tau^2] - E[\tau]^2 + (1 + r)Q\left\{E[X^2] - E[X]^2\right\}\end{aligned}$$

Now, rearranging $\mathrm{Var}[S] = E[S^2] - E[S]^2$ for $E[S^2]$ and substituting gives:

$$\begin{aligned}E[S^2] &= \mathrm{Var}[S] + E[S]^2 \\ &= E[\tau^2] - E[\tau]^2 + (1 + r)Q\left\{E[X^2] - E[X]^2\right\} + \left\{E[\tau] + (1 + r)QE[X]\right\}^2 \\ &= E[\tau^2] + (1 + r)Q\left\{E[X^2] + 2E[\tau]E[X] + (1 + r)QE[X]^2\right\} - (1 + r^2)QE[X]^2\end{aligned} \tag{3.2.2}$$

Equation 3.2.2 provides for the general case in which setup time is a random variable. However, in this study it will be assumed to be deterministic as a simplification measure, which results in the relationship in Equation 3.2.3:

$$E[S^2] = \tau^2 + (1 + r)Q\left\{E[X^2] + 2\tau E[X] + (1 + r)QE[X]^2\right\} - (1 + r^2)QE[X]^2 \tag{3.2.3}$$

From the M/G/1 queue, the expected time in the system for a random customer is given as [Buzacott and Shanthikumar, 1993]:

$$E[W] = \frac{\lambda E[S^2]}{2(1-\rho)} + E[S] \tag{3.2.4}$$

Which for this system becomes:

$$E[W] = \frac{\frac{D}{Q} E[S^2]}{2(1 - \frac{D}{Q} E[S])} + E[S] \tag{3.2.5}$$

where E[S] and E[S^2] are as given in Equations 3.2.1 and 3.2.2 and the identities λ = D/Q and ρ = D/Q E[S] were used.

The level of WIP, that is, the expected number of batches in the system is related to the expected time in the system E[W] from Equation 3.2.5 by invoking Little's Law [Tijms, 1994], i.e., E[WIP] = λ E[W]. Thus, given the setup time and quality level, this model can predict the WIP level of the manufacturing system.

3.3 Economic System Model

An economic model is now developed to capture the cost of improving the two decision variables, setup time and quality level, and the economic benefit to the system from reduced WIP levels. Since costs must be compared on the basis of equivalent time frames, the economic model may consider costs on a per-period basis or in terms of a present worth. This study uses the former, per-period costs, although costs can easily be converted into an equivalent present worth, if desired.

3.3.1 Nomenclature

The following symbols are introduced:

τ_0	setup time of initial system	[period]
T	normalized setup time (= τ/τ_0)	[$0 \leq T \leq 1$]
r_0	initial defect rate	[%]
R	normalized defect rate (= r/r_0)	[$0 \leq R \leq 1$]
h	holding cost rate for WIP inventory	[$/unit/period]
i	effective interest rate on investment	[%/period]
I_S	investment in setup reduction	[$, one time cost]
I_Q	investment in quality improvement	[$, one time cost]
a	setup investment function slope parameter	[period/$]
b	quality investment function slope parameter	[%/$]

3.3.2 Model Development

There are a variety of costs one might include in characterizing a production system. One particular view is to break down the costs into the following categories:

- Raw Material
- Direct Labour

- System Overhead
- WIP Holding Cost
- Cost of Investing in Improvements

The first, raw material, is a fixed cost per period since it has been assumed that demand per period is fixed and reworking defective items does not require extra material.

Direct labour is a variable cost in the long term. However, due to the practical difficulties in adjusting labour supply to compensate for small changes in server utilization resulting from improvements to the system, direct labour is assumed to be a fixed cost per period. Similarly, overhead costs for this production cell will be assumed to be fixed per period.

This leaves the cost of holding WIP and the cost of investing in improvements as the two variable costs in this model. The primary effect on the server of decreasing setup times and defect rates is to reduce batch service times (E[S]), and hence, server utilization (since utilization = $\lambda E[S]$). From basic queuing theory it is known that decreased server utilization results in decreased queue length; so cost of holding WIP inventory is directly affected by improvements to the system.

This model assumes that improvements occur immediately after one-time capital investments are made. However, WIP holding cost is a period cost. To make the two costs comparable, a Capitalized Cost formula [Fraser et al., 1997] is used to convert the one-time investment values into period costs, i.e.,

$$A = I\ i \qquad (3.3.1)$$

Where: A = period cost of investment

I = amount of investment

i = period interest rate

This formula assumes the cost of the investment is spread over an infinite time span. This in itself is not realistic, however, for finite study periods an effective interest rate (i_e) can be determined from the Capital Recovery factor (i_e = (A/P, i%, N) [Fraser, 1997]) and used in Equation 3.3.1.

The average time spent in the system by each batch (E[W]) was given in Equation 3.2.5. Multiplying by the arrival rate of units to the system (D) yields the average number of units of WIP inventory (via Little's Law), which can then be used to determine average WIP holding cost with the help of a holding cost rate (h). Thus, WIP holding cost per period becomes:

$$h\ D\ E[W] \tag{3.3.2}$$

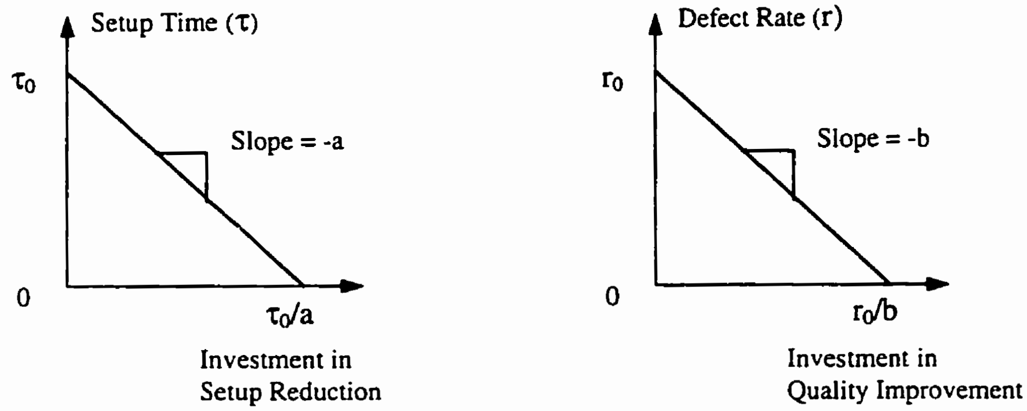
The expected total period cost for the system, E[TC], is the sum of these costs and can be written as:

$$E[TC] = h\ D\ E[W] + (I_S + I_Q)\ i + \{\text{Fixed Costs}\} \tag{3.3.3}$$

where E[W] is from Equation 3.2.5.

The first term in 3.3.3 represents the cost of holding WIP while the second represents the period cost of investing in improvements. The remaining costs in the system have been assumed not to be affected by changes in setup or quality level and are aggregated into the "{Fixed Costs}" term.

In order to relate investments in improvements to the level of the setup time and defect rate in this model, an investment function is necessary. The linear investment function is used in this chapter and in Chapter 4, while a convex investment function is examined in Chapters 5 and 6. The linear investment function is shown graphically in Figures 3.3a and 3.3b for setup and quality, respectively.



Figures 3.3a and 3.3b: Investment Functions for Setup and Quality Improvement

The investment-improvement functions can be written as:

$$I_S = \frac{\tau_0 - \tau}{a} \tag{3.3.4a}$$

$$I_Q = \frac{r_0 - r}{b} \tag{3.3.4b}$$

where the parameters 'a' and 'b' represent the marginal rate of improvement from investments in each factor. A greater value of these parameters represents a lower marginal cost of making improvements to each factor. It may be noted that a property of the linear investment function is that returns to scale are constant, that is, each increment of improvement has a constant cost. The convex investment function, discussed in Chapters 5 and 6, differs primarily in that it exhibits declining returns to scale behaviour.

Using normalized values for setup time and defect rate (i.e., $T = \tau/\tau_0$; $R = r/r_0$; $0 \leq T, R \leq 1$) changes these to:

$$I_S = \frac{(1 - T)\tau_0}{a} \tag{3.3.5a}$$

$$I_Q = \frac{(1 - R)r_0}{b} \tag{3.3.5b}$$

Substituting 3.3.5a, 3.3.5b and 3.2.4 into 3.3.3 gives the expression for the expected total cost per period:

$$\begin{aligned} E[TC] &= hD\left[\frac{\frac{D}{Q}E[S^2]}{2(1 - \frac{D}{Q}E[S])} + E[S]\right] + i\left[\frac{(1 - T)\tau_0}{a} + \frac{(1 - R)r_0}{b}\right] \\ &= hD\left[\frac{\frac{D}{Q}\left[\tau^2 + (1 + r)Q\left\{E[X^2] + 2\tau E[X] + (1 + r)QE[X]^2\right\} - (1 + r^2)QE[X]^2\right]}{2\left(1 - \frac{D}{Q}\{\tau + (1 + r)QE[X]\}\right)} \right. \\ &\quad \left. + \{\tau + (1 + r)QE[X]\}\right] + i\left[\frac{(1 - T)\tau_0}{a} + \frac{(1 - R)r_0}{b}\right] \end{aligned} \tag{3.3.6}$$

The "{Fixed Costs}" term has been omitted since it is a constant with respect to the decision variables and does not affect optimization results.

This equation, which includes the first and second moments of service time developed in Equations 3.2.1 and 3.2.2, predicts the expected cost of operating the manufacturing system shown in Figure 3.2 as a function of setup and quality levels, T and R. By establishing this relationship, levels of setup and quality can be found which minimize costs. This will be done in the next section.

3.4 'Naive' Investment Decisions

The example investment situation described in Chapter 1 is now re-examined and the decision strategies that would be applied in such a situation are developed more formally. To recap, a quality engineer was charged with the task of proposing investment projects which improve quality while a manufacturing engineer was responsible for suggesting projects to reduce setup times. The separate project proposals were then forwarded to company management for evaluation and approved or rejected based on their expected economic performance. An implicit assumption used by the decision maker was that the projects were independent of each other and could be correctly evaluated separately.

With the E[TC] function developed in the last section, optimal investment decisions appropriate for such projects under the assumption of independence of projects are now determined. Figure 3.4 shows the E[TC] function plotted as a function of investment in the two decision variables. This function behaves nicely due to the E[TC] function (Equation 3.3.6) being a second-order function of the decision variables. As such, critical points representing local minima within the feasible space of each decision variable (e.g., $0 \leq T \leq 1$ and $0 \leq R \leq 1$) will also represent the global minimizer of E[TC].

Consider the case of improvement in setup time alone. The goal is to find the setup time which minimizes E[TC]. The partial derivative of E[TC] with respect to the normalized setup time T is determined and set to zero:

$$\frac{\partial E[TC]}{\partial T} = hD\tau_0 \left[\frac{2(1 - D/Q\, E[S]) + D^2/Q^2\, E[S^2]}{2(1 - D/Q\, E[S])^2} + 1 \right] - i\frac{\tau_0}{a} \tag{3.4.1}$$

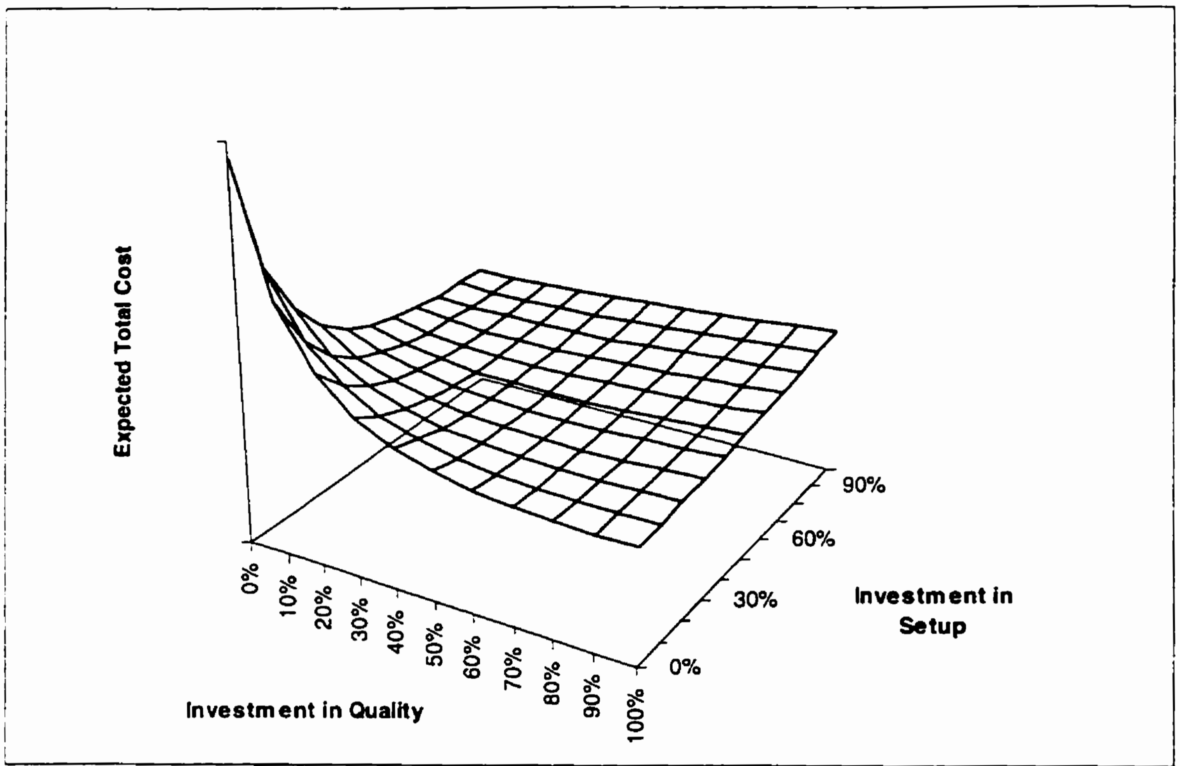


Figure 3.4: Shape of E[TC] Function vs. Decision Variables.

Solving Equation 3.4.1 for T provides two critical values:

$$T^* = \frac{Q}{D\tau_0}\{1 - D(1 + r_0)E[X]\} \pm \frac{1}{\tau_0}\sqrt{\frac{haQ^2(1-\rho)}{D(2i - hDa)}} \qquad (3.4.2)$$

Checking the second derivative shows the minimum E[TC] occurs at:

$$T_{min}^* = \frac{Q}{D\tau_0}\{1 - D(1 + r_0)E[X]\} + \frac{1}{\tau_0}\sqrt{\frac{haQ^2(1-\rho)}{D(2i - hDa)}} \qquad (3.4.3)$$

A consequence of the use of the linear investment function is that three different investment decisions are possible for each factor, namely:

1) No Investment (i.e., $T^*, R^* = 1$)

2) Some Investment (i.e., $0 < T^*, R^* < 1$)

3) Full Investment (i.e., $T^*, R^* = 0$)

Where T^*, R^* are the optimal normalized setup time and defect rate.

There are two 'transitions' between these three cases, namely, between "No" and "Some" investment, and between "Some" and "Full" investment. The location of these transitions can thought of as boundaries between the different investment decisions, in which the boundaries fall at the points where exactly a) no investment, and b) full investment occur. These points correspond to $T^* = 1$ and $T^* = 0$ respectively for the setup function and will be defined as the investment 'boundaries' in the following sections.

The transitions between the three investment decisions are a function of the parameter values in a system. Intuitively, one will expect that as the marginal cost of improving setup time increases, less and less setup improvement will be called for until the optimal decision will be to do no setup reduction. Similarly, as the marginal cost of reducing setup time drops, more setup reduction will be called for until setup is completely eliminated. This idea can be expressed in terms of critical values of the slope of the investment function curve, 'a.' For instance, rearranging Equation 3.4.3 provides an expression for the critical parameter 'a*' as a function of T^*:

$$a^* = \frac{2i}{hD}\left[\frac{\left\{1-\lambda\left[T^*\tau_0+(1+Rr_0)QE[X]\right]\right\}^2}{1+\lambda^2\left[(1+Rr_0)E[X^2]-(1+R^2r_0^2)QE[X]\right]+\left\{1-\lambda\left[T^*\tau_0+(1+Rr_0)QE[X]\right]\right\}^2}\right] \tag{3.4.4}$$

Quality level is not being changed in this case, so the variable R will be set to 1 (i.e., set to initial quality level).

A more convenient way to express this critical value is in terms of the cost to eliminate all setup, τ_0/a^*. Rearranging 3.4.4 in terms of this value gives:

$$\frac{\tau_0}{a^*} = \frac{hD\tau_0}{2i}\left[1 + \frac{1+\lambda^2\left[(1+Rr_0)E[X^2] - (1+R^2r_0^2)QE[X]\right]}{\left\{1-\lambda\left[T^*\tau_0 + (1+Rr_0)QE[X]\right]\right\}^2}\right] \tag{3.4.5}$$

Remembering that the transitions between the investment decision cases occurred when optimal setup time $T^* = 0$ or 1 exactly, the expressions for the location of the two investment boundaries becomes:

a) No Investment Boundary (i.e., $T^* = 1$)

$$\left.\frac{\tau_0}{a^*}\right|_{T^*=1} = \frac{hD\tau_0}{2i}\left[1 + \frac{1+\lambda^2\left[(1+r_0)E[X^2] - (1+r_0^2)QE[X]\right]}{\left\{1-\lambda\left[\tau_0 + (1+r_0)QE[X]\right]\right\}^2}\right] \tag{3.4.6a}$$

b) Full Investment Boundary (i.e., $T^* = 0$)

$$\left.\frac{\tau_0}{a^*}\right|_{T^*=0} = \frac{hD\tau_0}{2i}\left[1 + \frac{1+\lambda^2\left[(1+r_0)E[X^2] - (1+r_0^2)QE[X]\right]}{\left\{1-\lambda(1+r_0)QE[X]\right\}^2}\right] \tag{3.4.6b}$$

That is, if the cost to eliminate all setup is greater than $\tau_0/a^*|_{T^*=1}$, no investment should take place. Similarly, if the cost is less than $\tau_0/a^*|_{T^*=0}$, all setup time should be eliminated. Between these boundaries, setup time should be reduced but not eliminated to minimize total system operating costs.

A similar process in terms of defect rates yields the following equations for the critical costs to eliminate defects:

a) No Investment Boundary (i.e., R* = 1)

$$\left.\frac{r_0}{b^*}\right|_{R^*=1} = \frac{hDr_0QE[X]}{2i}\left[1 - \frac{2\lambda r_0 E[X]}{\{1-\lambda[\tau_0 + (1+r_0)QE[X]]\}} + \frac{1-\lambda^2(1+r_0^2)QE[X]^2}{\{1-\lambda[\tau_0 + (1+r_0)QE[X]]\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{\lambda(Q-D\tau_0)E[X^2]}{\{1-\lambda[\tau_0 + (1+r_0)QE[X]]\}^2}\right]$$

(3.4.7a)

b) Full Investment Boundary (i.e., R* = 0)

$$\left.\frac{r_0}{b^*}\right|_{R^*=0} = \frac{hDr_0QE[X]}{2i}\left[1 + \frac{1-\lambda^2 QE[X]^2}{\{1-\lambda[\tau_0 + QE[X]]\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{\lambda(Q-D\tau_0)E[X^2]}{\{1-\lambda[\tau_0 + QE[X]]\}^2}\right] \qquad (3.4.7b)$$

A convenient method of depicting the decision boundaries for the two factors is with the decision matrix shown in Figure 3.5. This decision matrix is used in the following manner. For a given manufacturing system under study, the engineers involved estimate the costs to eliminate both setup times and defects. These two costs are then plotted as a point, (τ_0/a, r_0/b), on this decision matrix. Where this point falls on this matrix relative to the decision boundaries determines the relevant investment decision. For instance, a point falling in the lower left zone indicates an optimal decision to fully eliminate both setup time and defects.

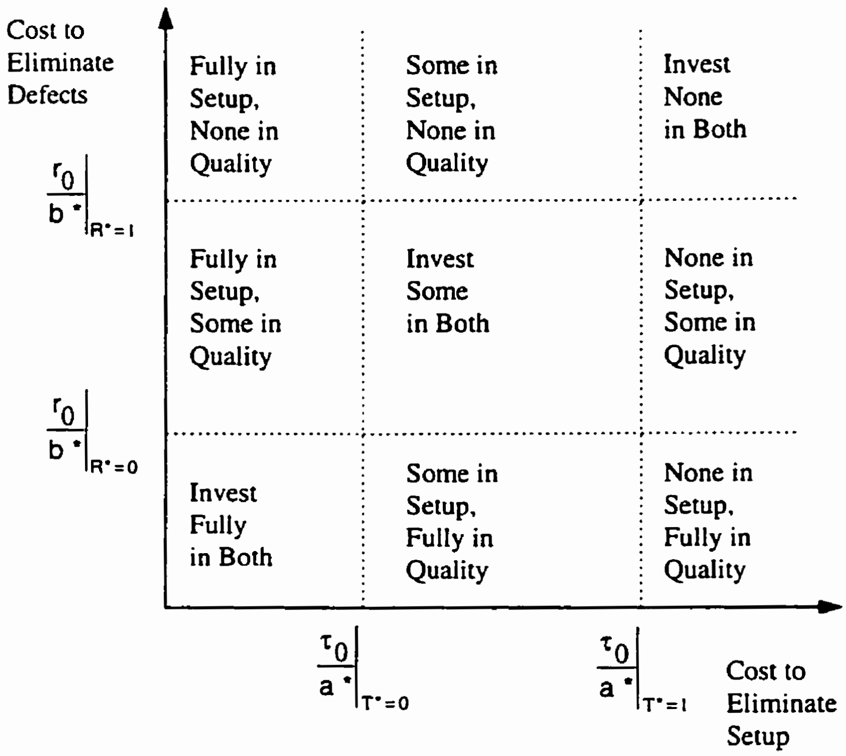


Figure 3.5: Schematic of Decision Matrix based on Costs to Eliminate Setup/Defects.

To summarize, in the case where projects are implicitly assumed to be independent of each other, each project has two critical costs of improvement. If the cost of improving a practice falls below the critical cost of making full improvements, then the maximum amount is invested in that practice. If the cost falls above the critical cost of making no improvements, then no investment is made. If the cost falls between the two critical values, then some investment is optimal. Investment in both the practices studied here, setup time reduction and quality improvement, is optimized by this strategy.

Unfortunately, there is a flaw in this method of making investment decisions. The two projects appear to the decision maker to be unconnected to each other and might reasonably be assumed independent for capital allocation decisions. However, both decision variables affect the physical system in the same way, that is, improvement of both setup times and quality levels reduce server utilization, which in turn reduces WIP levels and costs. Since the system costs are a function of both decision variables, implementing both setup time and quality level changes on the system simultaneously will alter the total system operating costs in ways not predicted in the individual project proposals. Because the operating costs assumed in the project proposals do not take simultaneous implementation of the decision variables into account, the optimal investment levels calculated in the proposals may not be correct. The next section will study the case when changes to both decision variables are considered. The difference in optimal decisions between the two approaches ('naive' vs. 'informed') will then be examined to estimate the likelihood of making incorrect investment decisions and the significance of the errors involved.

3.5 'Informed' Decision Making

To avoid the potential error described in the previous section, the expected total cost function must be optimized simultaneously for both decision variables in this problem. This will be the focus of this section.

3.5.1 Bivariate Optimization

The problem at hand is to find the point of minimum expected total cost subject to the constraints:

a) Setup time ≥ 0

b) Defect rate ≥ 0

c) Setup time $\leq \tau_0$

d) Defect rate $\leq r_0$

The first pair of constraints are non-negativity on the setup time and defect rate while the second pair stem from the assumption that the system will not degenerate from the initial conditions, that is, setup time and defect rates will not become worse than their original levels.

The two decision variables then can range from 0 to τ_0 and 0 to r_0, respectively. The space defined by these ranges becomes the feasible region for the decision variables in this problem. A sketch of the feasible region is shown in Figure 3.6. Initially the system starts with values of (τ_0, r_0) for setup and quality (upper right corner on the sketch) and moves down and to the left as improvements are made.

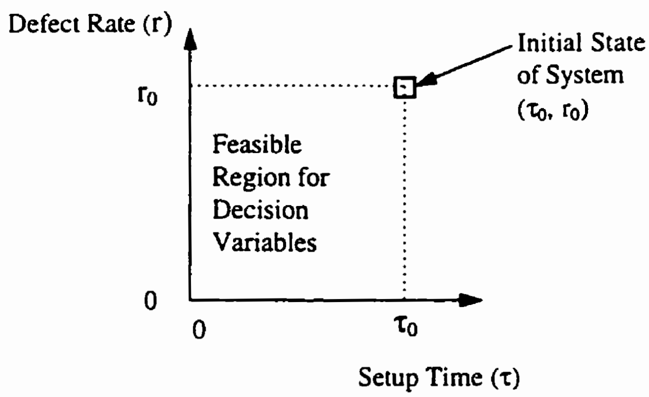


Figure 3.6: Feasible Region of E[TC] Optimization Problem.

To determine the necessary condition for the minimum expected total cost, the partial derivatives of the E[TC] function are set to zero to discover where any critical points might exist.

$$\frac{\partial E[TC]}{\partial T} = hD\tau_0\left[\frac{2(1-\lambda E[S]) + \lambda^2 E[S^2]}{2(1-\lambda E[S])^2} + 1\right] - i\frac{\tau_0}{a} \tag{3.5.1a}$$

$$\frac{\partial E[TC]}{\partial R} = -i\frac{r_0}{b} + hDr_0Q\left[\frac{\lambda^2 E[S^2]E[X]}{2(1-\lambda E[S])^2} + E[X] + \frac{\lambda\left\{E[X^2] + 2T\tau_0 E[X] + 2(1+Rr_0)QE[X]^2 - 2Rr_0E[X]\right\}}{2(1-\lambda E[S])^2}\right] \tag{3.5.1b}$$

Before solving this system, the sufficient condition is determined by examining the determinant of the Hessian of the E[TC] function:

$$\det H(E[TC]) = -\frac{D^4h^2\tau_0^2r_0^2}{4(1-\rho)^4}\left\{4\rho^2 Var[X] + \lambda^2 E[X^2]^2 + \frac{4}{Q}E[X]^2\right\} \tag{3.5.2}$$

Because this determinant is negative for all feasible parameter values, the Hessian is indefinite and the E[TC] function cannot be convex. Due to this lack of convexity, if a critical point occurred within the feasible region, it would not represent a local minimum of E[TC].

Therefore, it is concluded that a minimum expected total cost will always occur on the boundary of the feasible region of the problem. Furthermore, the non-convex nature of the E[TC] function precludes the use of convex programming techniques such as the Karush-Kuhn-Tucker method to determine analytically the optimal values of the decision variables [Peressini et al., 1988]. Because conventional optimization approaches are not applicable here, optima must be determined through a more fundamental approach which takes into account the special features of this problem. Such an approach is presented in the next section.

3.5.2 Investment Boundary Approach to Bivariate Optimization

The lack of convexity of the E[TC] function leads to the conclusion that optimal setup and quality levels must fall on the perimeter of the feasible region, shown in Figure 3.6. The first observation to be made from this conclusion is that the 'invest some in both' decision at the center region of the matrix in Figure 3.5 will never be optimal. For the 'some-some' decision to represent a minimum, a critical point within the feasible region would have to occur at a point of local convexity of the E[TC] function, and this cannot happen.

The second observation is that on the perimeter of the feasible region, one or both decision variables must assume their maximum or minimum limit, that is, 'no' investment or 'full' investment. This knowledge that one decision variable is forced to its limiting value will be exploited to delineate the different optimal decisions (i.e., where 'no' or 'full' investment is optimal) which occur, and to set a framework for finding the conditions which lead to these different decisions. This section develops this analysis.

In re-examining the feasible region for the decision variables shown in Figure 3.6, the 'left' and 'right' boundaries (i.e., setup time $\tau = 0$ and τ_0) represent the decisions to invest fully and to invest nothing in setup reduction, respectively. Similarly, the 'top' and 'bottom' boundaries (i.e., defect rate $r = r_0$ and 0) represent the decision to invest none and invest fully in quality improvement, respectively.

Four clearly defined investment decisions exist at the corners of the feasible region. Since the value of both decision variables is known at each of these points, the analysis will start at these points and will then be extended to the sides of the feasible region. The four cases may then be listed as:

Case 1: Invest Fully in Both (i.e., $T^* = 0, R^* = 0$)

Case 2: Invest None in Both (i.e., $T^* = 1, R^* = 1$)

Case 3: Invest Fully in Setup, None in Quality (i.e., $T^* = 0, R^* = 1$)

Case 4: Invest Fully in Quality, None in Setup (i.e., $T^* = 1, R^* = 0$)

As in the previous section, for a given set of problem parameters each of these cases will occur at critical values of the marginal cost to make improvements in each decision variable. For instance, consider for a moment cases 2 and 3. In both of these cases quality is not changed from its original level, so it is held constant as setup time is varied. Since setup time is the only variable in this instance, optimal setup time T^* can be determined by a similar process as was used to develop Equation 3.4.3. Leaving the quality level, R, unevaluated gives:

$$T_{min}^* = \frac{Q}{D\tau_0}\{1 - D(1 + Rr_0)E[X]\} + \frac{1}{\tau_0}\sqrt{\frac{haQ^2(1-\rho)}{D(2i - hDa)}} \tag{3.5.3}$$

Rearranging once again for τ_0/a^* gives:

$$\frac{\tau_0}{a^*} = \frac{hD\tau_0}{2i}\left[1 + \frac{1 + \lambda^2\left[(1 + Rr_0)E[X^2] - (1 + R^2r_0^2)QE[X]\right]}{\left\{1 - \lambda\left[T^*\tau_0 + (1 + Rr_0)QE[X]\right]\right\}^2}\right] \tag{3.5.4}$$

A similar process for quality level equations yields:

$$\frac{r_0}{b^*} = \frac{hDr_0QE[X]}{2i}\left[1 - \frac{2\lambda Rr_0E[X]}{\left\{1 - \lambda\left[T\tau_0 + (1 + Rr_0)QE[X]\right]\right\}} + \frac{1 - \lambda^2(1 + R^2r_0^2)QE[X]^2}{\left\{1 - \lambda\left[T\tau_0 + (1 + Rr_0)QE[X]\right]\right\}^2}\right]$$
$$+ \frac{hDr_0}{2i}\left[\frac{\lambda(Q - DT\tau_0)E[X^2]}{\left\{1 - \lambda\left[T\tau_0 + (1 + Rr_0)QE[X]\right]\right\}^2}\right] \tag{3.5.5}$$

A decision matrix analogous to that shown in Figure 3.5 will be generated by examining the four cases and the critical τ_0/a^*, r_0/b^* ratios in each case. At this point, a terminology will be introduced to represent these various critical values. Let:

$B^x{}_{y,z}$ ≡ critical investment cost to reduce a decision variable to zero

Where: x ≡ Factor (S ≡ setup, Q ≡ quality)

y ≡ Investment in that Factor

0 = No Investment Boundary

1 = Full Investment Boundary

z ≡ Investment in the other Factor

0 = No Investment

1 = Full Investment

Then the investment boundaries for the four cases become:

Case 1: Invest Fully in Both (i.e., T* = 0, R* = 0)

$$B_{1,1}^{S} = \frac{hD\tau_0}{2i}\left[1 + \frac{1 + \lambda^2\left[E[X^2] - QE[X]\right]}{\{1 - \lambda QE[X]\}^2}\right] \tag{3.5.6a}$$

$$B_{1,1}^{Q} = \frac{hDr_0QE[X]}{2i}\left[1 + \frac{1 - \lambda^2 QE[X]^2}{\{1 - \lambda QE[X]\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{\lambda QE[X^2]}{\{1 - \lambda QE[X]\}^2}\right] \tag{3.5.6b}$$

Case 2: Invest None in Both (i.e., T* =1, R* = 1)

$$B_{0,0}^{S} = \frac{hD\tau_0}{2i}\left[1 + \frac{1 + \lambda^2\left[(1 + r_0)E[X^2] - (1 + r_0^2)QE[X]\right]}{\{1 - \lambda[\tau_0 + (1 + r_0)QE[X]]\}^2}\right] \tag{3.5.7a}$$

$$\begin{aligned} B_{0,0}^{Q} = {} & \frac{hDr_0QE[X]}{2i}\left[1 - \frac{2\lambda r_0 E[X]}{\{1 - \lambda[\tau_0 + (1 + r_0)QE[X]]\}} + \frac{1 - \lambda^2(1 + r_0^2)QE[X]^2}{\{1 - \lambda[\tau_0 + (1 + r_0)QE[X]]\}^2}\right] \\ & + \frac{hDr_0}{2i}\left[\frac{\lambda(Q - D\tau_0)E[X^2]}{\{1 - \lambda[\tau_0 + (1 + r_0)QE[X]]\}^2}\right] \end{aligned} \tag{3.5.7b}$$

Case 3: Invest Fully in Setup, None in Quality (i.e., T* = 0, R* = 1)

$$B_{1,0}^{S} = \frac{hD\tau_0}{2i}\left[1 + \frac{1 + \lambda^2\left[(1 + r_0)E[X^2] - (1 + r_0^2)QE[X]\right]}{\{1 - \lambda(1 + r_0)QE[X]\}^2}\right] \tag{3.5.8a}$$

$$\begin{aligned} B_{0,1}^{Q} = {} & \frac{hDr_0QE[X]}{2i}\left[1 - \frac{2\lambda r_0 E[X]}{\{1 - \lambda(1 + r_0)QE[X]\}} + \frac{1 - \lambda^2(1 + r_0^2)QE[X]^2}{\{1 - \lambda(1 + r_0)QE[X]\}^2}\right] \\ & + \frac{hDr_0}{2i}\left[\frac{\lambda QE[X^2]}{\{1 - \lambda(1 + r_0)QE[X]\}^2}\right] \end{aligned} \tag{3.5.8b}$$

Case 4: Invest Fully in Quality, None in Setup (i.e., T* = 1, R* = 0)

$$B^{S}_{0,1} = \frac{hD\tau_0}{2i}\left[1 + \frac{1 + \lambda^2\left[E[X^2] - QE[X]\right]}{\left\{1 - \lambda\left[\tau_0 + QE[X]\right]\right\}^2}\right] \tag{3.5.9a}$$

$$B^{Q}_{1,0} = \frac{hDr_0QE[X]}{2i}\left[1 + \frac{1 - \lambda^2QE[X]^2}{\left\{1 - \lambda\left[\tau_0 + QE[X]\right]\right\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{\lambda(Q - D\tau_0)E[X^2]}{\left\{1 - \lambda\left[\tau_0 + QE[X]\right]\right\}^2}\right] \tag{3.5.9b}$$

These points have been drawn on the new decision matrix as shown in Figure 3.7.

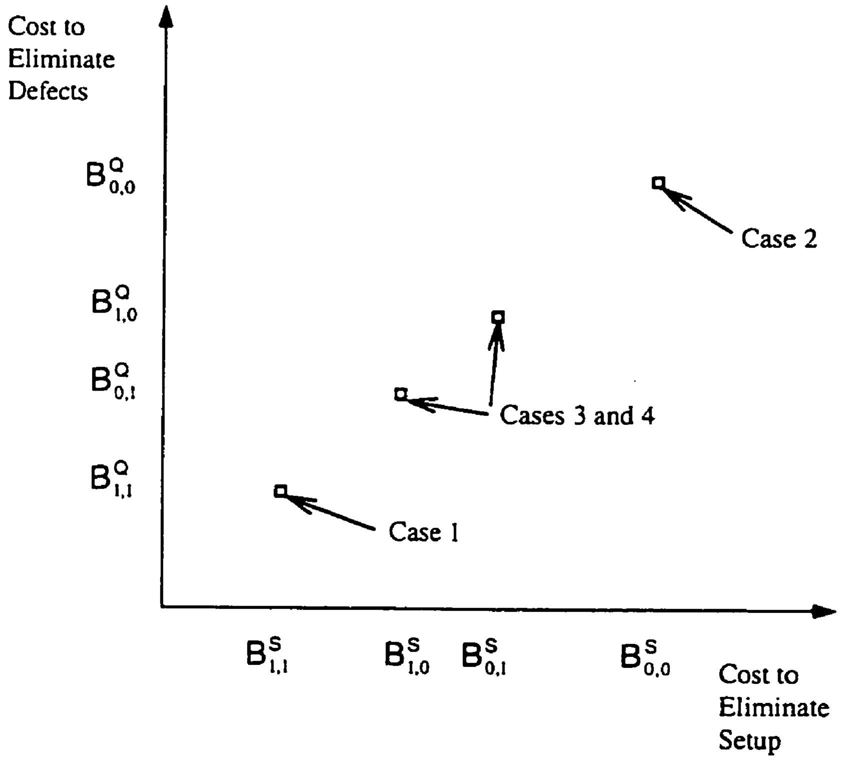


Figure 3.7: Initial Points on Decision Matrix

Suppose a candidate manufacturing system is evaluated and found to fall on the decision matrix in Figure 3.7 at the point called "Case 1". This case represents the optimal decision to invest fully in both setup and quality to completely improve them. Similarly, if the candidate system fell at "Case 2", the optimal decision would be to invest nothing in either improvement practice. Intuitively, it is expected that points falling below and to the left of "Case 1" should continue to result in an 'invest fully in both' decision, since the marginal cost of making improvements in each practice is decreasing in this direction. Similarly, points above and to the right of "Case 2" should continue to represent an 'invest none in both' decision since marginal costs are increasing. This intuitive result will be derived more formally below.

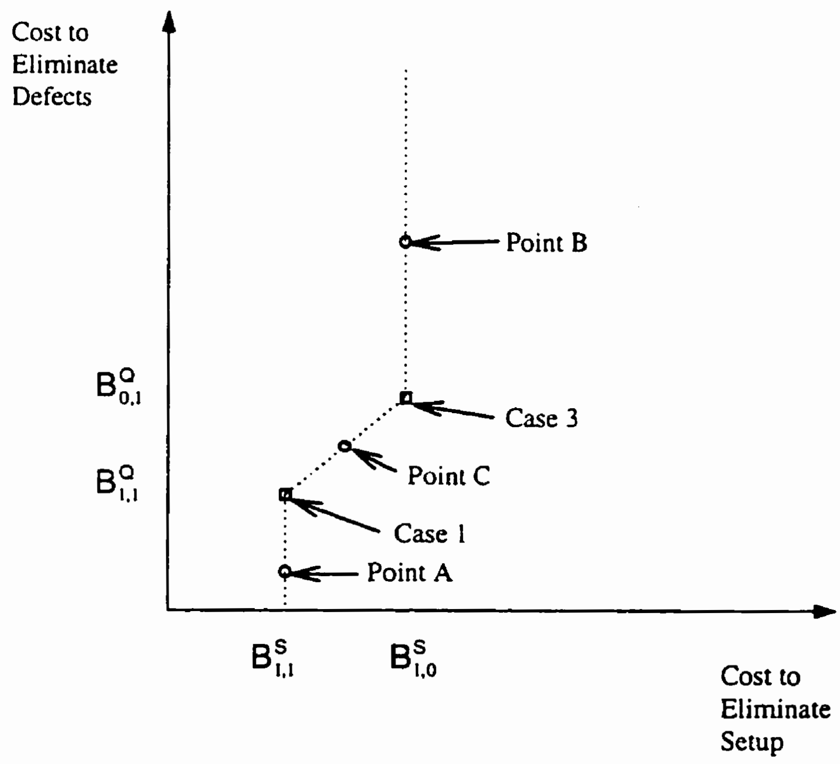


Figure 3.8: Development of Decision Boundary

Consider the decision boundary for the 'invest fully in setup' decision. This boundary must contain the points "Case 1" and "Case 3", which are shown in Figure 3.8. At "Case 1", the decision is to invest fully in quality (along with setup). Below the point "Case 1", the marginal cost of improving quality falls, so if the optimal decision at "Case 1" is to invest fully in quality, the optimal decision will still be to invest fully in quality below that point.

Suppose point 'A' falls on a line directly below "Case 1" as indicated in Figure 3.8. Being below "Case 1", the quality investment decision will not change along this line (i.e., it will continue to be to 'invest fully'). Since the level of investment in quality does not change along this line, and since on the vertical line the marginal cost of improving setup does not change, the decision boundary for 'full investment' in setup will lie along this line.

Similarly, point 'B' on a vertical line above "Case 3" can be considered. Upward along this line the marginal cost of improving quality increases, so the optimal decision for quality will continue to be to 'invest none' along this line. As the marginal cost of improving setup does not change on the vertical line, this line will represent another section of the decision boundary for 'full investment' in setup reduction.

Between "Case 1" and "Case 3", the decision boundary line can be plotted parametrically. At "Case 1" the optimal quality level, R*, is zero (i.e., all defects are eliminated) while at "Case 3" the optimal level becomes one. Point 'C' will represent the parametric point $(B^S_{1,R}, B^Q_{R,1})$ for $0 \leq R \leq 1$ where $B^S_{1,R}$ is from Equation 3.5.4 and $B^Q_{R,1}$ is from Equation 3.5.5 when $T = 1$. The line traced by point 'C' as R varies from zero to one then represents this segment of the 'full investment' decision boundary for setup reduction.

By considering the points where the optimal investment in setup is 'invest none' (i.e., "Case 2" and "Case 4"), this process can be repeated to develop the 'invest none' decision boundaries for setup reduction. The completed boundaries on setup decisions are shown in Figures 3.9a and 3.9b. Because the relative positions of the points for "Case 3" and "Case 4" may be reversed, two 'shapes' can result for the setup boundaries, which is why the boundaries may appear as in either figure.

Following an analogous procedure for quality will lead to the quality boundaries shown in Figure 3.10a (corresponding to the positions of the points "Case 3" and "Case 4" in Figure 3.9a) or Figure 3.10b (corresponding to Figure 3.9b).

By combining the decision matrices of Figures 3.9a and 3.10a, the bivariate decision matrix is obtained, and is shown in Figure 3.11. (Combining Figures 3.9b and 3.10b would lead to a similar matrix with the points "Case 3" and "Case 4" reversed.)

Figure 3.5 is also redrawn as Figure 3.12 in the same scale as Figure 3.11 in order to compare how decisions change for the cases where setup and quality are assumed to interact, and where they are considered independent of one another.

Two primary differences between the two systems (Figures 3.11 and 3.12) occur. The first is that when interactive effects between setup and quality improvements are considered, the boundaries separating investment decisions will tend to shift to points where the marginal cost of making improvements is less. For instance, a comparison of the 'invest fully in both' region (lower left corner) shows this region to be smaller in the case where interactions are considered. This means that full investment in both practices only takes place if the cost of making these improvements is less, when interactive effects are considered.

The second significant difference when interactions are considered is that the 'some-some' investment decision region disappears from the matrix. This is a consequence of the non-convexity of the E[TC] function (Equation 3.5.2). Optimal solutions must represent extreme points in the feasible region of the decision variables (shown in Figure 3.6), i.e., points which fall on the perimeter of the feasible region. Since no points inside the perimeter of the feasible region can represent optimal solutions, the decision matrix will not show a 'some-some' investment region.

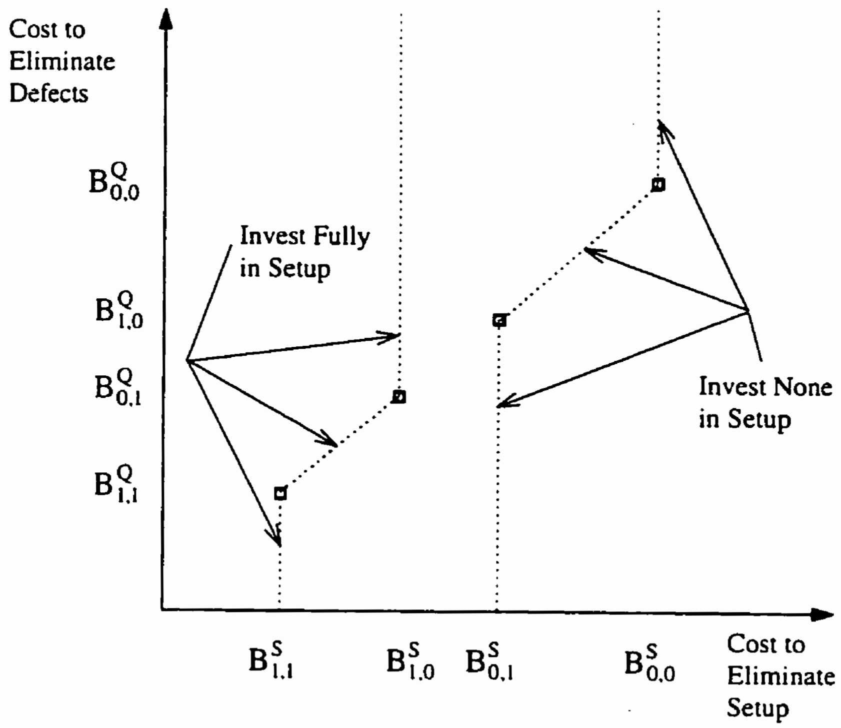


Figure 3.9a: Decision Boundaries for Setup Investment Alone

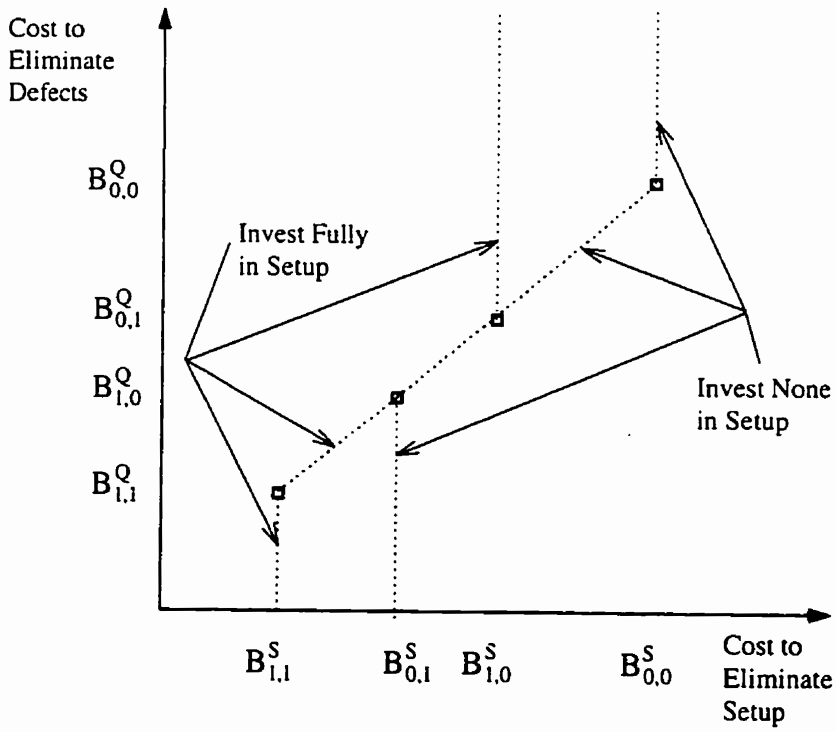


Figure 3.9b: Decision Boundaries for Setup Investment Alone (alternate configuration)

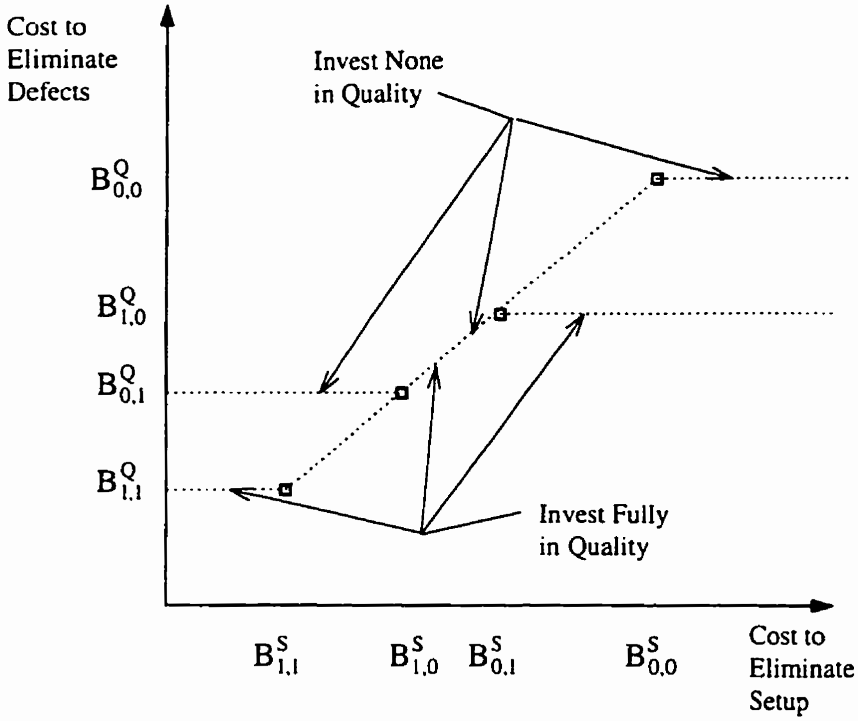


Figure 3.10a: Decision Boundaries for Quality Investment Alone)

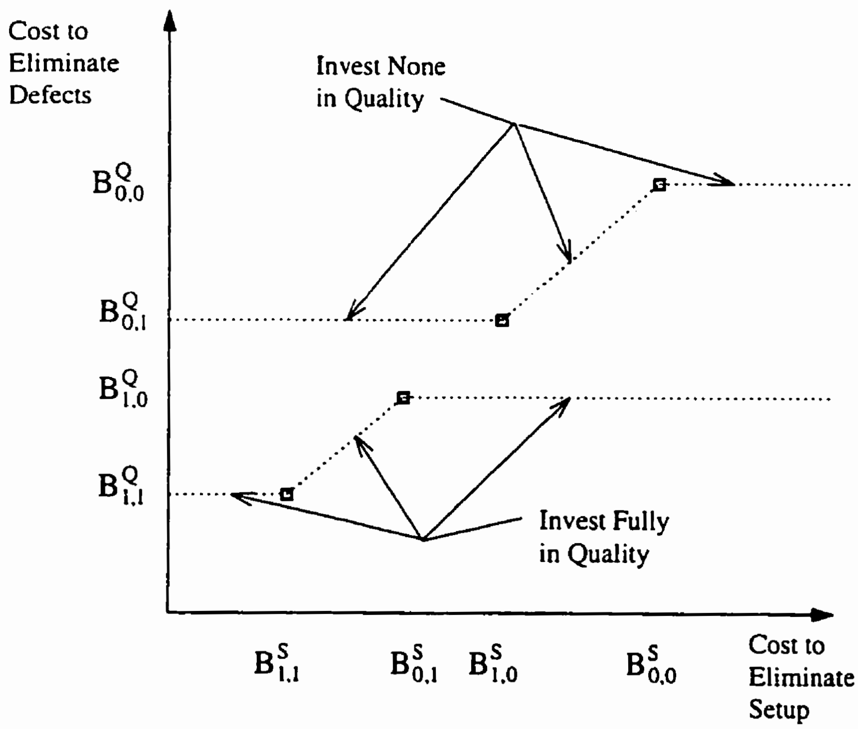


Figure 3.10b: Decision Boundaries for Quality Investment Alone (alternate configuration)

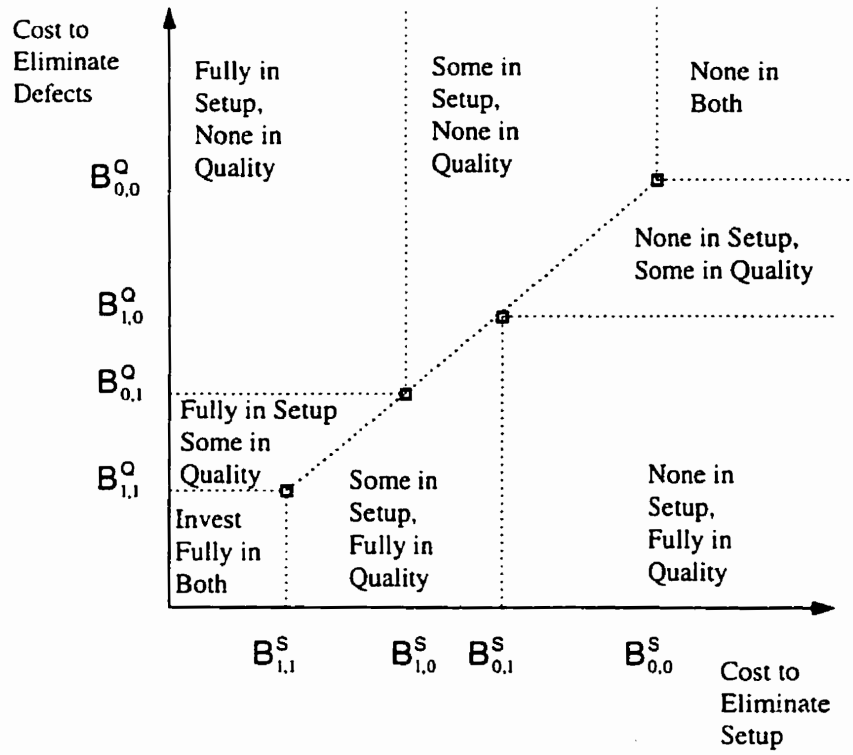


Figure 3.11: Decision Boundaries for Bivariate Setup and Quality Investment.

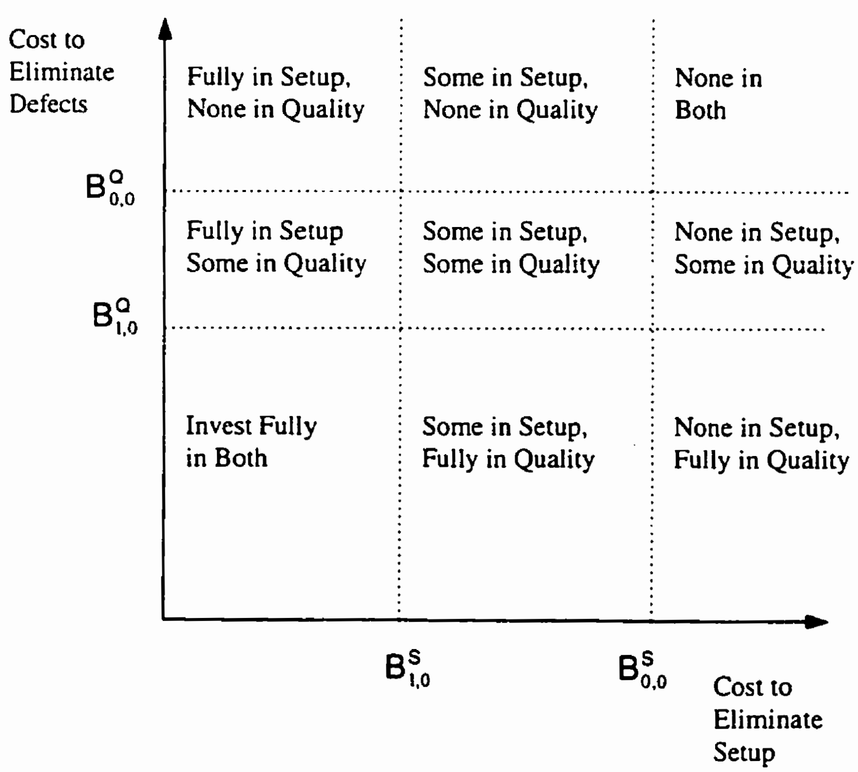


Figure 3.12: Decision Boundaries for Setup and Quality Investment, Neglecting Interactions.

The lack of a 'some-some' optimum investment region is a result of using a linear (i.e., constant returns to scale) investment-improvement function, as will be demonstrated in Chapter 5. In that chapter a decreasing returns to scale investment function is studied and a large 'some-some' optimum decision region is found on the decision matrix.

Returning to the example problem in which a quality and a manufacturing engineer were considering improvement projects in isolation of each other, these decision matrices show that since the regions of the 'independent' decision matrix are larger, the engineers are more likely to invest more money than if interactive behaviour is considered, and are less likely to realize the expected level of benefits.

These results show that setup reduction and quality improvement practices interact when they are implemented simultaneously on a manufacturing system. A result of the existence of this interaction is that over-investment can take place if the interaction is not taken into account. The next section will examine the factors which influence the potential of making an erroneous decision, and the magnitude of the potential over-investment.

3.6 Magnitude of Decision Error

In re-examining the example investment problem, the last section showed how decisions can be made erroneously, leading to investment of more resources than are justified. In this section, the factors which affect the magnitude of this potential error will be examined.

3.6.1 Analytic Examination of Decision Error

The factors affecting how much the decision shifts can be examined analytically in this case. Consider for the moment the setup decision boundaries as shown in Figure 3.13. The term 'Decision Error' will be loosely defined here as the prospect of coming to an erroneous investment decision when neglecting potential interactions. One way of examining Decision Error is to look at the amount of shift of the investment boundaries in the 'informed' case relative to the 'naive' case. The greater the shift in boundaries, the larger the regions where non-optimal investments take place and the greater the prospect of Decision Error occurring in investment decisions.

The shift in boundaries can be examined mathematically. If a term 'Boundary Shift' is coined, it can be defined as the relative amount that each boundary shifts from the 'naive' case to the most shifted value in the 'informed' case. This function then provides an indication (from 0 to 1) of how much the boundary has shifted as a result of considering interactions, with higher values reflecting a greater shift and hence a greater likelihood of a decision maker reaching an incorrect investment decision by ignoring interactive effects.

Following this definition, Boundary Shift (for the setup case) can be expressed mathematically as:

$$\text{Boundary Shift} = \frac{B^S_{0,0} - B^S_{0,1}}{B^S_{0,0}} \text{ (no-invest boundary), or } \frac{B^S_{1,0} - B^S_{1,1}}{B^S_{1,0}} \text{ (full-invest boundary)}$$

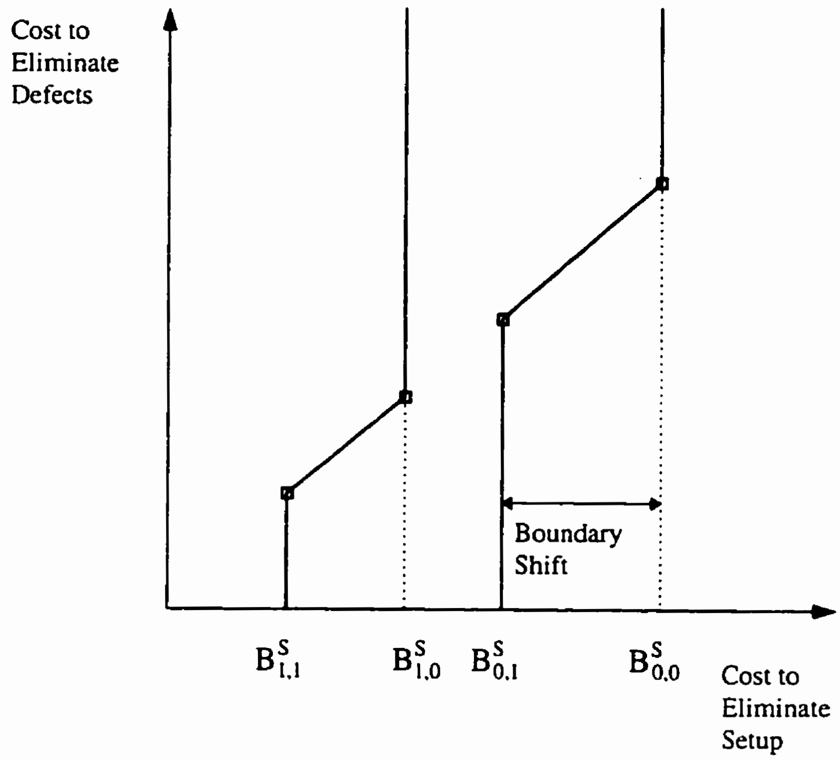


Figure 3.13: Boundary Shift between 'Naive' and 'Informed' Boundaries.

By returning to Equations 3.5.6 through 3.5.9, this Boundary Shift term can be determined analytically. For instance, for the setup time decision variable, full investment boundary,

$$\text{Boundary Shift} = \frac{B^S_{1,0} - B^S_{1,1}}{B^S_{1,0}}$$

$$= 1 - \frac{\{1 - \lambda(1 + r_0)QE[X]\}^2}{\{1 - \lambda QE[X]\}^2}$$

$$\times \left[\frac{\{1 - \lambda QE[X]\}^2 + 1 + \lambda^2 \left[E[X^2] - QE[X]\right]}{\{1 - \lambda(1 + r_0)QE[X]\}^2 + 1 + \lambda^2 \left[(1 + r_0)E[X^2] - (1 + r_0^2)QE[X]\right]} \right]$$

(3.6.1)

The sensitivity of the Decision Error to each parameter in the problem can be determined by checking the sign of the partial derivatives. A summary of the results for setup time Boundary Shift is given in Table 3.1. Increasing each of the parameters increases the Decision Error, with the exception of Batch Size, Q. (This result is intuitively reasonable since increasing each of the parameters other than Batch Size increases system utilization, while increased Batch Size leads to fewer setups and decreased system utilization, as will be shown next.)

Furthermore, the sensitivity of the Decision Error to the system utilization, ρ, can be determined. Because of the way the variable Boundary Shift was defined, it cannot be written specifically in terms of utilization. However, the total derivative of Boundary Shift with respect to utilization can be expressed as in Equation 3.6.2:

$$\frac{dBS}{d\rho} = \frac{\partial BS}{\partial D}\frac{\partial D}{\partial \rho} + \frac{\partial BS}{\partial Q}\frac{\partial Q}{\partial \rho} + \frac{\partial BS}{\partial \tau_0}\frac{\partial \tau_0}{\partial \rho} + \frac{\partial BS}{\partial r_0}\frac{\partial r_0}{\partial \rho} + \frac{\partial BS}{\partial E[X]}\frac{\partial E[X]}{\partial \rho} \qquad (3.6.2)$$

The various partial derivatives with respect to utilization can be evaluated from the definition of utilization:

$$\rho = {}^{D}\!/_{Q}\,\{T\tau_0 + (1 + Rr_0)QE[X]\}$$

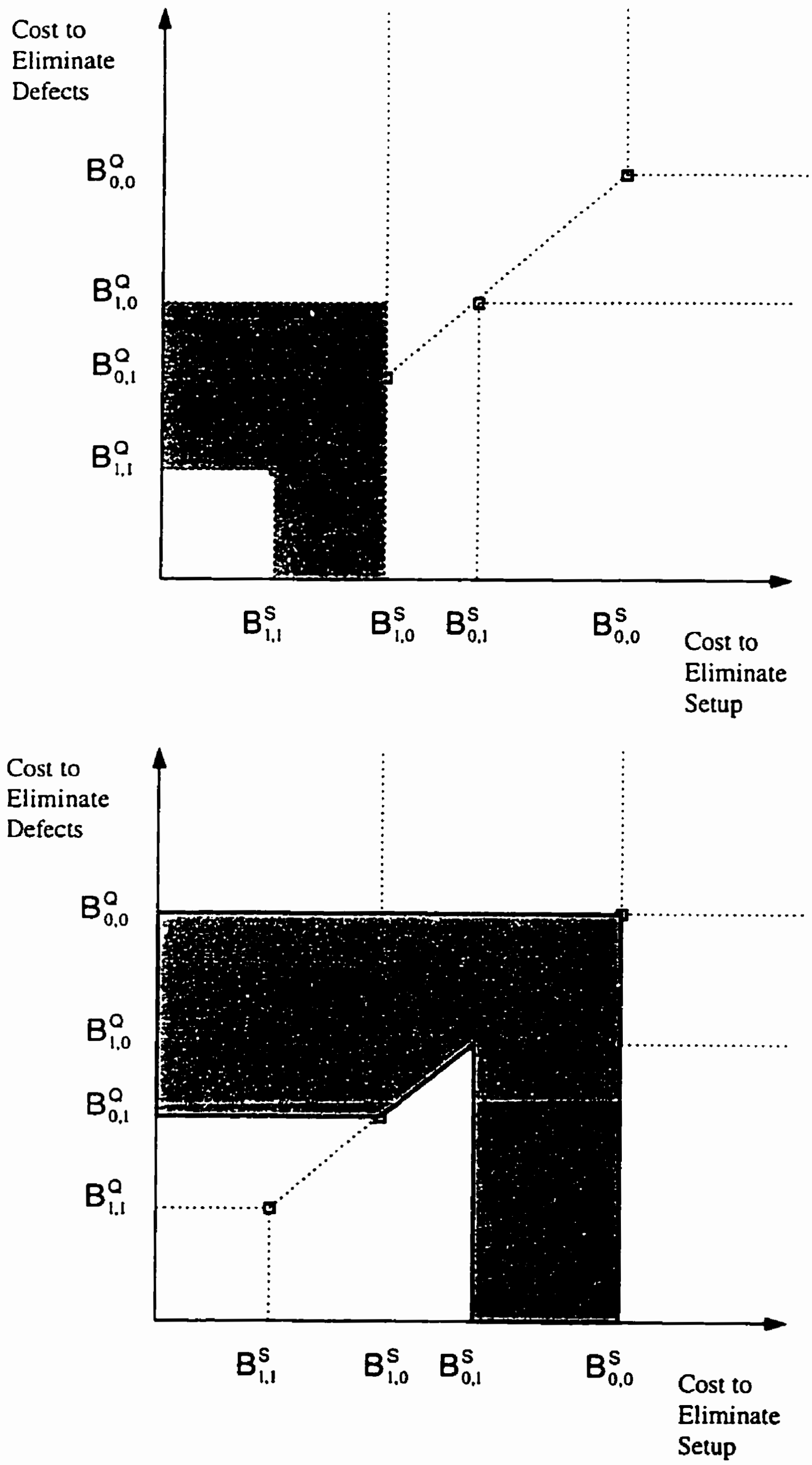
Table 3.1: Partial Derivatives for Sensitivity of Setup Time Decision Error

Variable	Boundary Shift Derivative	Utilization Derivative	Product Term
Higher Demand (D)	$\frac{\partial BS(T)}{\partial D} > 0$	$\frac{\partial D}{\partial \rho} = \frac{Q}{T\tau_0 + (1 + Rr_0)QE[X]} > 0$	$\frac{\partial BS(T)}{\partial D}\frac{\partial D}{\partial \rho} > 0$
Larger Lot size (Q)	$\frac{\partial BS(T)}{\partial Q} < 0$	$\frac{\partial Q}{\partial \rho} = \frac{-Q^2}{DT\tau_0} < 0$	$\frac{\partial BS(T)}{\partial Q}\frac{\partial Q}{\partial \rho} > 0$
Higher Mean Service Time (E[X])	$\frac{\partial BS(T)}{\partial E[X]} > 0$	$\frac{\partial E[X]}{\partial \rho} = \frac{1}{D(1 + Rr_0)} > 0$	$\frac{\partial BS(T)}{\partial E[X]}\frac{\partial E[X]}{\partial \rho} > 0$
Higher Initial Defect Rate (r_0)	$\frac{\partial BS(T)}{\partial r_0} > 0$	$\frac{\partial r_0}{\partial \rho} = \frac{1}{DE[X]R} > 0$	$\frac{\partial BS(T)}{\partial r_0}\frac{\partial r_0}{\partial \rho} > 0$
Higher Initial Setup Time (τ_0)	$\frac{\partial BS(T)}{\partial \tau_0} > 0$	$\frac{\partial \tau_0}{\partial \rho} = \frac{Q}{DT} > 0$	$\frac{\partial BS(T)}{\partial \tau_0}\frac{\partial \tau_0}{\partial \rho} > 0$

Table 3.1 also summarizes the signs of each of the derivatives in Equation 3.6.2, as well as the signs of the product terms. Since each of the product terms is positive, the total derivative of Boundary Shift with respect to system utilization is positive, indicating that decision makers will be more prone to making excess investments in improvements to a system as the system is more heavily utilized.

The concept of Decision Error can be shown graphically as well. Refer again to the decision matrices shown in Figure 3.11 and Figure 3.12. Considering the setup reduction decision boundaries alone for the moment, the 'naive' (ignoring interaction; Figure 3.12) and 'informed' (including interaction; Figure 3.11) decision boundaries coincide for higher levels of cost to eliminate defects (top portion of chart), while the 'informed' boundaries shift to the left as the cost to eliminate defects drops. This is because as the cost drops, investing in quality improvement provides benefits to the system (i.e., lower utilization of the server resulting in lower WIP costs) at a lower cost than investing in setup reduction, and so becomes the preferred investment. Since some gains have already been achieved, there is less potential benefit from investment in setup reduction. The boundaries shift to the left, reflecting the greater difficulty justifying investment in that factor.

Figures 3.14a and 3.14b show the differences between these 'naive' vs. 'informed' boundaries for the cases of ‘No Investment’ and ‘Full Investment’ decisions. As can be seen, the shifting of the 'informed' boundaries relative to the 'naive' boundaries creates significant regions in which the investment decision changes from "Invest Some" to "Invest None" and "Invest Fully" to "Invest Some". In each case the 'informed' decision results in less investment than the 'naive' decision, and hence, lower expected total operating costs for the system. Shaded areas in Figures 3.14a and b represent the regions where optimal investment drops. This behaviour is a general result for this system while the degree of change to the boundaries will depend on the parameter values in a particular system.



Figures 3.14a and 3.14b: Changes in "Full Investment" and "No Investment" Boundaries.

(The shaded areas represent regions of over-investment when interactions are ignored.)

3.6.2 Sensitivity Analysis

To help understand how the various parameters in this model affect this concept of Decision Error, a numerical sensitivity analysis has been performed. In this section the sensitivity analysis is given for the case of a linear investment function and fixed batch size; subsequent chapters will give similar analyses for cases with convex investment functions and variable batch sizes.

The parameter values for the base case and deviations are given in Table 3.2. These values are adapted from example values used by Porteus [1986b], Spence and Porteus [1987] and Kim et al. [1992]. The nominal levels (base case) result in a system utilization of 0.93. Since Demand, Batch Size, Defect Rate and Setup Time affect system utilization, the upper and lower levels of these parameters have been selected such that the high and low values change system utilization by 0.05 above and below the base utilization (e.g., from 0.88 to 0.98). The other variables studied, Holding Cost Rate, Interest Rate and Processing Time Coefficient of Variation, do not affect system utilization.

Table 3.2: Parameter Levels for Sensitivity Analysis

Parameter		Nominal	Upper	Lower
Demand	$D =$	100000	105000	95000
Batch Size	$Q =$	1000	650	2000
Defect Rate	$r_0 =$	0.24	0.32	0.16
Setup Time	$\tau_0 =$	0.001	0.0015	0.000
Holding Cost Rate	$h =$	10	100	1
Interest Rate	$i =$	0.25	0.5	0.01
Processing Time Coeff. of Variation	$cv =$	0.5	1	0
Expected Service Time	$E[X] =$	6.67E-06	(fixed)	

A definition of over-investment is created as:

$$\text{Over-Investment}\left(\frac{\tau_0}{a},\frac{r_0}{b}\right)=\frac{\text{'Naive' investment} - \text{'Informed' investment}}{\text{'Informed' investment}}$$

With this definition, the optimal level of investment is determined under both the 'naive' (univariate) and 'informed' (bivariate) assumptions for a given combination of setup elimination cost (τ_0/a) and defect elimination cost (r_0/b). If the optimal investment level in both cases coincides, over-investment is zero. If optimal investment in the 'naive' case is greater than that of the 'informed' case, over-investment is positive. It should be noted that since a ratio is involved, small optimal levels of 'informed' investment can lead to very high relative levels of over-investment, while the absolute levels of over-investment may be more modest.

As was shown through Equation 3.6.2 and Table 3.1, investments in one practice lead to 'informed' investments in the other practice equal to or less than 'naive' investments in that practice (i.e., equal investments or over-investment; never under-investment). With this result, investment level comparisons need only be made in regions of the decision matrix where the 'naive' strategy calls for investment in both practices. Any area where no investment is called for in one practice under the 'naive' strategy will also call for no investment in that practice under the 'informed' strategy. Such areas will lead to equal investment under the two strategies.

To examine the over-investment behaviour of this system, optimal investment levels for both strategies were calculated at each point in a grid covering the regions of the decision matrix where investment in both practices was called for in the 'naive' strategy. Numerical investigation during calculation of over-investment frequencies showed that the optimal investment levels were well behaved over the studied space, varying smoothly from no investment to full investment as a function of the costs to improve each practice. Consequently, optimal investment levels were calculated at each of 51 equally spaced values for the cost to eliminate setup and the cost to eliminate defects, ranging from a value of zero to the value representing the "no-invest" boundary for that practice in the 'naive' case. This led to a grid of 2601 (= 51^2) points covering the decision matrix where the 'naive' strategy called for investment.

Frequencies of these over-investment levels can be tabulated, such as given in Table 3.3. It is seen that the magnitudes of over-investment can be significant.

By sequentially changing each parameter to the upper and lower levels given in Table 3.2, similar frequencies of over-investment can be calculated to examine the sensitivity of the system to each parameter. Such calculations have been performed and the results are given graphically in Figures 3.15 to 3.21.

Table 3.3: Frequency of Over-Investments

Over-Investment Level (percentage of optimal investment)	Frequency of Observances (n = 2601)
< 5%	19%
5 - 50%	25%
50 - 100%	38%
100 - 150%	7%
150 - 200%	2%
200 - 250%	1%
250 - 300%	1%
300 - 350%	0%
350 - 400%	0%
400 - 450%	1%

The first four graphs are for the parameters which affect system utilization (i.e., demand, batch size, defect rate and setup time). As each parameter goes from the low-utilization level (i.e., low demand, high batch size, low defect rate and low setup time) to the high-utilization

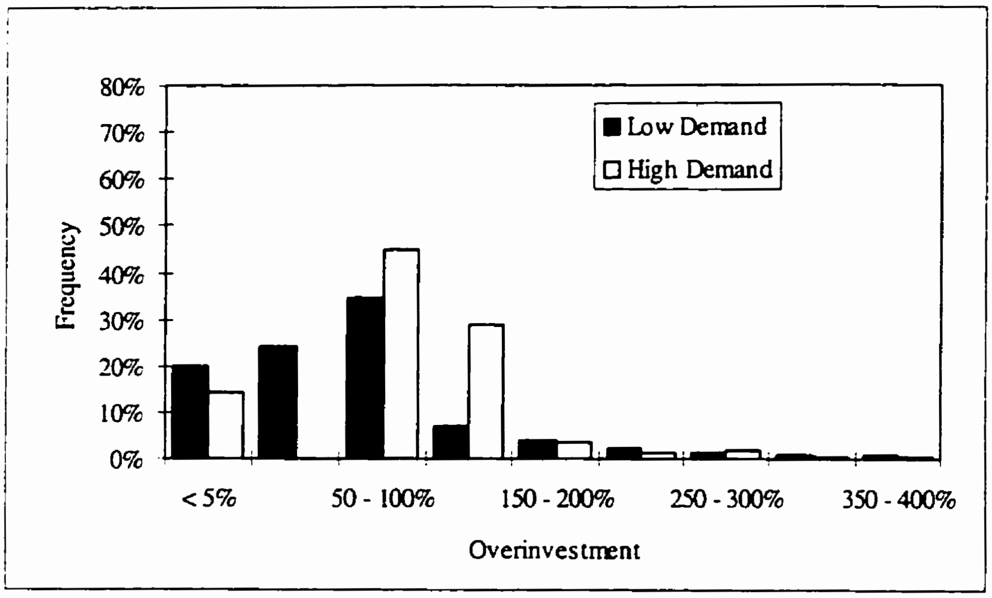


Figure 3.15: Sensitivity of Decision Error to Demand Levels

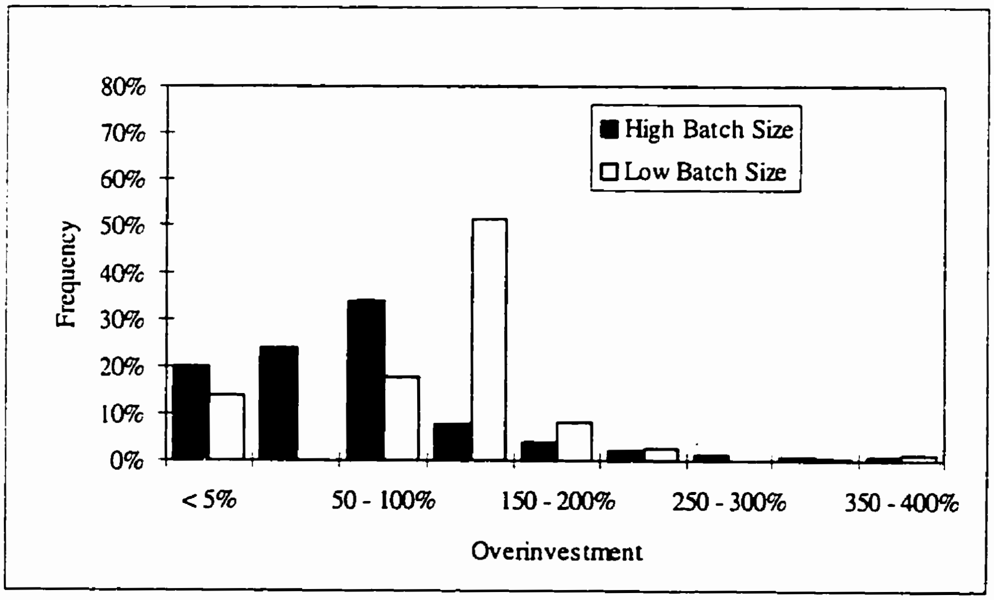


Figure 3.16: Sensitivity of Decision Error to Batch Size

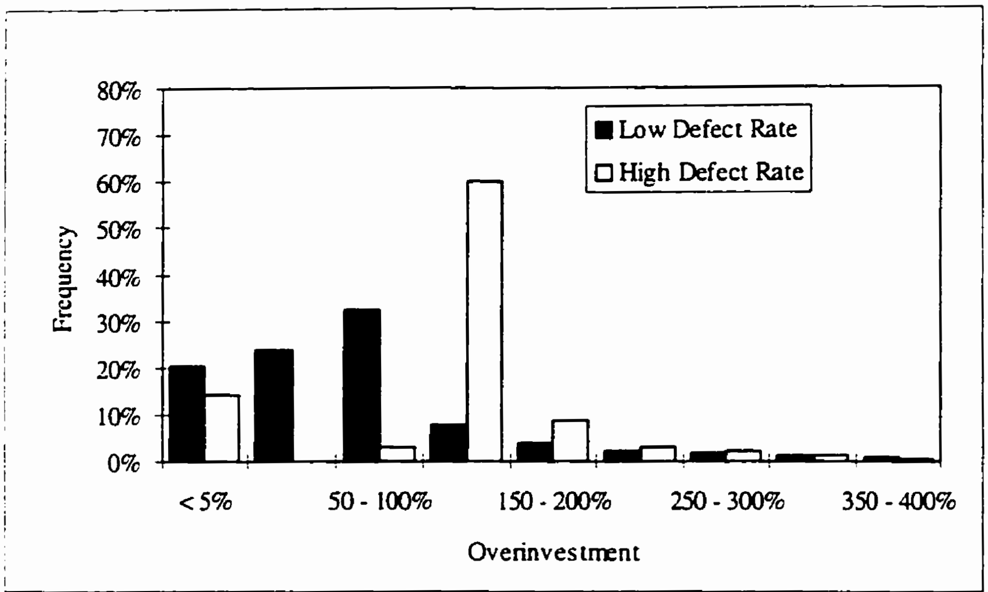


Figure 3.17: Sensitivity of Decision Error to Defect Rate

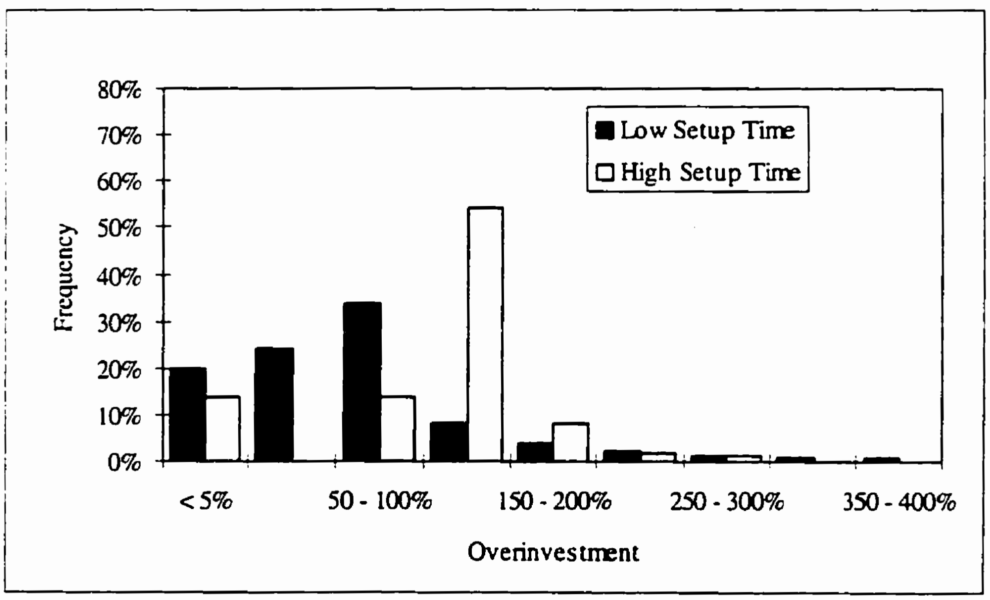


Figure 3.18: Sensitivity of Decision Error to Setup Time

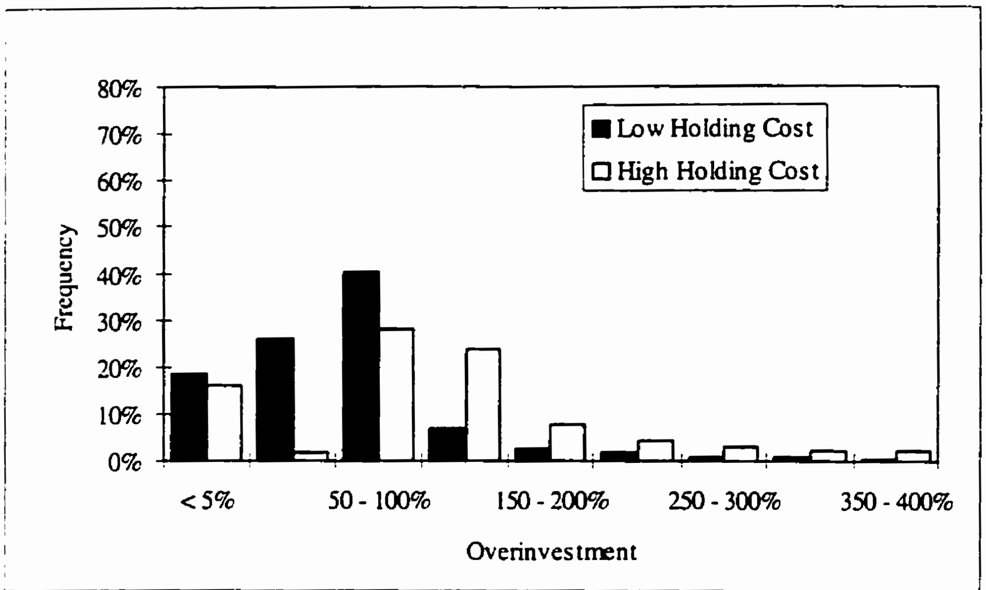


Figure 3.19: Sensitivity of Decision Error to Holding Cost Rate

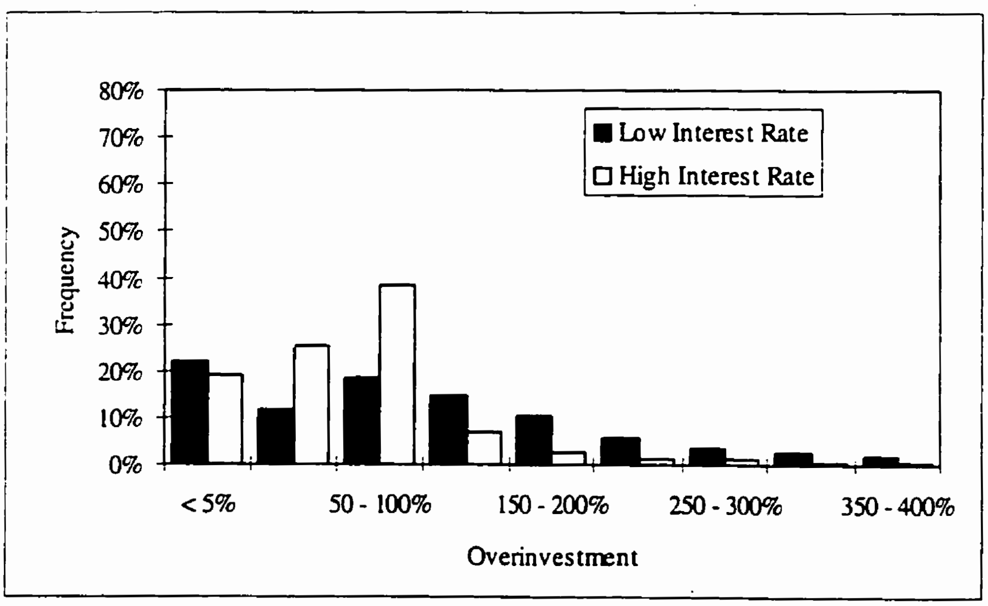


Figure 3.20: Sensitivity of Decision Error to Interest Rate

level, the over-investment frequency distributions shift to the right, indicating that the mean over-investment level increases. This result corresponds to the analytic result described in Equation 3.6.2, that is, the decision error increases as system utilization increases.

Figures 3.19 and 3.20 show the effects of changes to the holding cost rate and the interest rate. It can be seen that the mean over-investment shifts in opposite directions as the parameters go from low to high levels. This is expected since increased holding costs would lead to higher investment levels (since greater benefits are obtained from each unit of inventory reduction) while increased interest rates would lead to lower investment levels (due to higher costs for each unit of improvement).

The direction of the shift in mean over-investment can also be explained. Referring to Equations 3.5.6a through 3.5.9b, it can be seen analytically that the location of each decision boundary is proportional to the ratio h/i (i.e., the ratio of holding cost rate to interest rate). As either holding cost rate decreases or interest rate increases, the location of the decision boundaries decrease in proportion. This decrease shifts the regions of over-investment to areas of lower costs of making improvements. Similarly, as the boundaries shift towards the origin, the area of the regions where over-investment takes place decreases. These two trends tend to lower the levels and likelihood of over-investment and thus reduce the mean level of over-investment, as shown in Figures 3.19 and 3.20.

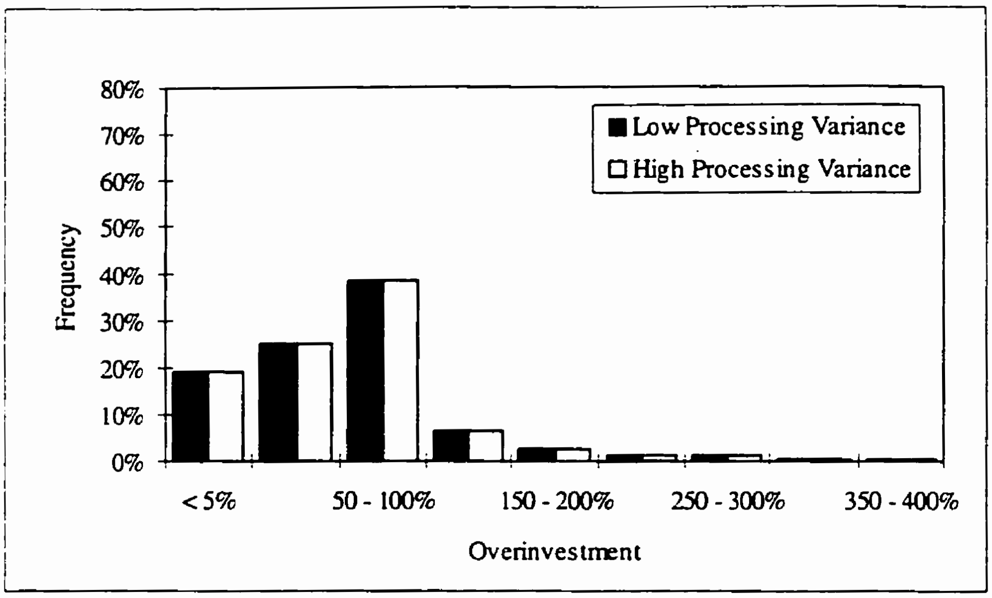


Figure 3.21: Sensitivity of Decision Error to Processing Time Variance

The effect of changing the coefficient of variation (cv) of the unit processing time (X) on over-investment was also examined (Figure 3.21). Essentially, no difference is seen in the over-investment frequencies. This can be explained by considering the definition of cv of batch service time (neglecting setup time and rework for the moment):

$$cv[S] = \frac{\sqrt{Var[S]}}{E[S]} \tag{3.6.3}$$

Where $Var[S] = Q\,Var[X]$

$E[S] = Q\,E[X]$

Q = batch size

then,

$$\begin{aligned} cv[S] &= \frac{\sqrt{Q\ Var[X]}}{Q\ E[X]} \\ &= \frac{\sqrt{Var[X]}}{\sqrt{Q}\ E[X]} \\ &= \frac{1}{\sqrt{Q}} cv[X] \end{aligned} \tag{3.6.4}$$

Now, consider the second moment of the batch service time, $E[S^2]$,

$$E[S^2] = Var[S] + E[S]^2$$

and, rearranging 3.6.3 for Var[S],

$$Var[S] = E[S]^2 cv[S]^2$$

giving,

$$E[S^2] = E[S]^2 (1 + cv[S]^2)$$

and substituting 3.6.4 yields,

$$E[S^2] = E[S]^2 \left(1 + \frac{1}{Q} cv[X]^2 \right) \tag{3.6.5}$$

This equation expresses the second moment of batch service time as a function of batch size (Q) and unit processing time coefficient of variation (cv[X]). By substituting it into the expected WIP cost (from Equation 3.3.6), we obtain the following:

$$E[\text{WIP Cost}] = hD\left\{\frac{D/_Q\, E[S]^2\left(1 + \frac{1}{Q}\,cv[X]^2\right)}{2\left(1 - D/_Q\, E[S]\right)} + E[S]\right\} \tag{3.6.6}$$

The effect of changes to cv[X] on expected WIP costs might be demonstrated by considering a numerical example. Suppose that Q is selected to be 1000 (the nominal example value from Table 3.2). Also, suppose that cv[X] changes from 1 to 0, e.g., such as a case in which processing time changes from being exponentially distributed to being deterministic. The quantity in the parentheses in the numerator of the first term of 3.6.6, i.e.,

$$1 + \frac{1}{Q}\,cv[X]^2$$

will change from a value of 1.001 to 1 as the variability of processing times is changed over this quite wide range.

Because the second term of 3.6.6 is not affected by changes to cv[X], the expected WIP holding cost will change by less than 0.1% in this numerical example. While optimal investment decisions will change as costs change, as expected costs are so insensitive to changes in unit processing time coefficient of variation when batch sizes are relatively large, investment decisions will also be insensitive to changes in this parameter.

This result for batch service time coefficient of variation is suspected to be less that that seen in practice, but this difference is not expected to affect the general conclusions drawn from the model. To illustrate this claim, suppose that batch service time variance was reduced to its limit, i.e., to zero. Service time would then be deterministic, and the second moment, $E[S^2]$, would equal $E[S]^2$. Expected waiting time in the system, E[W], from Equation 3.2.4 would then be:

$$
\begin{aligned}
E[W] &= \frac{\lambda E[S]^2}{2(1-\rho)} + E[S] \\
&= \frac{\rho}{1-\rho}\frac{E[S]}{2} + E[S] \\
&> 0
\end{aligned}
$$

Clearly, queuing behaviour would still occur, and be strongly affected by system utilization. Queue length, and hence the level of WIP holding costs available to justify investment in improvements, will decrease as batch service time variance decreases, but the relationships between system utilization and setup and quality improvements will be qualitatively the same. Therefore, even if the service time model used in this research does not adequately capture the processing time variances typically found in manufacturing systems, the insights regarding investment decision behaviour discovered in this research are still expected to be valid.

In summary, this section has used the linear investment function/fixed batch size model of a manufacturing cell to show that ignoring interactions can lead to significant levels of over-investment. It was shown that the factors which affect system utilization (namely, demand, batch size, setup time and defect rate) also affect the potential decision error, with increased decision error associated with increased utilization. A numerical sensitivity analysis has supported these results, as well as showing that mean over-investment is also affected by the ratio of holding cost rate to interest rate. Over-investment, however, appears to be essentially insensitive to the processing time coefficient of variation.

3.7 Discussion

In the previous sections a mathematical model was developed to represent the process of making simultaneous setup and quality improvements to a manufacturing system, and to predict how those changes affect the operating cost of the system. This model is characterized by having an M/G/1 queue to estimate WIP levels in the system, linear investment-improvement functions for setup and quality, and investment costs amortized through the use of the Capital Cost formula.

This model could not be optimized analytically through conventional methods to determine simultaneously the best level of the two decision variables due to the non-convexity of the expected total cost function. It was concluded that no local optima exist within the feasible region for the decision variables, so each optimum is characterized by falling on the boundary of the feasible decision space. However, by studying the points where investment decisions changed, the optimal improvement decision strategy for this system was determined.

To determine the optimal improvement strategy, the critical points where discontinuities exist in optimal investment levels were found. These discontinuities represent the 'boundaries' on investment decisions (e.g., such as where the optimal investment decision changes from investing 'some' to investing 'none'). By examining the difference in the location of these decision boundaries between the case where investments in the two decision variables were assumed independent and where possible interactions were considered, it was determined that setup and quality improvements interact. In terms of investment strategy, it was shown that this interaction can produce a significant chance of over-investing, and the magnitude of the over-investment can be dramatic. Neglecting the effect of the interaction always results in a level of investment equal to or greater than that when the interaction is not considered.

It was also shown that the factors which increase the utilization of the server also increase the risk of over-investing in improvement practices. Thus, on more heavily utilized systems there is a greater risk of investing more than is justified if interactions are not considered.

A numerical sensitivity analysis found that the mean level of over-investment increased with increases in product demand, initial setup time, initial defect rate, and interest rate, increased with decreases in batch size and inventory holding cost rate, and was insensitive to changes in processing time variance.

A non-intuitive result, that a 'some-some' investment is never optimal, was also found. As will be shown in Chapter 5, this result is due to the use of the linear investment function, which led to a non-convex total cost function.

Earlier literature [e.g., Hall, 1983; Schonberger, 1982] recognized empirically that optimal batch sizes decrease as improvements are made and that batch size reduction is an important variable in the reduction of WIP inventories. With the advent of EOQ-based models, Porteus [1986b] also showed optimal batch size to be affected strongly by the setup and quality levels of the system. While EOQ-based models neglect WIP costs, Karmarkar [1987] showed that optimal batch size similarly decreases as system utilization drops (as in the case where setup and quality improvements are made) in models including WIP inventories. In this Chapter, batch size was modeled as a fixed parameter, while empirical and analytical evidence suggests that optimal batch sizes change significantly as system improvements are made. Chapter 4 investigates the behaviour of the system when batch size is a decision variable.

Chapter 4: Variable Batch Size, Linear Investment Function

In this chapter, the expected total cost (E[TC]) model of the last chapter (Equation 3.3.6) has been extended to study the case in which batch size is a decision variable rather than a parameter. This produces a system in three decision variables: setup time (T), defect rate (R) and batch size (Q). The linear investment-improvement function is used in this chapter while systems with convex investment functions are studied in subsequent chapters.

Reductions in batch size have been linked to reductions in setup times in both empirical studies [e.g., Schonberger, 1982; Hall, 1983] and analytical analyses [e.g., Porteus, 1986b]. Karmarkar [1987] has also demonstrated that WIP level, and hence WIP costs, in a manufacturing system are very dependent on batch size. Since the model developed in this study balances savings in WIP costs against investments in improvements to the system, batch size is expected to be an important variable in minimizing total system costs. For this reason, the effect of variable batch size on optimal investment decisions is studied in this chapter.

As a variable, batch size is inherently different than setup time or quality level. Changing setup time or quality level requires an investment of resources, while changes to batch size is assumed to require no investment. Because of the empirical evidence discussed in Chapter 2 linking setup time and batch size reductions, these two variables are optimized together. As was done in the last chapter, two specific cases are examined: the 'naive' investment decision case where investments in setup and quality are assumed to be independent of one another, and the 'informed' investment decision case where investments in both practices are considered simultaneously. Differences between optimal decisions in the two cases indicate the existence of interactions between the two improvement practices.

In the next section, the optimal solution for the 'naive' case is developed, followed by the solution for the 'informed' case in Section 4.2. Section 4.3 examines the decision error in this model version, as well as providing a sensitivity analysis. Finally, the results from this model are discussed in Section 4.4.

4.1 'Naive' Investment Decisions

Just as in the analysis of the previous chapter, in the 'naive' case the two improvement practices, setup and quality improvement, are assumed to be independent of one another in terms of their effects on both the physical and economic systems. Optimal investment decisions are found by determining the strategies for each variable which minimize expected total costs of the system.

By linking batch size optimization to setup time optimization, two types of independent investment projects are considered: optimal investment in setup time reduction, with batch size as a secondary optimization variable, and, optimal investment in quality improvement. Each of these cases is addressed in the following sub-sections.

4.1.1 Setup Time and Batch Size Optimization

In this case, the expected total cost is assumed to be a function of two decision variables, T and Q, only. Quality level, R, is treated as a parameter with its value set to the initial defect rate for the system (i.e., R = 1).

Before looking for critical points, the Hessian of the E[TC] of the system can be examined to determine if the total cost function is convex in the feasible region of the decision variables. The determinant of the Hessian of E[TC] is:

$$\det H\big(E[TC](T,Q)\big) = -\frac{\lambda^4 D^2 h^2 \tau_0^2 \left[(1 + Rr_0)E[X^2] - (1 + R^2 r_0^2)E[X]^2\right]^2}{(1-\rho)^4} \tag{4.1.1}$$
$$\leq 0$$

This result shows that the Hessian is indefinite under all circumstances so that the E[TC] function is always non-convex. If any critical points exist in the system, they cannot represent local minima. Because of this result, the global minimum of the constrained optimization problem must fall on the boundary of the feasible region (shown in Figure 4.1).

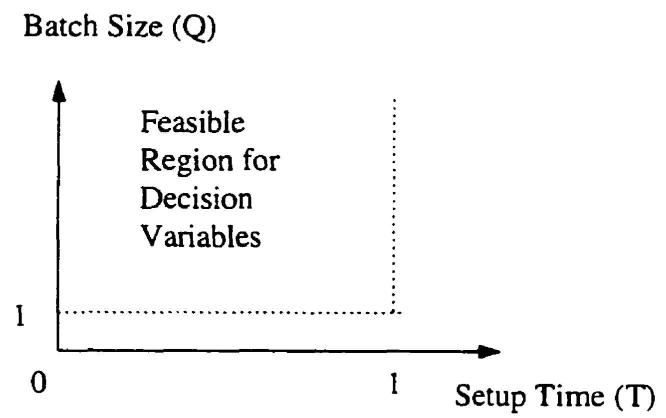


Figure 4.1: Feasible Region for (T, Q) Optimization

No method has been found to analytically minimize a constrained, non-convex function. Fortunately, the location of feasible values for the optimal combinations of T and Q can be determined from heuristic arguments. The following Properties will be used to help construct the optimal solution to this problem.

Property 4.1: The values of (T, Q) which result in minimum E[TC] fall on the boundaries of the feasible region.

Proof: From Equation 4.1.1, no local minima can exist, thus global minima must fall on the boundary of the feasible region. ∎

Property 4.2: The values of (T, Q) which minimize E[TC] do not occur where $Q = \infty$.

Proof: From Equations 3.2.1, 3.2.3 and 3.2.5, the limit of expected waiting time becomes:

$$\lim_{Q \to \infty} E[W] = \lim_{Q \to \infty} \left(\frac{D/_Q \, E[S^2]}{2\left(1 - D/_Q \, E[S]\right)} + E[S] \right)$$

$$= \lim_{Q \to \infty} \left(\frac{D\tau^2/_Q + D(1 + r)\left\{E[X^2] + 2\tau E[X] + (1 + r)QE[X]^2\right\} - D(1 + r^2)E[X]^2}{2\left(1 - D\tau/_Q - D(1 + r)E[X]\right)} \right.$$

$$\left. + \tau + D(1 + r)E[X] \right)$$

$$= +\infty$$

Thus, expected waiting time increases to infinity as $Q \to \infty$. Since E[TC] is proportional to expected waiting times, $Q \to \infty$ implies $E[TC] \to \infty$ which clearly cannot lead to a minimum. ∎

Property 4.3: The values of (T, Q) which minimize E[TC] do not occur where $T = 0$ and $Q > 1$ (i.e., along left boundary of Figure 4.1).

Proof: Equation 3.2.5 for the expected waiting time can be rearranged (for $T = 0$) as:

$$E[W] = \frac{D(1 + r_0)\mathrm{Var}[X]}{2\left\{1 - D(1 + r_0)E[X]\right\}} + \frac{(1 + r_0)\left\{D(1 + r_0)E[X]^2 + E[X]\right\}}{2\left\{1 - D(1 + r_0)E[X]\right\}} Q$$

Clearly, the expected waiting time is an increasing function in Q when $T = 0$, and since the expected total cost is proportional to Q, E[TC] is not minimized for any values of $Q > 1$ when $T = 0$. ∎

Corollary 4.3: When $T = 0$, E[TC] is minimized at $Q = 1$.

Proof: Given that a minimum E[TC] cannot occur for $Q > 1$ (Property 4.3), it must occur at $Q = 1$. ∎

Property 4.4: The values of (T, Q) which minimize E[TC] must occur along the lines $T = 1$ (i.e., along right boundary of feasible region) or $Q = 1$ (i.e., along bottom boundary).

Proof: By Properties 4.1, 4.2 and 4.3, the global minimum cannot reside anywhere else on the feasible region. ∎

These Properties, especially Property 4.4, greatly simplify the task of finding expressions for optimal values of the decision variables since either the value of T is fixed, or the value of Q is fixed. Two optimal investment decisions then exist in this system:

1. Don't Invest, i.e., $T^* = 1$, $Q^* > 1$[1]
2. Invest, i.e., $T^* < 1$, $Q^* = 1$

(This list excludes the trivial case where optimal batch size in the initial, unimproved system is one. In such a case, there would be limited WIP inventory costs to justify investment in setup reduction.)

Since each of these optimal decisions is fixed in one variable, each case becomes an optimization problem in one variable. The optimal values of the decision variables can be found through the calculus of minimization:

<u>Case 1:</u> $T^* = 1$, $Q^* > 1$ (i.e., 'Don't Invest')

The necessary condition is:

[1] For reference, it may be repeated here that the normalized setup time, T, falls in the range [0, 1] while the batch size, Q, falls in the range [1, ∞].

$$\frac{\partial E[TC]}{\partial Q} = \frac{\lambda h}{2(1-\rho)^2}\Big[E[X][2-(1+r_0)DE[X]]\big\{Q[1-(1+r_0)DE[X]]-2\tau_0(1+r_0)D\big\}$$
$$-\lambda\tau_0\big\{D(1+r_0)D^2E[X^2]-(1+r_0^2)D^2E[X]^2+\tau_0[1-(1+r_0)DE[X]]\big\}\Big]$$
$$=0$$

From which the optimum batch size is found as:

$$Q^* = \frac{D\tau_0}{1-(1+r_0)DE[X]} + \sqrt{\frac{D\tau_0}{(1+r_0)E[X]}} \times \sqrt{\frac{\tau_0}{1-(1+r_0)DE[X]} + D\Big[(1+r_0)E[X^2]-(1+r_0^2)E[X]^2\Big]} \tag{4.1.2}$$

While the sufficient condition becomes:

$$\left.\frac{\partial^2 E[TC]}{\partial Q^2}\right|_{Q^*} = \frac{h\sqrt{\tau_0 D}\,[2-(1+r_0)DE[X]]^{3/2}[(1+r_0)E[X]]^{3/2}}{\sqrt{\tau_0 + D\Big[(1+r_0)E[X^2]-(1+r_0^2)E[X]^2\Big][1-(1+r_0)DE[X]]}} > 0$$

So that Q* from (4.1.2) minimizes E[TC].

<u>Case 2:</u> $T^* < 1$, $Q^* = 1$ (i.e., ' Invest')

The necessary condition becomes:

$$\frac{\partial E[TC]}{\partial T} = \frac{\lambda h D\tau_0}{2(1-\rho)^2}\Big\{1 + D\big((1+r_0)D^2E[X^2]-(1+r_0^2)D^2E[X]^2\big)\Big\} - \frac{\tau_0}{2a}(2i - hDa)$$
$$=0$$

From which the optimum setup time is found as:

$$T^* = \frac{1 - (1 + r_0)DE[X]}{D\tau_0} - \frac{1}{D\tau_0}\sqrt{\frac{hDa\left[1 + (1 + r_0)D^2E[X^2] - (1 + r_0^2)D^2E[X]^2\right]}{2i - hDa}} \qquad (4.1.3)$$

The sufficient condition becomes:

$$\left.\frac{\partial^2 E[TC]}{\partial T^2}\right|_{T^*} = \frac{D\tau_0^2}{a}\frac{(2i - hDa)^3}{\sqrt{hDa\left[1 + (1 + r_0)D^2E[X^2] - (1 + r_0^2)D^2E[X]^2\right]}} > 0$$

Now that the behaviour of these two cases is defined, the next question is where, as a function of the cost of eliminating setup times, τ_0/a, does the optimum decision change from 'Don't Invest' to 'Invest'? The following Properties will be used to answer this question.

Property 4.5: E[TC] in the 'Don't Invest' case is constant with respect to τ_0/a

Proof: Substituting T = 1 (i.e., representing no investment) into the E[TC] function gives:

$$E[TC]\Big|_{T=1} = hD\left(\frac{D\tau_0^2/Q + D(1 + r_0)\left\{E[X^2] + 2\tau_0E[X] + (1 + r_0)E[X]^2\right\} - D(1 + r_0^2)E[X]^2}{2\left(1 - D\tau_0^2/Q - D(1 + r_0)E[X]\right)}\right.$$
$$\left. + \tau_0 + D(1 + r_0)QE[X]\right)$$

Inspection shows this expression does not contain the 'a' parameter, so that E[TC] is constant with respect to τ_0/a. ∎

Property 4.6: E[TC] in the 'Invest' case is increasing in τ_0/a.

Proof: Substituting Q = 1 into the E[TC] function gives:

$$E[TC]\Big|_{Q=1} = hD\left(\frac{DT^2\tau_0^2 + D(1+r_0)\left\{E[X^2] + 2T\tau_0E[X] + (1+r_0)E[X]^2\right\} - D(1+r_0^2)E[X]^2}{2\left(1 - DT\tau_0 - D(1+r_0)E[X]\right)}\right.$$

$$\left. + T\tau_0 + D(1+r_0)E[X]\right) + i\,(1-T)\frac{\tau_0}{a}$$

Since $0 \le T^* < 1$, the last term in this function shows that E[TC] is linearly increasing in τ_0/a.

∎

Property 4.7: E[TC] for the 'Don't Invest' case crosses the E[TC] for the 'Invest' case in exactly one location.

Proof: i) Assume $\tau_0/a = 0$. Since the cost of eliminating setup time is zero, the 'Invest' case strategy will be to totally eliminate setup time (i.e., $T^* = 0$). Then,

$$E[TC]_{Invest} = hD\left(\frac{D(1+r_0)\left\{E[X^2] + (1+r_0)E[X]^2\right\} - D(1+r_0^2)E[X]^2}{2\left(1 - D(1+r_0)E[X]\right)} + D(1+r_0)E[X]\right)$$

$$< hD\left(\frac{D\tau_0^2/Q + D(1+r_0)\left\{E[X^2] + 2\tau_0E[X] + (1+r_0)E[X]^2\right\} - D(1+r_0^2)E[X]^2}{2\left(1 - D\tau_0^2/Q - D(1+r_0)E[X]\right)}\right.$$

$$\left. + \tau_0 + D(1+r_0)QE[X]\right)$$

$$= E[TC]_{Don't\ Invest}$$

When $\tau_0/a = 0$, $E[TC]_{Don't\ Invest} > E[TC]_{Invest}$

ii) Assume $\tau_0/a \to \infty$.

$$\lim_{\tau_0/a \to \infty} E[TC]_{\text{Invest}}$$

$$= \lim_{\tau_0/a \to \infty} \left\{ hD\left(\frac{DT^2\tau_0^2 + D(1+r_0)\left\{E[X^2] + 2T\tau_0 E[X] + (1+r_0)E[X]^2\right\} - D(1+r_0^2)E[X]^2}{2(1 - DT\tau_0 - D(1+r_0)E[X])} + T\tau_0 + D(1+r_0)E[X] \right) + i(1-T)\frac{\tau_0}{a} \right\}$$

$$= \infty$$

for $T < 1$. Thus, when $\tau_0/a \to \infty$, $E[TC]_{\text{Don't Invest}} < E[TC]_{\text{Invest}}$ since $E[TC]_{\text{Don't Invest}}$ remains at a finite level independent of τ_0/a due to Property 4.5.

Since the E[TC] function in the 'Don't Invest' case is constant with respect to τ_0/a, and in the 'Invest' case is linearly increasing in τ_0/a (Property 4.6), and since the E[TC] of the 'Invest' case is lower than that of the 'Don't Invest' case for $\tau_0/a = 0$ but is higher when $\tau_0/a \to \infty$, the two E[TC] functions must cross for a single value of τ_0/a, denoted τ_0/a^*. ∎

Property 4.8: The value τ_0/a^* described in Property 4.7 represents the location where the optimal investment decision changes from 'Invest' to 'Don't Invest'.

Proof: By definition, the optimal investment decision is the one which minimizes E[TC]. By Property 4.7, E[TC] is minimized by the 'Invest' decision over the range $0 \le \tau_0/a \le \tau_0/a^*$ and by the 'Don't Invest' decision over the range $\tau_0/a^* \le \tau_0/a < \infty$. Thus, the optimal investment decision changes at τ_0/a^*. ∎

Property 4.9: The value τ_0/a^* is located where the value of the two E[TC] functions is equal.

Proof: Since the two E[TC] functions cross at a single location (Property 4.7), they must assume the same value at that location. Since this location also represents the point at which the optimal investment decision changes (Property 4.8), this point must occur when the two E[TC]s are equal. ∎

Property 4.9 allows the calculation of the value τ_0/a^*, that is, the location of the decision boundary between the two optimal decision cases. Equating the expected total costs gives:

$$E[TC]_{Invest} = E[TC]_{DontInvest} \quad \Rightarrow$$

$$hD\left(\frac{DT^{*2}\tau_0^2 + D(1+r_0)\left\{E[X^2] + 2T^*\tau_0 E[X] + (1+r_0)E[X]^2\right\} - D(1+r_0^2)E[X]^2}{2\left(1 - DT^*\tau_0 - D(1+r_0)E[X]\right)} + T^*\tau_0 + D(1+r_0)E[X]\right) + i\left(1 - T^*\right)\frac{\tau_0}{a}$$

$$= hD\left(\frac{D\tau_0^2/Q^* + D(1+r_0)\left\{E[X^2] + 2\tau_0 E[X] + (1+r_0)E[X]^2\right\} - D(1+r_0^2)E[X]^2}{2\left(1 - D\tau_0^2/Q^* - D(1+r_0)E[X]\right)} + \tau_0 + D(1+r_0)Q^*E[X]\right)$$

or

$$
\begin{aligned}
\frac{\tau_0}{a^*} = \frac{hD}{i\left(1 - T^*\right)} \Bigg(& + \tau_0 (1 - T^*) + D(1 + r_0) E[X](Q^* - 1) \\
& + \frac{D\tau_0^2 / Q^* + D(1 + r_0)\left\{E[X^2] + 2\tau_0 E[X] + (1 + r_0) E[X]^2\right\} - D(1 + r_0^2) E[X]^2}{2\left(1 - D\tau_0^2 / Q^* - D(1 + r_0) E[X]\right)} \\
& \frac{DT^{*2}\tau_0^2 + D(1 + r_0)\left\{E[X^2] + 2T^*\tau_0 E[X] + (1 + r_0) E[X]^2\right\} - D(1 + r_0^2) E[X]^2}{2\left(1 - DT^*\tau_0 - D(1 + r_0) E[X]\right)} \Bigg)
\end{aligned}
\tag{4.1.4}
$$

Where T* and Q* are from Equations 4.1.3 and 4.1.2, respectively.

These results can be incorporated into a decision matrix similar to that of the last chapter. Such a graph is shown in Figure 4.2.

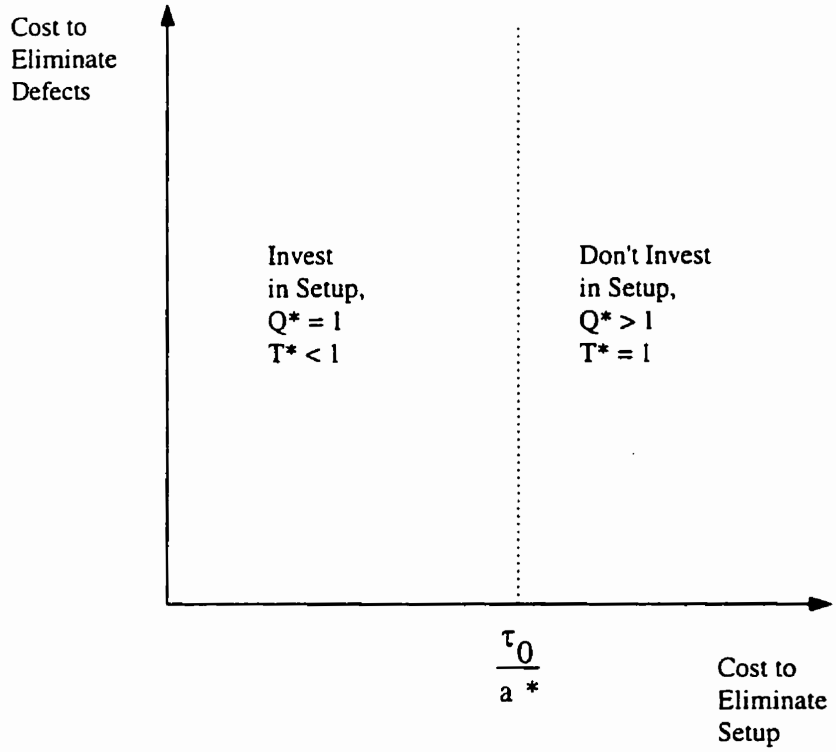


Figure 4.2: Decision Matrix for Setup Investment, 'Naive' Case.

To the left of the value τ_0/a^*, given in Equation 4.1.4, the optimal strategy is to invest in setup reduction such that $Q^* = 1$ and T^* is as given in Equation 4.1.3. To the right of this boundary value, the optimal strategy is to invest nothing so that $T^* = 1$ and Q^* is as given in Equation 4.1.2.

Intuitively, these results can be explained in this way. When the cost to eliminate setup (τ_0/a) is relatively high, the optimal decision is to do nothing, i.e., 'Don't Invest'. As this cost decreases, the point will be reached where the first increment of investment is justified and setup time is reduced. Reduced setup time leads to a lower optimum batch size, which provides additional cost savings. These extra savings justify more setup reduction which leads to smaller batch sizes and yet more savings. In effect, this is a system with positive feedback.

Since this model assumes the linear investment function holds, the marginal cost of improving setup time is constant. Once the first increment of investment takes place, each additional increment is also justified until batch size is reduced to one, in which case no further reductions in batch size are possible so no additional reductions in setup time would result.

The optimum investment decision is then either to not invest at all in setup reduction, or to invest until batch size is reduced to a single unit, confirming the analytic development of this section. When investment is made, setup time will be reduced substantially but may not reach zero. This is because batch size cannot be reduced below a single unit so there will always be some WIP in the system and the last bit of WIP savings cannot be realized to justify total elimination of setup time. Numerical experiments have shown that the optimal setup time is quite small, typically $T^* < 0.001$. For practical purposes this optimal setup time can be assumed to be not significantly different from zero (i.e., $T^* = 0$), and the equivalent optimal decision is to 'Invest Fully'.

4.1.2 Quality Level Optimization

The second part of the 'naive' case is to consider investments in quality improvement. Since quality improvement is assumed to be independent of setup times and batch sizes, total costs of the system are optimized with respect to the single variable R. Setup time is assumed to be maintained at the initial level (i.e., $T = 1$), and batch size, Q, is left as a parameter.

Under these conditions, investment decisions for quality are the same as was determined in Section 3.4 of the last chapter. The critical values for costs of eliminating defects were given in Equations 3.4.7. Figure 4.3 illustrates where these boundaries fall on the decision matrix.

Superimposing the matrices of Figures 4.2 and 4.3 provides the final decision matrix for the 'Naive' investment case, as shown in Figure 4.4.

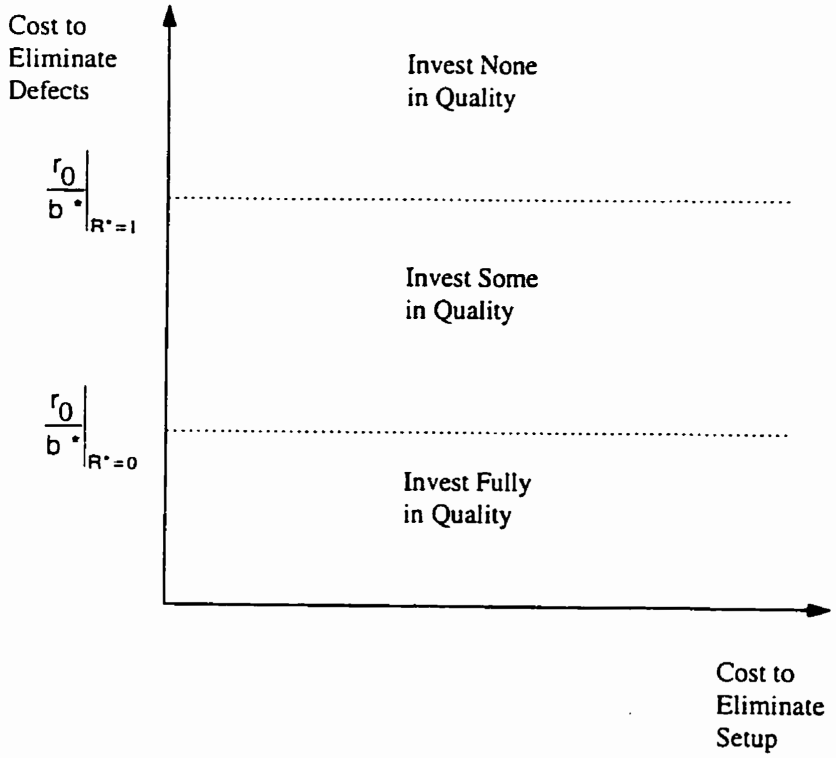


Figure 4.3: Decision Matrix for Quality Investment, 'Naive' Case.

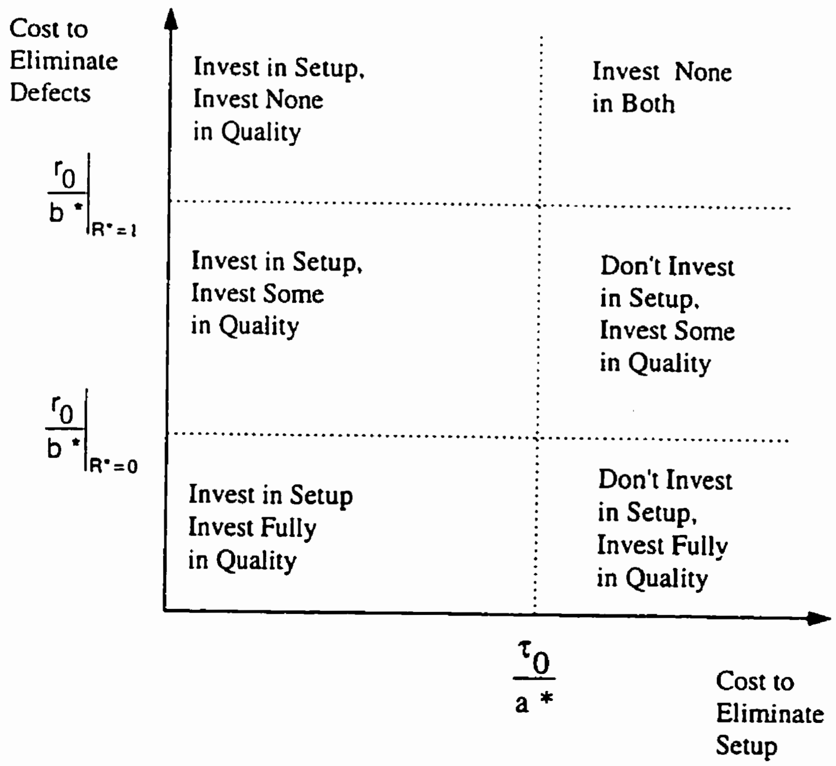


Figure 4.4: Complete Decision Matrix for 'Naive' Investment Case.

This completed 'naive' case decision matrix is analogous to that of Figure 3.4 of the last chapter. The difference is that this decision matrix only contains regions for six different optimal decisions; because of the linkage between setup time reduction and batch size, in this system it is never optimal to invest 'some' in setup reduction.

In the next section, the 'informed' case decision matrix is developed for comparison with the 'naive' case matrix developed here.

4.2 'Informed' Investment Decisions

In this section, optimal investment decisions are determined for the 'informed' case, that is, for the case where total costs are minimized simultaneously over the three decision variables.

4.2.1 Convexity of the E[TC] function in three variables

Once again, the optimal investment decision is the one which minimizes the expected total cost of operating the production system. As a first test, the E[TC] function in three variables is examined for convexity by checking the determinant of the principal minors of the Hessian of the E[TC] function.

The complexity of the full Hessian of three variables can be dispensed with by utilizing the property that a function is convex when its Hessian matrix is positive definite (or semi-definite). In the three variable case, the Hessian matrix is positive definite if and only if all its principal minors are strictly positive (or positive semi-definite if the first two are strictly positive and the third equals zero). [Peressini, et al. 1988]

In this case, the second principal minor (Δ_2) is:

$$\Delta_2 = \det\begin{bmatrix} \frac{\partial^2 E[TC]}{\partial T^2} & \frac{\partial^2 E[TC]}{\partial T \partial R} \\ \frac{\partial^2 E[TC]}{\partial T \partial R} & \frac{\partial^2 E[TC]}{\partial R^2} \end{bmatrix} \tag{4.2.1}$$

This equation will be recognized as being equivalent to Equation 3.5.2, so that

$$\Delta_2 = -\frac{D^4 h^2 \tau_0^2 r_0^2}{4(1-\rho)^4}\left\{4\rho^2 \mathrm{Var}[X] + \lambda^2 E[X^2]^2 + \frac{4}{Q}E[X]^2\right\} < 0 \tag{4.2.2}$$

Because the second principal minor is strictly negative for all parameter values, the Hessian matrix cannot become positive definite and the E[TC] function is nowhere convex. Even if critical points are found within the feasible space of this problem, they could not represent minima of the expected total cost. So, as with the bivariate optimization problem of the last chapter, the optimal values of the decision variables must fall on the boundaries of the feasible region for this problem.

4.2.2 Development of Decision Matrix

A decision matrix similar to that of Section 3.5 is developed here for the 'informed' case of this model. As the total cost function here is also non-convex, the decision matrix will be developed through arguments regarding the behaviour of optimal decisions under different circumstances.

Because of the complexity of the decision matrix in this trivariate case, the complete matrix is given now as Figure 4.5. Justification for this matrix will proceed below, in which each of the discrete decision boundary segments labeled on Figure 4.5 is developed mathematically.

<u>Boundary Segment I</u>

To start with, consider a case when the cost of eliminating defects, r_0/b, is arbitrarily high. In this case, the optimal decision regarding quality improvement would be to invest nothing, so that $R^* = 1$. This situation is then equivalent to that of Section 4.1.1, that is, the optimization problem is to minimize E[TC] as a function of T and Q alone.

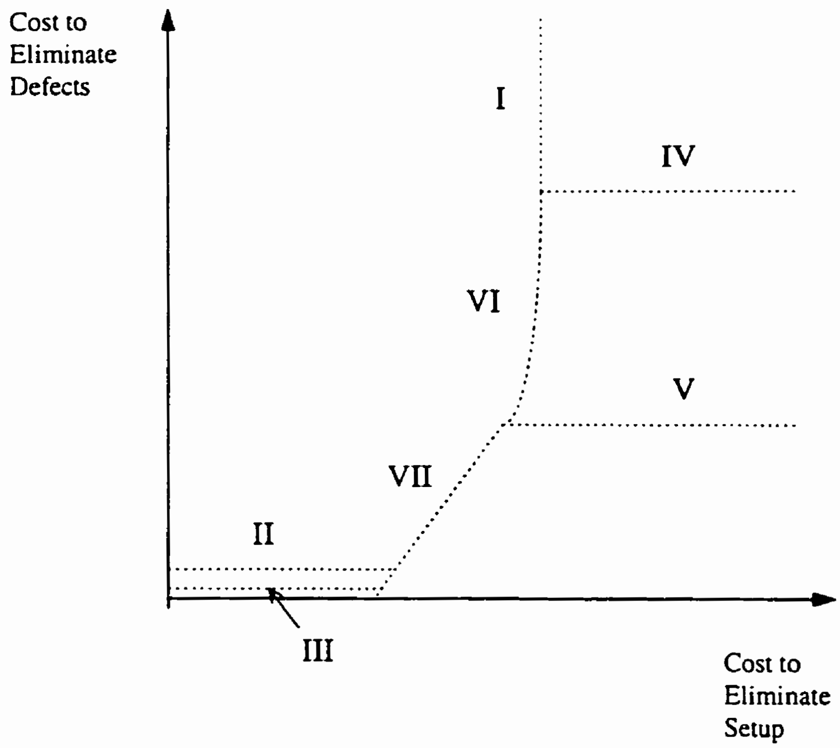


Figure 4.5: 'Informed' Case Decision Matrix Showing Boundary Segments
(Note: Roman numerals refer to line segments discussed in the text.)

Two optimal decisions will then exist. For low values of τ_0/a, the optimal levels of the decision variables are ($T^* = 0$, $Q^* = 1$, $R^* = 1$), with T^* determined from Equation 4.1.3, and for high values of τ_0/a the optimal decision is ($T^* = 1$, $Q^* > 1$, $R^* = 1$) where Q^* is given by Equation 4.1.2. The boundary between these two decisions occurs at τ_0/a^* given by Equation 4.1.4. Since τ_0/a^* does not depend on r_0/b, this boundary will plot as a vertical line on the decision matrix, as long as $R^* = 1$.

Boundary Segments II and III

As r_0/b decreases, the cost of improving quality decreases. At some point the optimal decision will change from no investment in quality improvement (i.e., $R^* = 1$) to some investment (i.e., $R^* < 1$).

Starting from a high level of r_0/b, two cases can be differentiated: τ_0/a below τ_0/a^*, and, τ_0/a above τ_0/a^*. As was shown in Section 4.1.1, each of these cases resulted in a different optimal decision. The effect of decreasing r_0/b in each of these cases will be examined now.

Case 1: $\tau_0/a < \tau_0/a^*$ (i.e., $T^* = 0$, $Q^* = 1$)

For simplicity, suppose that $\tau_0/a = 0$ is considered first (i.e., cost of eliminating setup time is zero). Optimal setup time will then be zero, optimal batch size will continue to be one unit and with these two variables fixed, the location of the quality investment boundaries can be found from Equation 3.5.5, here with $T^* = 0$ and $Q^* = 1$:

$$\frac{r_0}{b^*} = \frac{hDr_0E[X]}{2i}\left[1 - \frac{2DRr_0E[X]}{\{1 - D(1 + Rr_0)E[X]\}} + \frac{1 - D^2(1 + R^2r_0^2)E[X]^2}{\{1 - D(1 + Rr_0)E[X]\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{DE[X^2]}{\{1 - D(1 + Rr_0)E[X]\}^2}\right] \tag{4.2.3}$$

The boundary for the 'no investment' in quality decision (i.e., $R^* = 1$, Boundary Segment II) falls at:

$$\left.\frac{r_0}{b^*}\right|_{R^*=1} = \frac{hDr_0E[X]}{2i}\left[1 - \frac{2Dr_0E[X]}{\{1 - D(1 + r_0)E[X]\}} + \frac{1 - D^2(1 + r_0^2)E[X]^2}{\{1 - D(1 + r_0)E[X]\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{DE[X^2]}{\{1 - D(1 + r_0)E[X]\}^2}\right] \tag{4.2.4}$$

And the boundary for the 'full investment' in quality decision (i.e., R* = 0, Boundary Segment III) falls at:

$$\left.\frac{r_0}{b^*}\right|_{R^*=0} = \frac{hDr_0E[X]}{2i}\left[1 + \frac{1 - D^2E[X]^2}{\{1 - DE[X]\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{DE[X^2]}{\{1 - DE[X]\}^2}\right] \tag{4.2.5}$$

As will be shown in the example numerically-derived decision matrices of the next section, the position of these boundaries is so close to the horizontal axis that when plotted to scale they are indistinguishable from the axis.

<u>Case 2:</u> $\tau_0/a > \tau_0/a^*$ (i.e., T* = 1, Q* > 1)

For an arbitrarily large value of τ_0/a, the optimal decision will be to not invest in setup reduction, fixing T* at one. Under this condition, optimal batch size, Q*, from Equation 4.1.2 can be restated as a function of R* as:

$$Q^* = \frac{D\tau_0}{1 - (1 + R^*r_0)DE[X]} + \sqrt{\frac{D\tau_0}{(1 + R^*r_0)E[X]}} \times \sqrt{\frac{\tau_0}{1 - (1 + R^*r_0)DE[X]} + D\left[(1 + R^*r_0)E[X^2] - (1 + R^{*2}r_0^2)E[X]^2\right]} \tag{4.2.6}$$

Since Q* is not a function of τ_0/a or r_0/b, it assumes a fixed value for any value of R*. R*, on the other hand, is a function of r_0/b and Q*. R* can be shown to be a minimizer of E[TC] by examining the sufficient condition:

$$\frac{\partial^2 E[TC]}{\partial R^2} = \frac{hE[X]r_0^2\lambda^2}{\left(1 - D\left[\tau_0 + (1 + Rr_0)E[X]\right]\right)^3} \times \left[\left\{Q^2\left[Q - D^2E[X]^2\right] - \left[DQE[X] - (Q - D\tau_0)\right]^2\right\}E[X] + DQ(Q - D\tau_0)E[X^2]\right] \tag{4.2.7}$$

Since Equation 4.2.7 is non-negative for any feasible value of Q, or R, E[TC] is convex in R. Critical points will represent minimizers of the E[TC], and the optimal investment decision.

First, note that three sub-cases can be distinguished:

1. high r_0/b, leading to R* = 1 (i.e., no investment)
2. low r_0/b, leading to R* = 0 (i.e., full investment)
3. moderate r_0/b, leading to 0 < R* < 1 (i.e., some investment)

Boundary Segment IV

The first sub-case, where r_0/b is high and R* =1, was discussed in Section 4.1.1 where Q* was determined in Equation 4.1.2. The location of the boundary between this sub-case and the third sub-case, where some investment takes place, can be found in the same manner as in Chapter 3. Using Q* from Equation 4.1.2 (which is invariant to r_0/b), R* is set to one and r_0/b* calculated from Equation 3.5.5 becomes:

$$\left.\frac{r_0}{b^*}\right|_{R^*=1,Q^*} = \frac{hDr_0Q^*E[X]}{2i}\left[1-\frac{2\,{}^{D}\!/_{Q^*}\,r_0E[X]}{\left\{1-{}^{D}\!/_{Q^*}\left[\tau_0+(1+r_0)Q^*E[X]\right]\right\}}\right.$$
$$\left.+\frac{1-\left({}^{D}\!/_{Q^*}\right)^2(1+r_0^2)Q^*E[X]^2}{\left\{1-{}^{D}\!/_{Q^*}\left[\tau_0+(1+r_0)Q^*E[X]\right]\right\}^2}\right]$$
$$+\frac{hDr_0}{2i}\left[\frac{{}^{D}\!/_{Q^*}\,(Q^*-D\tau_0)E[X^2]}{\left\{1-{}^{D}\!/_{Q^*}\left[\tau_0+(1+r_0)Q^*E[X]\right]\right\}^2}\right] \tag{4.2.7}$$

It can be seen that the location of this quality decision boundary does not include a τ_0/a term so that this boundary will lie horizontally on the decision matrix over the range $\tau_0/a^* < \tau_0/a < \infty$ (where τ_0/a^* is given in Equation 4.1.4).

Boundary Segment V

The second sub-case, for low r_0/b, is treated similarly. Equation 4.1.2 for Q^* when $R^* = 0$ rather than 1 becomes:

$$\left.Q^*\right|_{R=0} = \frac{D\tau_0}{1-DE[X]} + \sqrt{\frac{D\tau_0}{E[X]}}\sqrt{\frac{\tau_0}{1-DE[X]} + D\left[E[X^2]-E[X]^2\right]} \tag{4.2.8}$$

The 'full investment' boundary in this sub-case is also found from Equation 3.5.5 by setting $R^* = 0$ and using Q^* from 4.2.8:

$$\left.\frac{r_0}{b^*}\right|_{R^*=0,Q^*} = \frac{hDr_0Q^*E[X]}{2i}\left[1 + \frac{1-\left(\frac{D}{Q^*}\right)^2 Q^*E[X]^2}{\left\{1-\frac{D}{Q^*}\left[\tau_0 + Q^*E[X]\right]\right\}^2}\right] + \frac{hDr_0}{2i}\left[\frac{\frac{D}{Q^*}(Q^* - D\tau_0)E[X^2]}{\left\{1-\frac{D}{Q^*}\left[\tau_0 + Q^*E[X]\right]\right\}^2}\right] \tag{4.2.9}$$

The third sub-case, which represents the decision region bounded by Boundary Segments IV, V and VI, involves simultaneous optimization over R and Q. The necessary conditions are:

$$\begin{aligned}\frac{\partial E[TC]}{\partial R} &= -\frac{r_0}{2b}\{2i - hDbQE[X]\} \\ &+ \frac{\lambda hr_0}{2(1-\rho)^2}\left\{D(Q - D\tau_0)\left[E[X^2] - 2Rr_0E[X]^2\right] + D^2Q\left[(1+Rr_0)^2 - 2\right]E[X]^3 + Q^2E[X]\right\} \\ &= 0\end{aligned} \tag{4.2.10a}$$

$$\begin{aligned}\frac{\partial E[TC]}{\partial Q} &= \frac{\lambda h}{2(1-\rho)^2}\Big[E[X][2-(1+Rr_0)DE[X]]\left\{Q[1-(1+Rr_0)DE[X]] - 2\tau_0(1+Rr_0)D\right\} \\ &\quad -\lambda\tau_0\left\{(1+Rr_0)D^2E[X^2] - (1+R^2r_0^2)D^2E[X]^2 + \tau_0[1-(1+Rr_0)DE[X]]\right\}\Big] \\ &= 0\end{aligned} \tag{4.2.10b}$$

Solving these equations simultaneously leads to an eighth-order polynomial for Q*, which, when written out takes up approximately one page. An analytic solution for this equation has not been found. The solution for R* is similarly complicated and will be left for the moment, but it will be noted that Equations 4.2.10 are independent of τ_0/a so solutions will not change along (horizontal) lines of constant r_0/b. Algorithms will be presented to solve for these variables in the discussion of Boundary Segments VI and VII, below.

By invoking Property 4.9 of the last section, that is, that E[TC] is continuous over the decision matrix, the remaining boundaries can be developed.

Boundary Segments VI and VII

The 'vertical' boundary, that is the boundary between the optimal decision to invest nothing in setup or to invest in setup and decrease batch size to one unit, can be extended for two more cases, for the range ($T^* = 1, 0 < R^* < 1$) and ($T^* = 1, R^* = 0$). These two cases are developed as:

Case 1: Over range $\left.\frac{r_0}{b^*}\right|_{T^*=1, R^*=0} < \frac{r_0}{b^*} < \left.\frac{r_0}{b^*}\right|_{T^*=1, R^*=1}$, i.e., Boundary Segment VI

Property 4.9 provides that $E[TC]_{left} = E[TC]_{right}$. It is known that on the left $T_L = T_L^*$, $Q_L^* = 1$, $R_L^* = 1$ while on the right of the boundary $T_R^* = 1$, $Q_R = Q_R^*$, $0 < R_R^* < 1$. Substituting these values into the E[TC] function (Equation 3.3.6) yields

$$E[TC]_{left} = hD\left[\frac{D\left[T_L^{*2}\tau_0^2 + (1 + r_0)\left\{E[X^2] + 2T_L^*\tau_0 E[X] + (1 + r_0)E[X]^2\right\} - (1 + r_0^2)E[X]^2\right]}{2\left(1 - D\left\{T_L^*\tau_0 + (1 + r_0)E[X]\right\}\right)} + \left\{T_L^*\tau_0 + (1 + r_0)E[X]\right\}\right] + i\left[\frac{(1 - T_L^*)\tau_0}{a}\right]$$

(4.2.11a)

$$
E[TC]_{right} = hD\left[\frac{\frac{D}{Q_R^{*}}\left[\tau_0^2 + (1 + R_R^{*}r_0)Q_R^{*}\left\{E[X^2] + 2\tau_0 E[X] + (1 + R_R^{*}r_0)Q_R^{*}E[X]^2\right\} - (1 + R_R^{*2}r_0^2)Q_R^{*}E[X]^2\right]}{2\left(1 - \frac{D}{Q_R^{*}}\left\{\tau_0 + (1 + R_R^{*}r_0)Q_R^{*}E[X]\right\}\right)} + \left\{\tau_0 + (1 + R_R^{*}r_0)Q_R^{*}E[X]\right\}\right] + i\left[\frac{(1 - R_R^{*})r_0}{b}\right]
$$

(4.2.11b)

Equating $E[TC]_{right}$ to $E[TC]_{left}$ and rearranging yields:

$$
\frac{r_0}{b^*} = \frac{\tau_0}{a^*}\frac{(1 - T_L^{*})}{(1 - R_R^{*})} + \frac{h}{i}\frac{D}{(1 - R_R^{*})}
\left\{\frac{D\left(T_L^{*2}\tau_0^2 + (1 + r_0)\left\{E[X^2] + 2T_L^{*}\tau_0 E[X] + (1 + r_0)E[X]^2\right\} - (1 + r_0^2)E[X]^2\right)}{2\left[1 - D\left(T_L^{*}\tau_0 + (1 + r_0)E[X]\right)\right]}\right.
$$
$$
- \frac{\frac{D}{Q_R^{*}}\left(\tau_0^2 + (1 + R_R^{*}r_0)Q_R^{*}\left\{E[X^2] + 2\tau_0 E[X] + (1 + R_R^{*}r_0)Q_R^{*}E[X]^2\right\} - (1 + R_R^{*2}r_0^2)Q_R^{*}E[X]^2\right)}{2\left[1 - \frac{D}{Q_R^{*}}\left(\tau_0 + (1 + R_R^{*}r_0)Q_R^{*}E[X]\right)\right]}
$$
$$
\left. + T_L^{*}\tau_0 - \tau_0 + (1 + r_0)E[X] - (1 + R_R^{*}r_0)Q_R^{*}E[X]\right\}
$$

(4.2.12)

This equation is still in terms of the unknown values T_L^*, R_R^* and Q_R^*, which are functions of r_0/b^* and τ_0/a^*. The solution to this system for r_0/b^* as a function of τ_0/a^* (i.e., finding the position of the boundary in terms of the vertical and horizontal axes of the decision matrix) involves solving an eighth-order polynomial equation, which is analytically impractical. However, an algorithm may be used to plot the location of the boundary on the decision matrix:

Algorithm 4.1: Determination of Boundary Segment VI

1) Choose a value R*, $0 < R^* < 1$

2) find r_0/b^* from Equation 4.2.3

3) find Q* from Equation 4.1.2

4) substitute values for R*, r_0/b^*, Q* into Equation 4.2.11b to calculate $E[TC]_{right}$

5) equate $E[TC]_{right}$ to $E[TC]_{left}$ as in 4.2.12 and solve for τ_0/a^*

6) plot r_0/b^* vs. τ_0/a^* for the range of $0 < R^* < 1$ in Step 1.

With this algorithm, the location of decision boundary representing Boundary Segment VI can be calculated and plotted.

Case 2: Over range $\dfrac{r_0}{b^*} < \left.\dfrac{r_0}{b^*}\right|_{T^*=1, R^*=0}$, i.e., Boundary Segment VII

In this case it is assumed that the regions in the lower left of the decision matrix (where Q* = 1 and R* < 1) are of insignificant size and are ignored. Since R_R^* has a constant value of 1 in this region, and Q_R^* also assumes a constant value, using the assumption that $T_L^* = 0$ and Property 4.9 for continuous total costs across the decision space leads to:

$$E[TC]_{left} = E[TC]_{right}$$

$$\Rightarrow \frac{r_0}{b^*} = \frac{\tau_0}{a^*} + \frac{hD}{i}\left\{ \frac{D\left((1+r_0)E[X^2] + 2r_0E[X]^2\right)}{2\left[1 - D(1+r_0)E[X]\right]} \right.$$

$$\left. \frac{D/_{Q^*}\left(\tau_0^2 + Q^*\left\{E[X^2] + 2\tau_0E[X] + (Q^* - 1)E[X]^2\right\}\right)}{2\left[1 - D/_{Q^*}\left(\tau_0 + Q^*E[X]\right)\right]} - \tau_0 + (1+r_0)E[X] - Q^*E[X] \right\}$$

Because none of the parameters in this equation other than the first depend on r_0/b^* or τ_0/a^*, this result leads to a decision boundary which is linear over the range

$0 < \frac{r_0}{b^*} < \left.\frac{r_0}{b^*}\right|_{T^*=1, R^*=0}$ and can be plotted directly as a function of τ_0/a^*, representing Boundary Segment VII in Figure 4.5.

By assuming that the boundaries at $\left.\frac{r_0}{b^*}\right|_{T^*=0, R^*=1}$ and $\left.\frac{r_0}{b^*}\right|_{T^*=0, R^*=0}$ (i.e., Boundary Segments II and III) are independent of τ_0/a, they will plot horizontally on the decision matrix from the value $\tau_0/a = 0$ until the 'no-invest' boundary for setup is reached. Together, these Boundary Segments complete the decision matrix, as shown in Figure 4.6.

Comparing this decision matrix to that of the 'naive' case in the last section shows two major differences. First, the boundary between not investing and investing in setup curves to the left as the cost of improving quality decreases. As the cost of improving quality decreases, and the amount of improvement performed on quality increases, the decision to invest in setup reduction only occurs for lower and lower marginal costs of making setup improvements. This corresponds to the leftward shift seen in the setup decision boundaries in the bivariate case of Chapter 3. Thus, an interaction between the practices is seen in this trivariate case as well.

The second significant difference between the 'naive' and 'informed' decision matrices is that in the latter when investment is called for in setup reduction, improvements only occur in quality when the cost of making those improvements is dramatically lower than when no setup improvement has occurred.

The next section will illustrate the differences between these two cases through numerical examples and will estimate the magnitudes of decision errors.

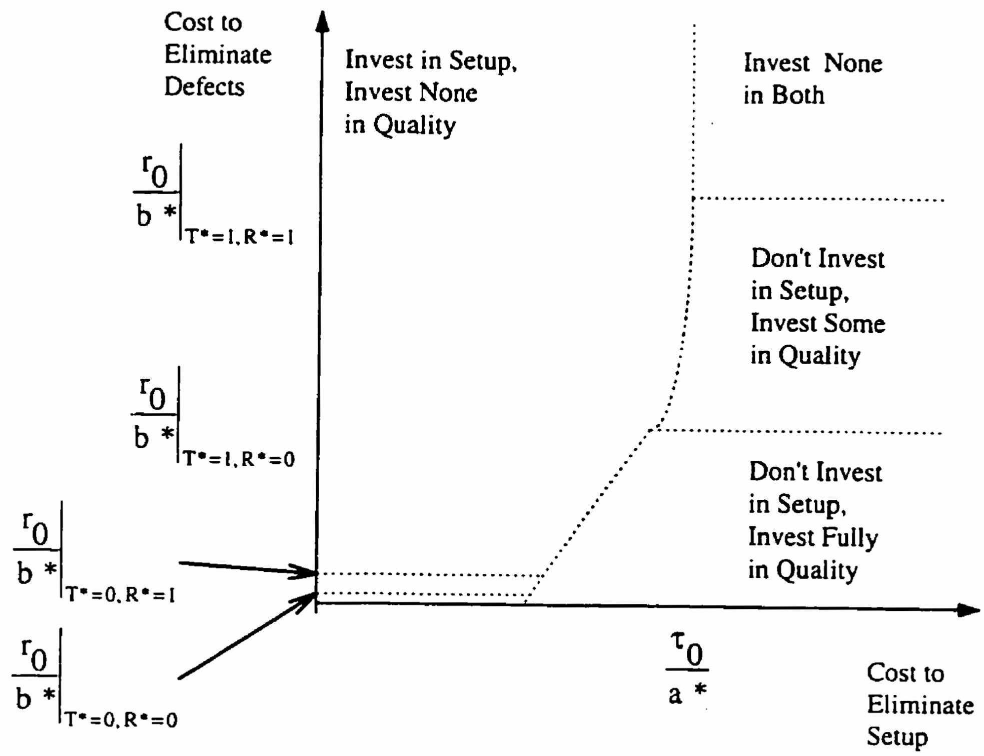


Figure 4.6: Completed Decision Matrix for 'Informed' Case.

4.3 Decision Error and Sensitivity

Just as in the comparison between the 'naive' and 'informed' cases in the last chapter where batch size was fixed, the system with a variable batch size also shows interactions between improvement practices, and suggests over-investment is the consequence of ignoring these interactions. This section will examine this decision error and its sensitivity to various system parameters.

4.3.1 Decision Error

Due to the complexity of the shapes of the various decision regions on the decision matrices, analytic expressions for the decision error have not been found. Instead, numerical cases will be examined.

The Figures in this section are based upon the example problem parameters given in Table 3.2. To ease comparisons, all decision matrices are plotted to a normalized scale such that the point where investment starts to take place in the 'naive' case is scaled to the value of one. This scaling permits the shapes of the decision matrices under different circumstances to be more easily compared since the relative sizes will change as parameter values change.

To begin with, the decision matrices for setup reduction under the 'naive' (Figure 4.7) and 'informed' (Figure 4.8) cases have been plotted. These graphs were calculated numerically, and show the full/none investment decision behaviour predicted in the analytic derivations of the last two sections. Figure 4.8 shows the three Boundary Segments (i.e., I, VI and VII) described in Figure 4.5, as well as a tiny deviation where Boundary Segment VII meets the horizontal axis. This deviation is due to optimal setup time, T^*, being slightly different from zero in this region while the analytic derivation of the last section made the assumption that T^* was equal to zero. The degradation of solution accuracy due to this assumption is not felt to be significant.

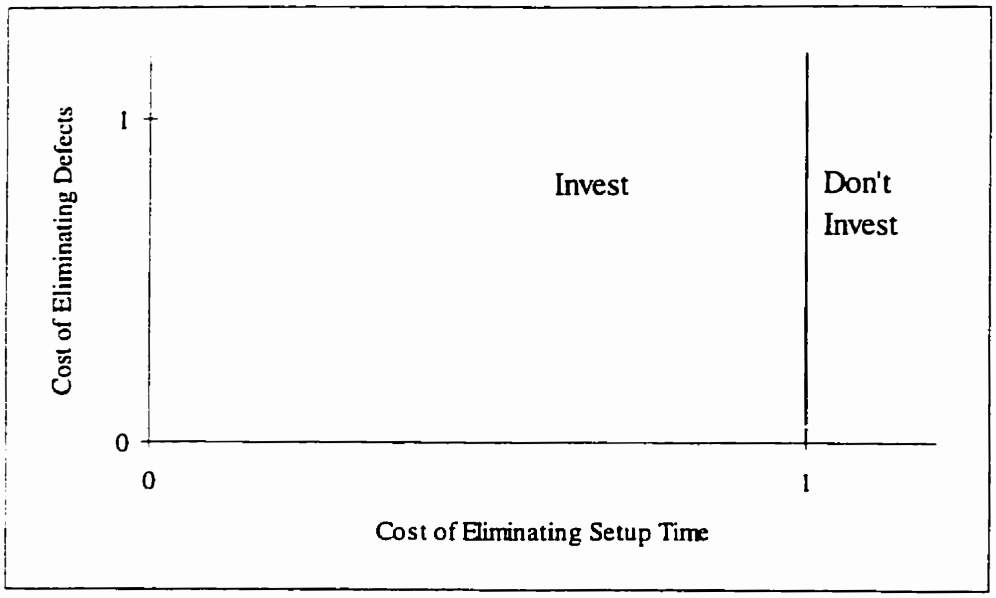


Figure 4.7: Normalized 'Naive' Case Setup Investment Decisions

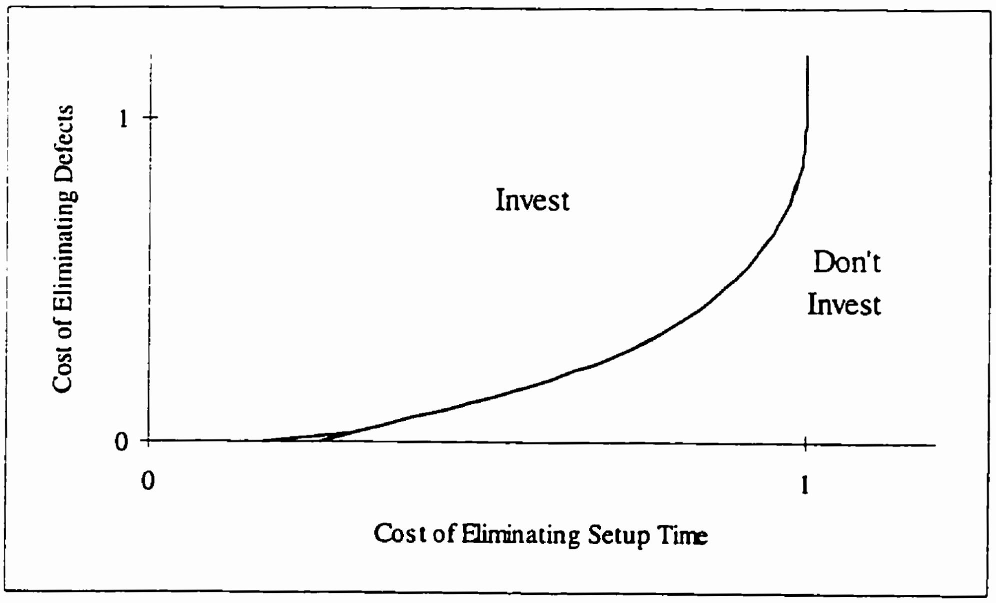


Figure 4.8: Normalized 'Informed' Case Setup Investment Decisions

The 'naive' and 'informed' setup boundaries are overlaid in Figure 4.9. The region between the two boundaries is the area of the matrix in which investment will be made if interactions are ignored, but no investment will be made if interactions are considered. This region may be thought of as where the setup investment decision error occurs.

Figures 4.10 and 4.11 similarly show the boundaries in the 'naive' and informed' cases, respectively, for investment in quality. The most striking aspect of Figure 4.11 is that in the region where investment is made in setup reduction (i.e., above and to the left of the "J" boundary), there is no investment in quality improvement. This behaviour is a consequence of the linear investment function and the non-convex total cost function. It may also be observed that in the 'informed' case, Boundary Segments II and III from Figure 4.5 are not visible. These boundaries do exist in Figure 4.11, but are so close to the axis that they are not visible at the scale of the plot.

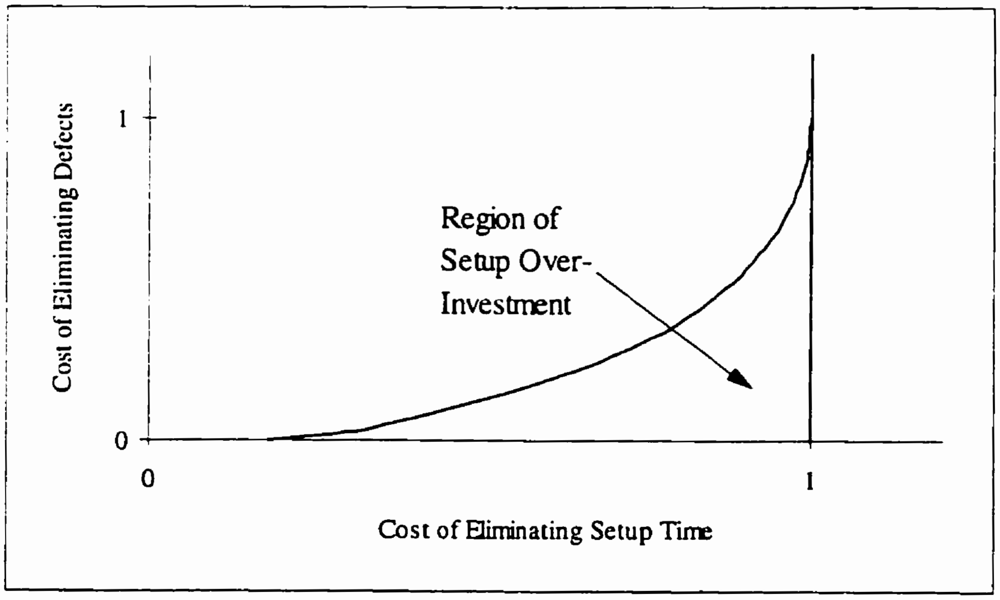


Figure 4.9: Difference in Setup Investment Decisions, 'Naive' to 'Informed' Cases

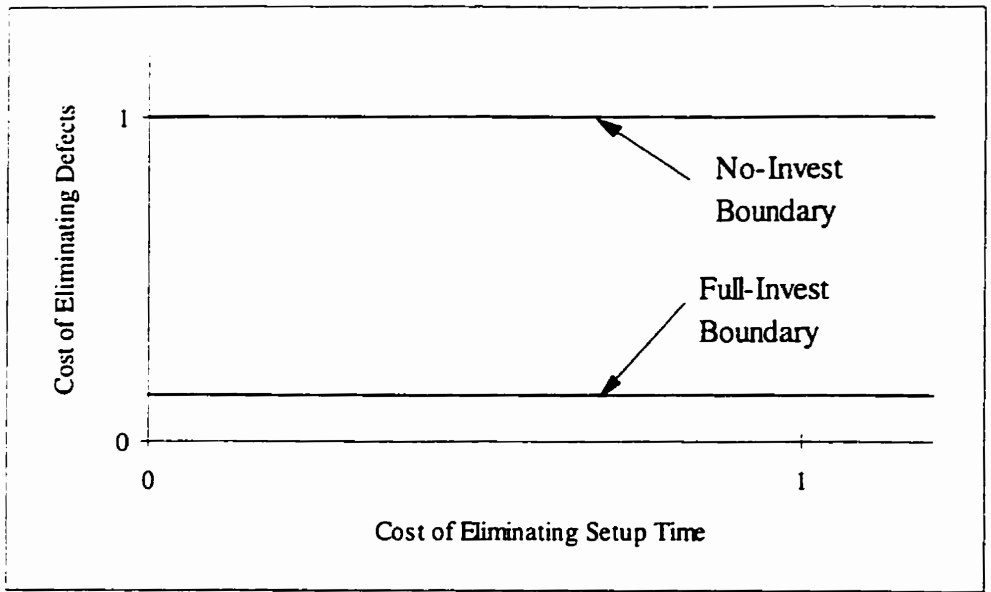


Figure 4.10: Normalized 'Naive' Case Quality Investment Decisions

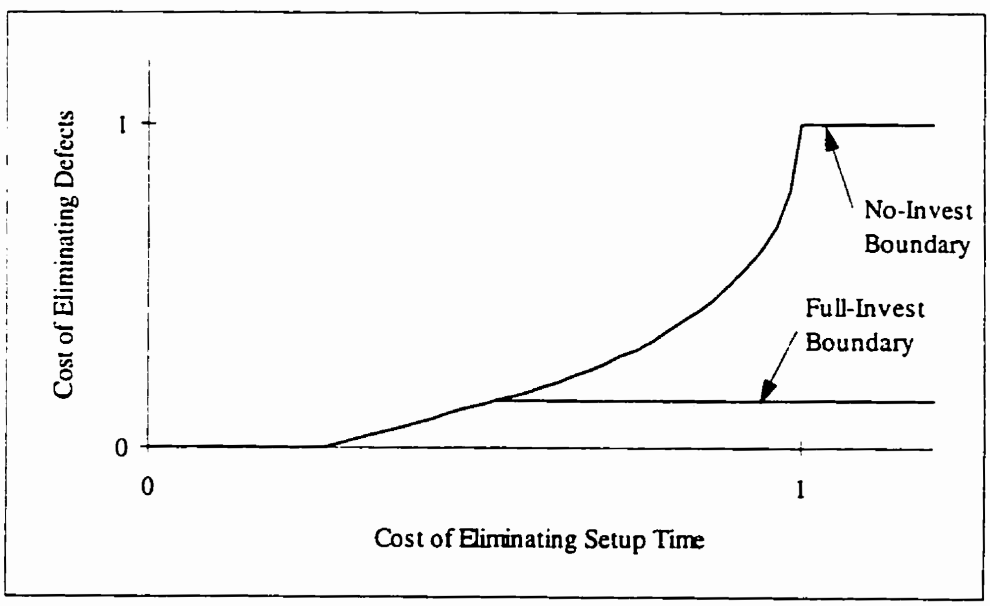


Figure 4.11: Normalized 'Informed' Case Quality Investment Decisions

The differences in the regions between the 'naive' and 'informed' cases for quality investment are plotted in Figure 4.12. Here, there are fairly substantial sized areas where considering interactions will change the optimal investment decision from 'invest some' to 'invest none', or 'invest fully' to 'invest none'. Comparing Figures 4.9 and 4.12 shows that in the region of the decision matrix where the 'naive' strategy calls for at least some investment in both practices (i.e., over the rectangle bounded by {(0,0), (0,1), (1,1), (1,0)}), ignoring potential interactions leads to either over-investment in setup reduction, or over-investment in quality improvement.

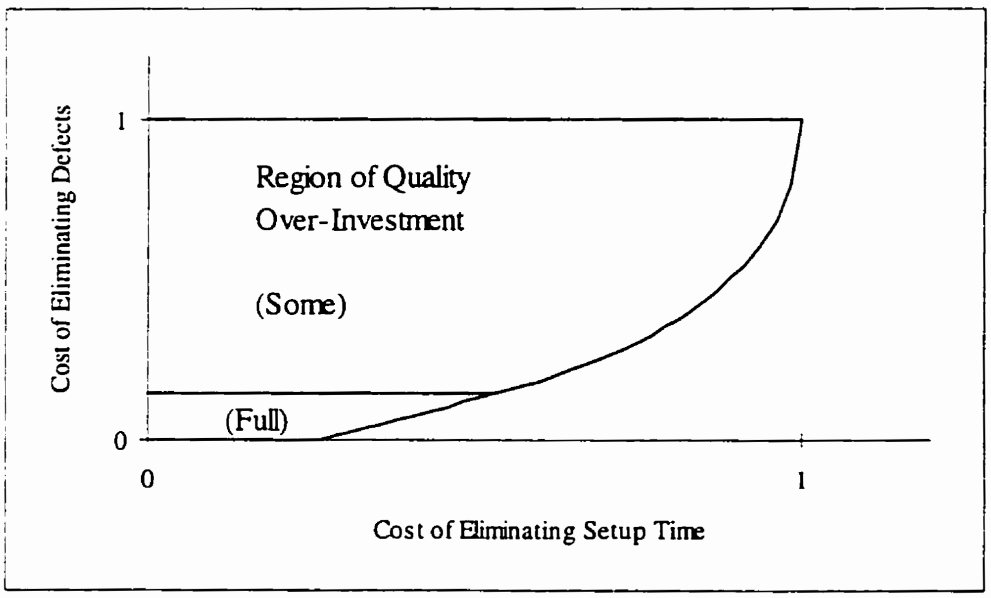


Figure 4.12: Difference in Quality Investment Decisions, 'Naive' to 'Informed' Cases

The frequencies of various levels of over-investment (i.e., for the 'naive' case relative to the 'informed' case) for this numerical example have been calculated and are summarized in Table 4.1. Once again, the levels of potential over-investment are significant. As in the model from Chapter 3 in which batch size was fixed, no situations exist in which ignoring interactions leads to under-investment.

Table 4.1 Frequencies of Over-Investment from Numerical Example

Over-Investment Level (percentage of optimal investment)	Frequency of Observances (n = 2601)
< 5%	18%
5 - 50%	31%
50 - 100%	16%
100 - 150%	5%
150 - 200%	3%
200 - 250%	3%
250 - 300%	4%
300 - 350%	5%
350 - 400%	2%
400 - 450%	1%

4.3.2 Sensitivities

The sensitivity of the distributions of over-investment to changes in Demand (Figure 4.13), Initial Defect Rate (Figure 4.1), Initial Setup Time (Figure 4.15), Holding Cost Rate (Figure 4.18), Interest Rate (Figure 4.19) and Processing Time Coefficient of Variation (Figure 4.20) has been calculated using the low and high values of each parameter given in Table 3.2. The histograms of these distributions are presented below.

Changes to the three parameters which affect system utilization, namely demand, initial defect rate and initial setup time, show little or no effect on the distributions of over-investment (Figures 4.12, 4.14 and 4.15, respectively). To investigate why, the changes in decision boundary positions have been plotted for the setup reduction boundaries (Figure 4.16) and quality improvement boundaries (Figure 4.17). These two Figures are plotted for high and low values of demand, and Figures plotted for high and low values of initial setup times and initial defect rates show similar behaviour.

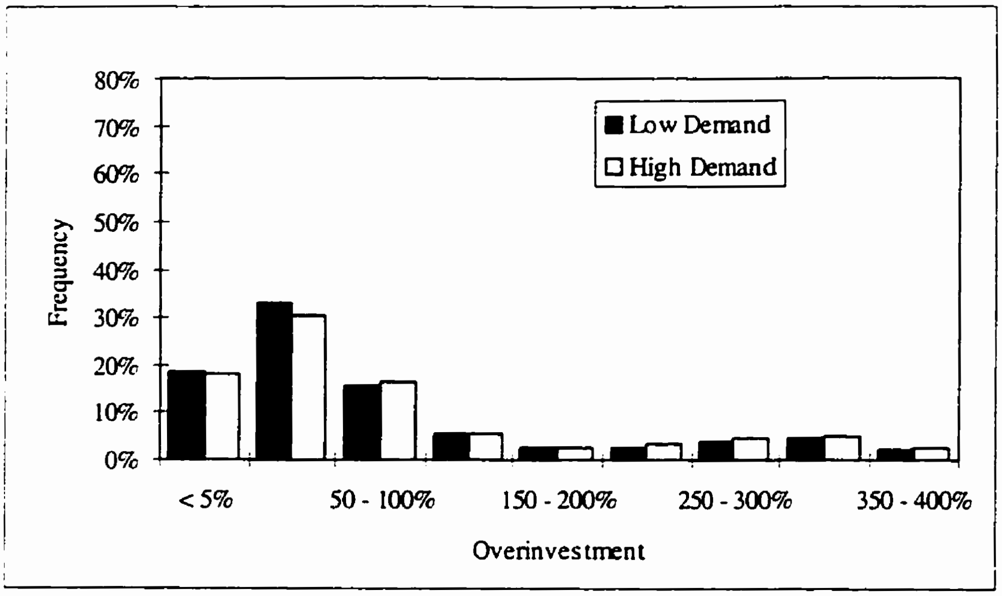


Figure 4.13: Sensitivity of Decision Error Distribution to Demand Changes

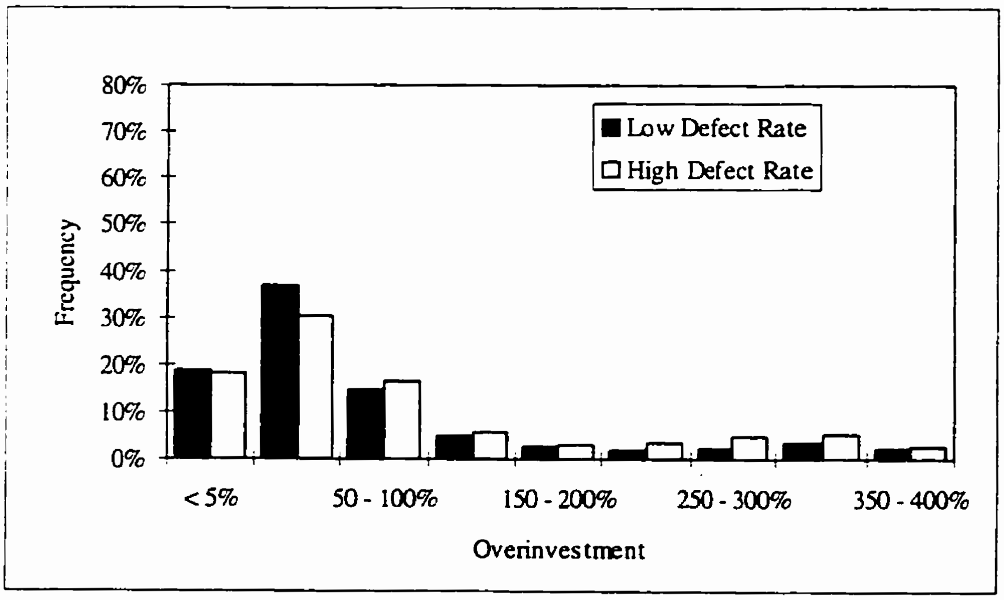


Figure 4.14: Sensitivity of Decision Error Distribution to Defect Rate Changes

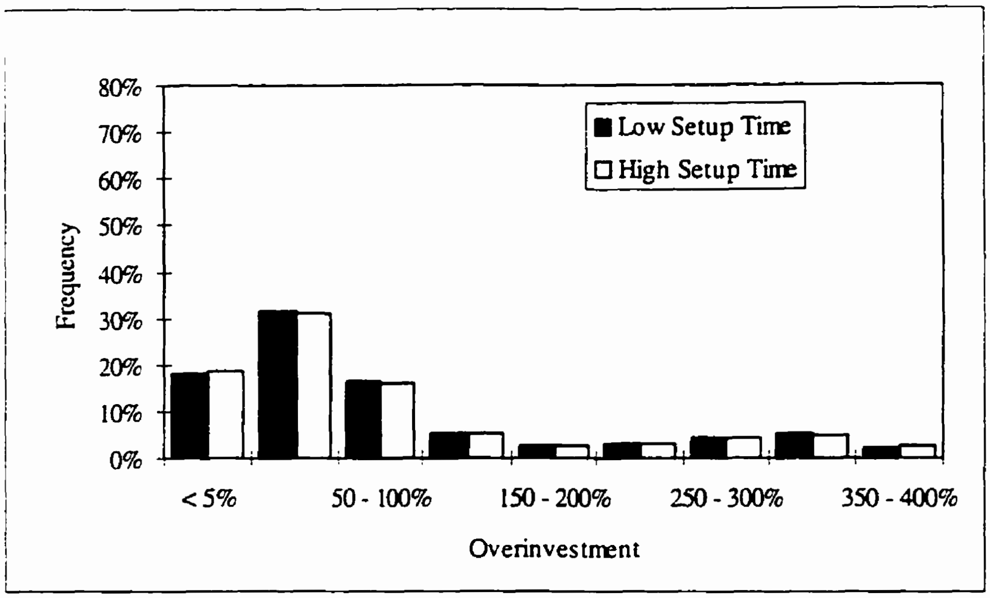


Figure 4.15: Sensitivity of Decision Error Distribution to Setup Time Changes

As shown in Figure 4.16, as demand increases, only the very bottom of the decision boundary shifts, and shifts to the left. This results in decreasing the size of the region where setup investment is called for (i.e., the upper left region) and slightly increases the size of the region where decision error takes place. Hence, for setup reduction, increasing demand leads to increased levels of decision error.

Figure 4.17 shows the changes to the investment regions for quality improvement as demand increases. There are two separate regions here where the investment decision changes. In the region on the left, the optimal decision changes from 'invest none' to 'invest fully' as demand increases. This suggests a decreased level of decision error. The other region, on the right, has the optimal investment decision go from 'invest fully' to 'invest some' as demand increases. This would suggest a higher level of decision error as demand increases. Together, it is not clear whether there will be a net increase in the over-investment distributions for quality improvement, or a net decrease as demand increases, nor is there a clear indication of how the net over-investment distribution for both setup and quality will behave as demand is changed.

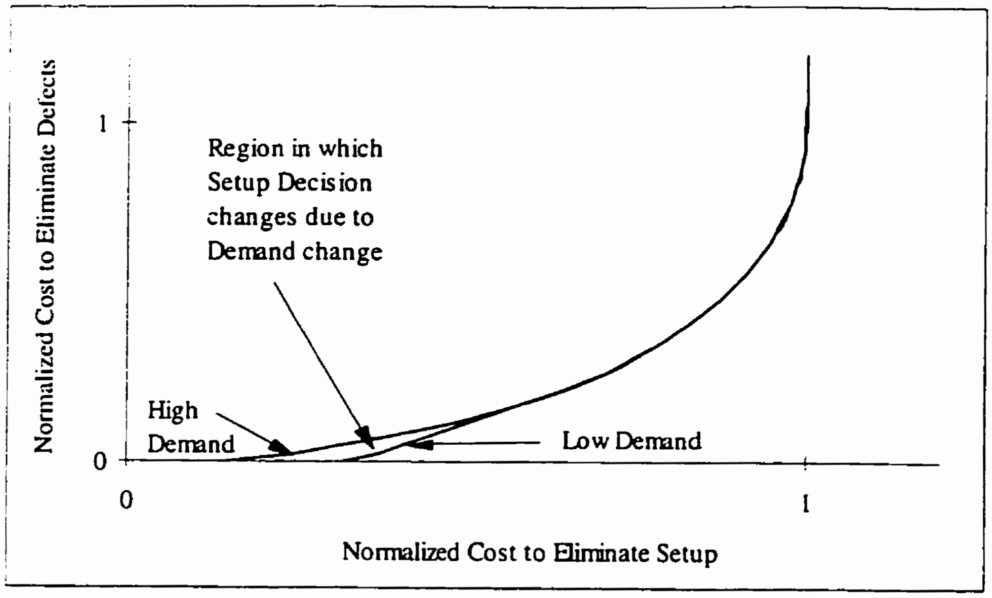


Figure 4.16: Decision Matrix for High and Low Demand Rates

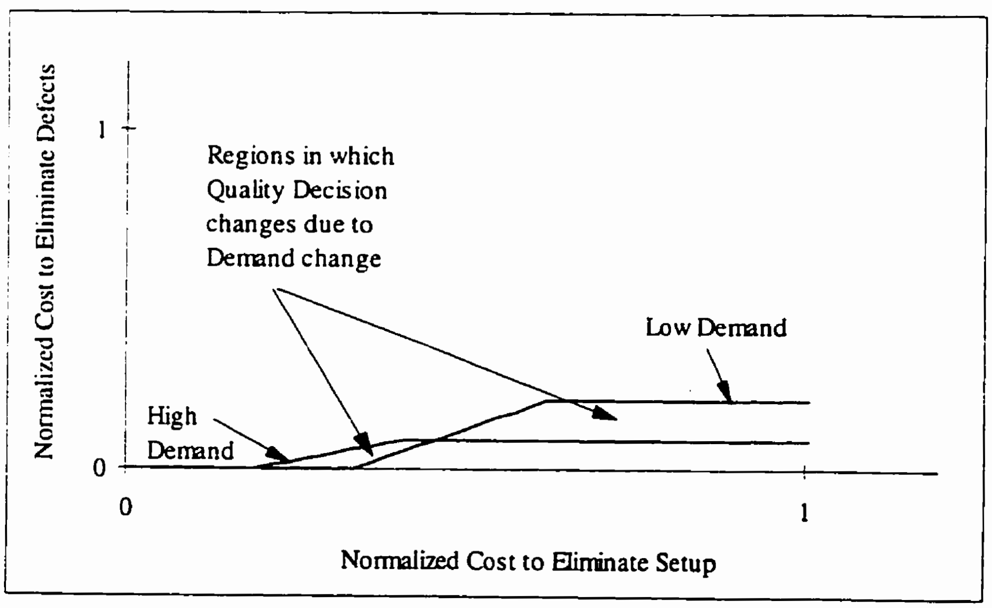


Figure 4.17: Decision Matrix for High and Low Defect Rates

A further factor to explain the lack of sensitivity is that the size of the regions where changes in optimal decisions take place is relatively small compared to the size of the decision matrix. Since the distribution of over-investments shown in the sensitivity graphs use matrix area as a measure of frequency, limited changes in graph areas will lead to limited changes in over-investment frequencies.

Since changes to demand lead to ambiguous conclusions regarding the distribution of over-investment, the lack of sensitivity to this parameter shown in the distributions of Figure 4.13 is not surprising. As well, since setup times and defect rates have a similar effect on the underlying queuing model used in this research, that is, increases in each of these three parameters lead to increased system utilization, queue lengths and WIP holding costs, the lack of sensitivity of the over-investment distributions to changes in initial defect rate (Figure 4.14) and initial setup time (Figure 4.15) are consistent with the lack of sensitivity to demand.

The sensitivity of decision errors to changes in holding cost rate and interest rate is shown in Figures 4.18 and 4.19, respectively. These graphs show a significant effect on distributions of over-investments due to the relative changes in costs of holding WIP inventories vs. the costs of making improvements as these two parameters are changed. The sensitivities shown here are consistent in direction with those of the model from Chapter 3 (Figures 3.14e and 3.14f), and for the same reasons. Just as in the results of Chapter 3, opposing sensitivities are found for holding cost rate and interest rate. These parameters affect the locations of the decision boundaries according to the ratio h/i (i.e., the ratio of holding cost rate to interest rate), so the opposite sensitivities found here are expected.

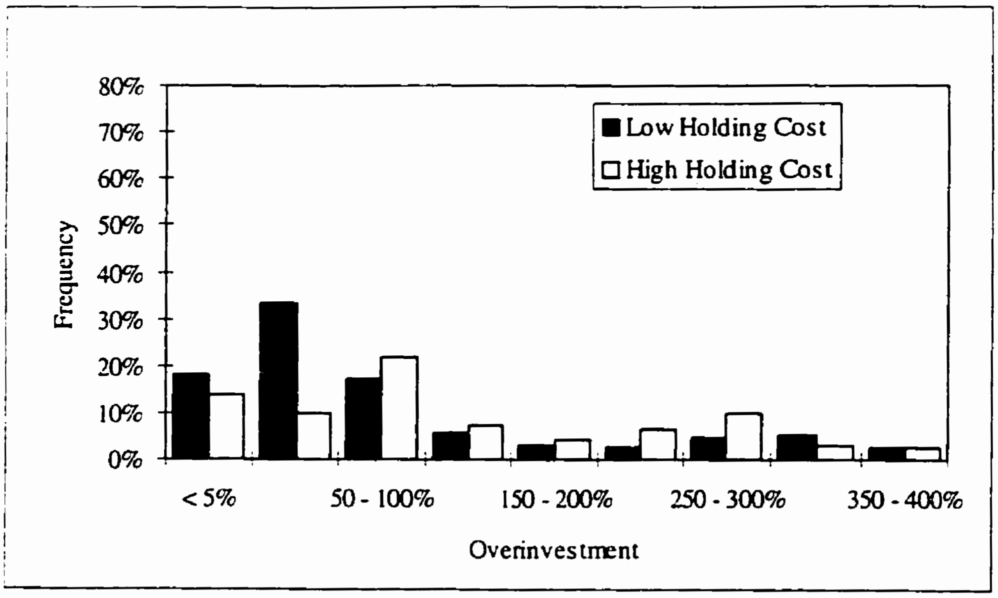


Figure 4.18: Sensitivity of Decision Error Distribution to Holding Cost Changes

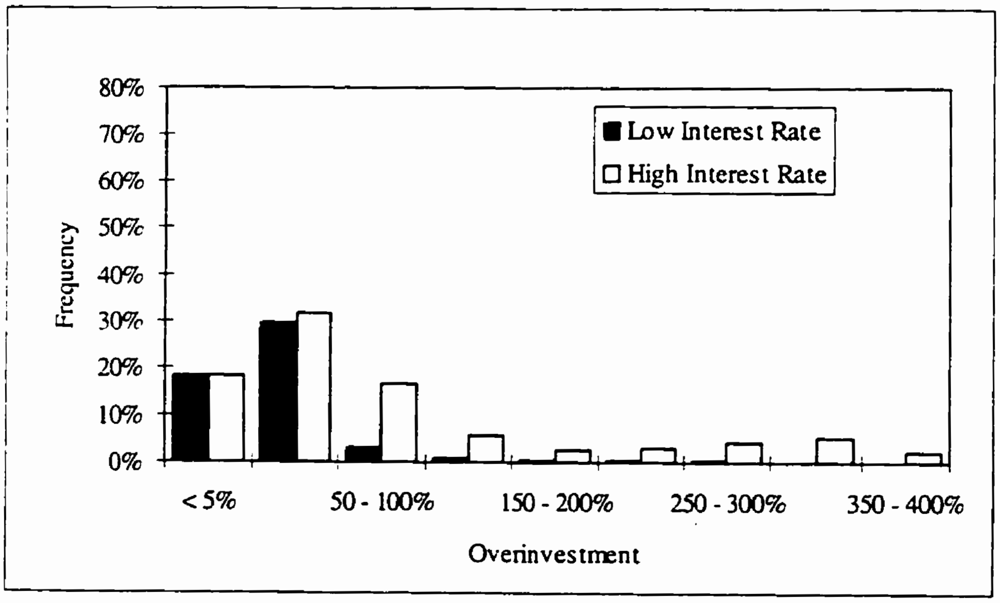


Figure 4.19: Sensitivity of Decision Error Distribution to Interest Rate Changes

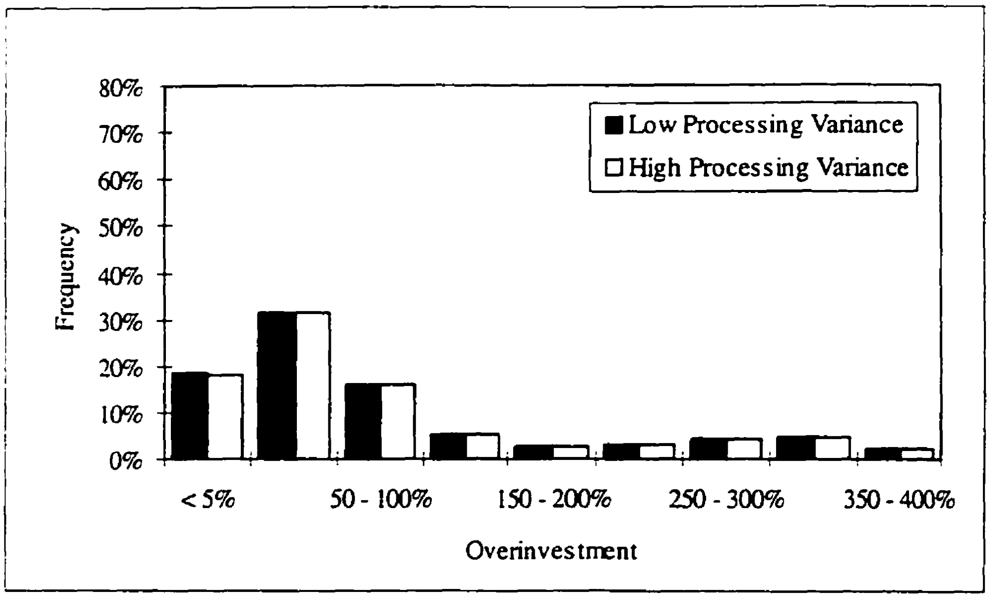


Figure 4.20: Sensitivity of Decision Error Distribution to Processing Variance Changes

The final sensitivity case studied is that for processing time variance, with the distribution of over-investments shown in Figure 4.20. Little or no sensitivity is seen in this graph. There are two explanations for this result, both having to do with the behaviour of batch size over the decision matrix. When the cost of setup reduction is relatively high, no setup reduction takes place (e.g., the lower right region of the decision matrix) and optimal batch size is relatively large. Under this condition, batch service time coefficient of variation becomes quite small, as was demonstrated in Equations 3.6.3 through 3.6.6 of the last chapter.

The other region of the decision matrix, where investment takes place in setup reduction, has an optimal batch size of one unit. Under this condition, queue length will be a function of batch processing time variance, but WIP costs are minimal since each batch of work in the queue represents only a single unit. These minimal changes to WIP cost will lead to minimal changes in optimal investment decisions as a function of processing time variance, and hence, little or no sensitivity as is shown in Figure 4.20.

4.4 Discussion

In this chapter, the expected total cost (E[TC]) model of Chapter 3 was extended to include a third decision variable: batch size. Empirical evidence suggests that batch size changes are strongly linked to changes in setup time, and since batch size changes are assumed to not require capital investment, the batch size variable was optimized with the setup time variable.

Again, two investment strategies were examined: the 'naive' strategy in which optimal investment decisions were calculated for setup improvement and quality improvement independently of each other, and the 'informed' strategy in which investment levels in setup and quality were optimized simultaneously.

In the 'naive' case, the E[TC] function was shown to be always non-convex in terms of the setup and batch size decision variables. Conventional convex optimization methods were not applicable. However, through the development of a series of Properties, it was shown that this system has two optimal decisions: 1) 'Don't Invest', in which no investment is made in setup reduction and an optimal batch size, $Q^* > 1$ (in general) results, or, 2) 'Invest', in which enough investment is made in setup reduction to reduce the optimum batch size to one unit. In numerical experiments, it was shown that in the 'invest' case, setup time was essentially reduced to zero, i.e., the optimal decision is to invest fully. It may be noted here also that E[TC] is convex in Q alone. Optimal 'naive' investment in quality is the same as in Chapter 3, that is, optimal decisions can be 'invest none', 'invest some', or, 'invest fully', depending on the parameter values in the particular problem.

The 'informed' case optimization was similarly challenging due to the continued non-convexity of the E[TC] function. A 'piece-wise' approach was taken to develop the decision matrix and closed-form expressions for the location of the decision boundaries could not be obtained in all instances due to the increasingly complex mathematics of the system.

Comparison of the two cases shows that, just as in Chapter 3, the improvement practices interact with each other, and over-investments of significant magnitude can result if this information is not taken into account. Both the 'naive' and 'informed' cases show binary investment in setup reduction: either no investment takes place, or effectively full investment is called for. Batch size shows similar behaviour, either batch size is some significant quantity, or it is one unit.

Decisions to invest in quality improvement were similarly extreme. If investment in setup was called for, no investment would take place in quality, unless the cost of making quality improvements was very close to zero.

This either/or investment behaviour is not intuitive, and would not be expected in practice. The explanation is that the linear investment function tends to bias the system towards non-convexity and extreme point optima. A similar system with a convex investment function is studied in Chapter 6, and depending upon the degree of convexity of the investment function, does not show such extreme behaviour.

It was also found that for the entire area of the decision matrix where the 'naive' strategy calls for investment in both practices, over-investment occurred in either setup or quality improvement. There is no region where the 'naive' and 'informed' strategies call for the same optimal investment levels.

Distributions of over-investments when interactions are ignored were found to be essentially insensitive to changes in system demand, initial setup time, initial defect rate and processing time variance. The insensitivity to processing time variance is a result of reduction of batch processing time coeffient of variation for larger batch sizes, and due to there being very little queue and hence very little effect on the system when batch size is one unit. Changes to the other parameters, demand and initial setup and quality levels do affect the location of the boundaries where investment decisions change, but when scales are normalized (as is the case in the comparison of over-investment levels), the decision matrices show little sensitivity to changes in these parameters (i.e., the shapes of the decision matrices do not change very much).

Over-investment distributions are sensitive, however, to changes in the holding cost and interest rates, just as was found in Chapter 3.

In the next chapter, the model of Chapter 3 is extended to include a convex investment function, with fixed batch size, while the effects of variable batch size are again examined in Chapter 6, which also uses a convex investment function.

Chapter 5: Fixed Batch Size, Convex Investment Function

In this chapter, a system characterized by fixed batch size and a convex investment-improvement function is examined. Intuitively, a convex investment function is more appealing than the linear investment function since it represents the situation where each increment of improvement to the production system requires ever-increasing increments of investment. Decision variables, once again, are setup time (T) and quality level (R).

5.1 Convex Investment Function

As was described in Chapter 2, a number of investment functions have been used in the literature. Table 2.2 provided a summary. The choice of the functional form used seems to be related primarily to the solution method used. For instance, models which were solved via geometric programming all used a power function form for the investment function. In fact, with the exception of a step-function used by Trevino et al. [1993], none of the papers listed in Table 2.2 refer to any empirical measurements of setup or quality improvement costs as a basis for selection of the forms of investment functions used. There appears to be no compelling reason to select one functional form over another.

Since there is no 'best' investment function, a search can be made for a function with the most useful properties. To start this search, the following criteria are specified for an acceptable convex investment function, I(T,R):

1) Non-negativity, i.e.,

$$I \geq 0 \qquad \forall \; 0 \leq T, R \leq 1 \qquad (5.1.1)$$

2) Zero initial investment, i.e.,

$$I = 0 \qquad \text{for } T, R = 1 \qquad (5.1.2)$$

3) Positive improvements from investment, i.e.,

$$\frac{\partial I}{\partial T} < 0, \quad \frac{\partial I}{\partial R} < 0 \qquad \forall \ 0 \leq T, R \leq 1 \qquad (5.1.3)$$

4) Strict convexity, i.e.,

$$\frac{\partial^2 I}{\partial T^2} > 0, \quad \frac{\partial^2 I}{\partial R^2} > 0 \qquad \forall \ 0 \leq T, R \leq 1 \qquad (5.1.4)$$

In addition to meeting these criteria, an investment function should keep the mathematical complexity of the overall expected total cost function to a minimum so that the likelihood of obtaining useful analytic solutions is increased. To examine this issue, the E[TC] function of Equation 3.3.3 is reproduced here:

$$E[TC] = h\, D\, E[W] + i\,(I_S + I_Q) \qquad (5.1.5)$$

In order to minimize E[TC], the first derivatives of this function are taken and set to zero. For the purposes of brevity, only the derivation in terms of the setup time decision variable will be given in this section; derivation in terms of quality level is similar. The derivative with respect to setup time becomes:

$$\frac{\partial E[TC]}{\partial T} = hD\frac{\partial E[W]}{\partial T} + i\frac{\partial I_S}{\partial T} = 0$$

Substitution of the expected waiting time function, E[W], and simplifying gives:

$$2\lambda(1-\lambda E[S])E[S] + \lambda^2 E[S^2] + (1-\lambda E[S])^2 + \frac{i}{\tau_0 hD}(1-\lambda E[S])^2 \frac{\partial I_S}{\partial T} = 0 \qquad (5.1.6)$$

E[S] is a first order polynomial in terms of T (from Equation 3.2.1), while $E[S^2]$ is a second order polynomial in T (from Equation 3.2.3). Examination of the first three terms in 5.1.6 shows that they are second order polynomials in T, while the degree of the last term will depend upon the specific investment function chosen. Since the coefficient of that term is already a second order polynomial, if the derivative of the investment function is not 'zeroth' order in T, the complexity of the whole equation increases. To illustrate this idea, suppose that $\partial I_S/\partial T$ had the form $a_0 + a_1 T$ (where a_0, a_1 are arbitrary coefficients). The last term in 5.1.6 would then be a third order polynomial in T, requiring solution of a cubic equation. As the order of $I_S(T)$ increases, so does the order of $\partial E[S]/\partial T$ and $\partial E[TC]/\partial T$.

In light of these considerations for investment function behaviour, the following form for an investment function (in this case for setup reduction) is proposed:

$$I_S = \frac{a_0 + a_1 T + a_2 T^2}{1-\rho} \qquad (5.1.7)$$

where a_0, a_1, a_2 are arbitrary constants.

Taking the derivative and substituting into 5.1.6 gives:

$$2\lambda(1-\lambda E[S])E[S] + \lambda^2 E[S^2] + (1-\lambda E[S])^2 + \frac{i}{\tau_0 hD}\Big[(a_1 + 2a_2 T)(1-\lambda E[S]) + \lambda\tau_0(a_0 + a_1 T + a_2 T^2)\Big] = 0$$

Examination of this equation shows that it is a second order polynomial in T. Thus, the investment function 5.1.7 does not increase the order of the derivative of the E[TC] system, or the difficulty in obtaining a solution.

The form of this function is illustrated in Figure 5.1. The parameter $\frac{a_0}{1-\rho}$ determines the location of the intercept of the curve with the investment axis, and represents the amount of investment necessary to completely eliminate setup time. It is analogous to the a/τ_0 quantity associated with the linear investment function of the previous chapters. The convexity of the investment function is controlled by the magnitude of the a_1 parameter relative to the a_0 parameter.

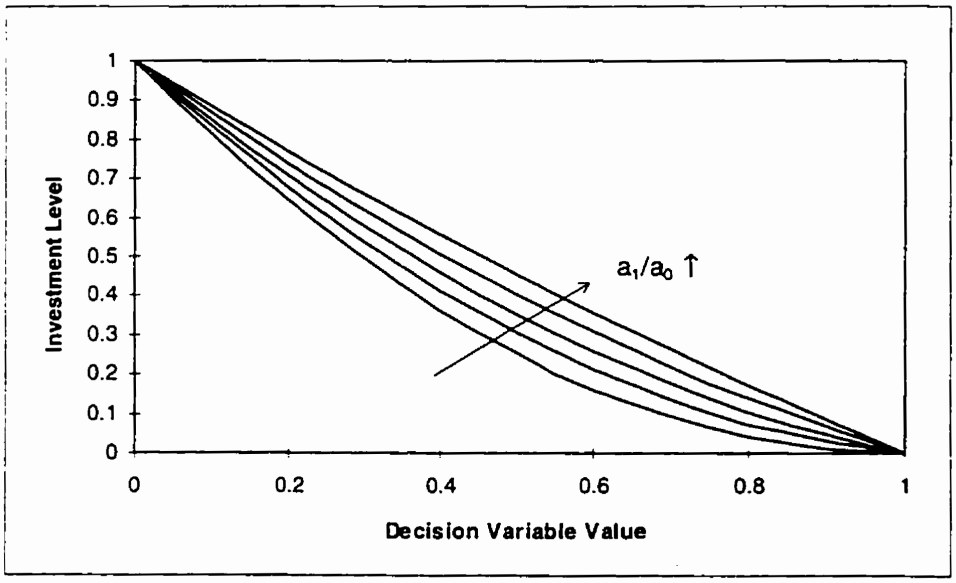


Figure 5.1: Convexity of Investment Function by value of a_1/a_0

This investment function form is influenced by the utilization, ρ, of the production system. For instance, as utilization increases, the level of investment necessary to make improvements also increases, as shown in Figure 5.2. Intuitively, this behavior can be justified in the following way. In a setup reduction project, typical tasks would be activities such as modifying tooling to provide standardized clamping or having operators rehearse better setup procedures. In a system with low utilization, making the system available for physical modifications as well as operator practice sessions is not difficult. However, with higher utilization improvement tasks would increasingly involve conflicts with production and would require additional resources or overtime, each representing higher costs, to complete the same level of improvements in a similar time frame. A similar type of behaviour was found by Kim [1990] in an EPQ-based model. That model predicted investment in setup reduction becoming increasingly more difficult to justify as server utilization increased. Thus, an investment function dependent upon system utilization is seen as reasonable.

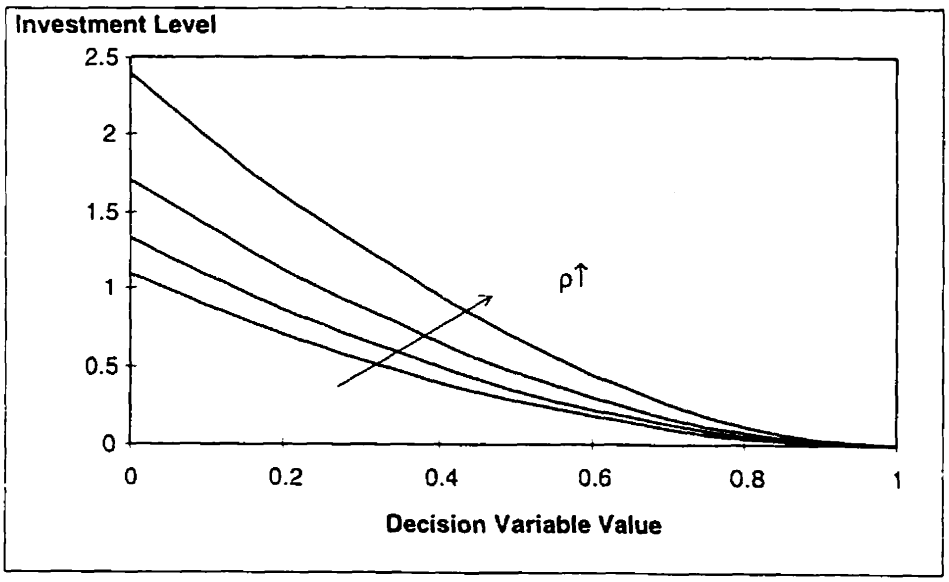


Figure 5.2: Dependence of Investment Function on System Utilization

The coefficients a_0, a_1, a_2 in 5.1.7 must be selected such that the criteria for an acceptable investment function (i.e., Equations 5.1.1 through 5.1.4) are satisfied. The constraints on these coefficients are developed here with the help of the four criteria:

1) Non-negativity, i.e.,

$I \geq 0$

$$\Rightarrow \quad a_0 + a_1 T + a_2 T^2 \geq 0 \qquad \forall \; 0 \leq T \leq 1 \tag{5.1.8}$$

since $(1 - \rho) > 0$

$$\Rightarrow \quad a_0 > 0 \qquad \text{when } T = 0 \tag{5.1.9}$$

2) Zero initial investment, i.e.,

$I = 0 \quad \text{for } T = 1$

$$\Rightarrow \quad a_0 + a_1 + a_2 = 0$$

$$\text{or } a_2 = -(a_0 + a_1) \tag{5.1.10}$$

3) Positive improvements from investment, i.e.,

$$\frac{\partial I}{\partial T} < 0$$

Let $Z = 1 - \lambda(1 + Rr_0)QE[X] \;\Rightarrow\; (1 - \rho) = (Z - \lambda\tau_0 T)$; then,

$$\frac{\partial I_S}{\partial T} = \frac{\left[a_1 - 2(a_0 + a_1)T\right](Z - \lambda\tau_0 T) + \left[a_0 + a_1 T - (a_0 + a_1)T^2\right]\lambda\tau_0}{(1-\rho)^2} < 0$$

$$\Rightarrow \quad a_1\left[Z - 2ZT + \lambda\tau_0 T^2\right] < a_0\left[\lambda\tau_0 - 2ZT + \lambda\tau_0 T^2\right] \tag{5.1.11}$$

This expression relates the values of the coefficients a_0 and a_1. To refine this relationship, consider the quantity $\left[Z - 2ZT + \lambda\tau_0 T^2\right]$:

$$\left[Z - 2ZT + \lambda\tau_0 T^2\right] = \begin{cases} Z \quad (> 0) & \text{for } T = 0 \\ -Z + \lambda\tau_0 \quad (< 0) & \text{for } T = 1 \end{cases}$$

This suggests that there is at least one sign change in the range [0,1]. To find the location of the zero of $\left[Z - 2ZT + \lambda\tau_0 T^2\right]$, denoted $\hat{T}$, the quadratic formula is used to find a single root in [0,1] at:

$$\hat{T} = \frac{Z}{\lambda\tau_0}\left(1 - \sqrt{1 - \frac{\lambda\tau_0}{Z}}\right)$$

Thus,

$$\left[Z - 2ZT + \lambda\tau_0 T^2\right] \begin{cases} > 0 & \text{for } 0 \leq T < \hat{T} \\ < 0 & \text{for } \hat{T} < T \leq 1 \end{cases}$$

and Equation 5.1.11 can be rearranged in terms of a_1/a_0 as:

$$\frac{a_1}{a_0} \begin{cases} < -\dfrac{\left[\lambda\tau_0 - 2ZT + \lambda\tau_0 T^2\right]}{\left[Z - 2ZT + \lambda\tau_0 T^2\right]} & \text{for } 0 \leq T < \hat{T} \\ > -\dfrac{\left[\lambda\tau_0 - 2ZT + \lambda\tau_0 T^2\right]}{\left[Z - 2ZT + \lambda\tau_0 T^2\right]} & \text{for } \hat{T} < T \leq 1 \end{cases} \tag{5.1.12}$$

Equation 5.1.12 forms a constraint on the value a_1/a_0. This constraint is sketched over the range $0 \leq T \leq 1$ in Figure 5.3.

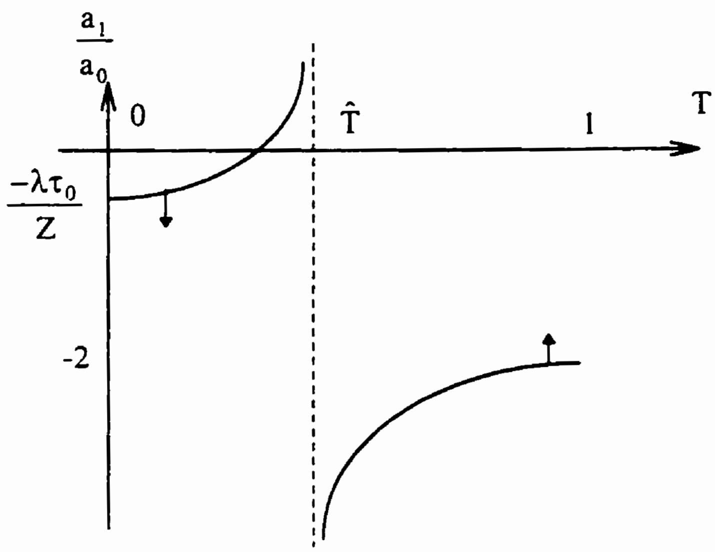


Figure 5.3: Constraints on $\frac{a_1}{a_0}$ from Positive-Return-from-Investment Criteria

Since the variable T can vary over the range [0,1], and the constraints 5.1.12 must be satisfied for all values of T, the term a_1/a_0 must satisfy:

$$\frac{a_1}{a_0} \begin{cases} < -\dfrac{\lambda\tau_0}{Z} & \forall\ 0 \le T < \hat{T} \\ > -2 & \forall\ \hat{T} < T \le 1 \end{cases} \tag{5.1.13}$$

These expressions provide both an upper and lower bound on the value of a_1/a_0.

4) Strict convexity, i.e.,

$$\frac{\partial^2 I}{\partial T^2} > 0 \Rightarrow$$

$$\frac{\partial^2 I}{\partial T^2} = \frac{-2(a_0 + a_1)}{(1-\rho)} + \frac{2\lambda\tau_0\left[a_1 - 2(a_0 + a_1)T\right]}{(1-\rho)^2} + \frac{2\lambda^2\tau_0^2\left[a_0 + a_1 T - (a_0 + a_1)T^2\right]}{(1-\rho)^2}$$

implies that

$$(a_0 + a_1)(Z - \lambda\tau_0 T)^2 + \lambda\tau_0\left[a_1 - 2(a_0 + a_1)T\right](Z - \lambda\tau_0 T)$$
$$+ \lambda^2\tau_0^2\left[a_0 + a_1 T - (a_0 + a_1)T^2\right] \quad > 0$$

or,

$$\frac{a_1}{a_0} < -\frac{Z^2 - \lambda^2\tau_0^2}{Z(Z - \lambda\tau_0)} \tag{5.1.14}$$

Both this constraint and 5.1.13 impose an upper bound on a_1/a_0. A question arises of which (if either) constraint is tighter, or, alternately, which (if either) of these upper bounds is redundant. To answer this question, consider the upper bound constraint from 5.1.13, along with some algebraic rearrangements:

$$
\begin{aligned}
\frac{a_1}{a_0} &< -\frac{\lambda\tau_0}{Z} \\
&= -\frac{\lambda\tau_0(Z - \lambda\tau_0)}{Z(Z - \lambda\tau_0)} \\
&= -\frac{\lambda\tau_0 Z - \lambda^2\tau_0^2}{Z(Z - \lambda\tau_0)} \\
&= -\frac{Z^2 - \lambda^2\tau_0^2}{Z(Z - \lambda\tau_0)} + \frac{Z^2 - \lambda\tau_0 Z}{Z(Z - \lambda\tau_0)} \\
&= -\frac{Z^2 - \lambda^2\tau_0^2}{Z(Z - \lambda\tau_0)} + 1
\end{aligned}
$$

The first term in the last line will be recognized as the right hand side of 5.1.14. Thus, the constraint 5.1.14 is tighter than 5.1.13 by one unit. Therefore, the convexity condition (i.e., Equation 5.1.14) will impose a tight upper bound on a_1/a_0 and the upper bound from 5.1.13 is redundant.

To summarize, a satisfactory investment-improvement function for setup reduction is given by:

$$
I_S = \frac{a_0 + a_1 T - (a_0 + a_1) T^2}{1 - \rho} \tag{5.1.15}
$$

$$
\begin{aligned}
\text{s.t.} \quad & -2 < \frac{a_1}{a_0} < -\frac{Z^2 - \lambda^2\tau_0^2}{Z(Z - \lambda\tau_0)} \\
& a_0 > 0 \\
& \text{where} \quad Z = 1 - \lambda(1 + Rr_0)QE[X]
\end{aligned}
$$

By a completely analogous procedure, a convex investment function for quality improvement, I_Q, which exhibits the same behavior as shown in Figures 5.2 can be developed as:

$$I_Q = \frac{b_0 + b_1 R - (b_0 + b_1) R^2}{1 - \rho} \tag{5.1.16}$$

$$\text{s.t.} \quad -2 < \frac{b_1}{b_0} < -\frac{Y^2 - \lambda^2 r_0^2 Q^2 E[X]^2}{Y(Y - \lambda r_0 QE[X])}$$

$$b_0 > 0$$

$$\text{where} \quad Y = 1 - \lambda(T\tau_0 + QE[X])$$

With these convex investment functions, the E[TC] function is complete and will now be used to analyse the 'naive' and 'informed' investment cases.

5.2 'Naive' Investment Decisions

Once again, the 'naive' investment decision case is the one in which investments in setup reduction and quality improvement are made independently of each other. For each practice, optimal investments are calculated while assuming no investment takes place in the other.

Then, for setup time, E[TC] is minimized with respect to T through the necessary condition:

$$\left.\frac{\partial E[TC]}{\partial T}\right|_{R=1} = 0 \quad \Rightarrow$$

$$T^* = \frac{1 - D(1 + r_0)E[X]}{\lambda \tau_0} + \frac{1}{D\tau_0}\sqrt{\frac{Q}{hD^2\tau_0^2 + 2Qi(a_0 + a_1)}} \times$$

$$\Big[\; 2Q^2 i(a_0 + a_1)[1 - D(1 + r_0)E[X]]^2 - D\tau_0\big(hD\tau_0 Q + 2D\tau_0 i a_0 + 2Qia_1\big) \qquad (5.2.1)$$

$$+hD^4\tau_0^2\Big[(1 + r_0^2)E[X]^2 - (1 + r_0)E[X^2]\Big] + 2D^2(1 + r_0)QE[X]\tau_0 i a_1 \;\Big]^{\frac{1}{2}}$$

The sufficient condition is checked by substituting 5.2.1 into the second derivative of E[TC]:

$$\left.\frac{\partial^2 E[TC]}{\partial T^2}\right|_{T^*} = \sqrt{\frac{\left[hD^2\tau_0^2 + 2Qi(a_0 + a_1)\right]^3}{Q}} \times \Big[\; 2Q^2 i(a_0 + a_1)[1 - 2D(1 + r_0)E[X]]$$

$$- D^2\tau_0^2\big(hQ + 2i(a_0 + a_1)\big) + hD^4\tau_0^2\Big[(1 + r_0^2)E[X]^2 - (1 + r_0)E[X^2]\Big]$$

$$+2D^2(1 + r_0)QE[X]\tau_0 i a_1 + 2D^2Q^2E[X]^2 i(a_0 + a_1)(2 + r_0)^2 \;\Big]^{-\frac{1}{2}}$$

$$\geq 0$$

So that 5.2.1 is the setup time which minimizes E[TC].

The decision boundaries are found in terms of a_0, which reflects the cost of eliminating all setup time in the convex investment function of Equation 5.1.15. (This value is analogous to the value τ_0/a used in the linear investment function of Chapters 3 and 4.) Rearranging 5.2.1 for a_0* gives:

$$a_0{}^* = \frac{\left(D^2\tau_0^2 h + 2Qia_1\right)}{2Qi} + \frac{1}{2Qi\left[D\tau_0 T^{*2} + 2Q\left\{D(1+r_0)E[X]-1\right\}T^* + D\tau_0\right]}\Big\{$$
$$-2Q^2\left(D(1+r_0)E[X]-1\right)\left(D\tau_0 h + ia_1\right) - D\tau_0\left(D^2\tau_0^2 h + 2Qia_1\right) \tag{5.2.2}$$
$$+D^3\tau_0 hQ\left[(1+r_0)E[X^2] + \left((Q-1)(1+r_0)^2 + 2(1+r_0) - 2\right)E[X]^2\right]\Big\}$$

The 'no-invest' boundary occurs when T* = 1, or,

$$a_0{}^*\Big|_{T^*=1} = \frac{\left(D^2\tau_0^2 h + 2Qia_1\right)}{2Qi} + \frac{1}{4Qi\left[D\tau_0 + Q\left\{D(1+r_0)E[X]-1\right\}\right]}\Big\{$$
$$-2Q^2\left(D(1+r_0)E[X]-1\right)\left(D\tau_0 h + ia_1\right) - D\tau_0\left(D^2\tau_0^2 h + 2Qia_1\right) \tag{5.2.3}$$
$$+D^3\tau_0 hQ\left[(1+r_0)E[X^2] + \left((Q-1)(1+r_0)^2 + 2(1+r_0) - 2\right)E[X]^2\right]\Big\}$$

While the 'full-invest' boundary occurs at T* = 0, or

$$a_0{}^*\Big|_{T^*=0} = \frac{\left(D^2\tau_0^2 h + 2Qia_1\right)}{2Qi} + \frac{1}{2QiD\tau_0}\Big\{$$
$$-2Q^2\left(D(1+r_0)E[X]-1\right)\left(D\tau_0 h + ia_1\right) - D\tau_0\left(D^2\tau_0^2 h + 2Qia_1\right) \tag{5.2.4}$$
$$+D^3\tau_0 hQ\left[(1+r_0)E[X^2] + \left((Q-1)(1+r_0)^2 + 2(1+r_0) - 2\right)E[X]^2\right]\Big\}$$

Optimal investment in quality is found similarly. The necessary condition on E[TC] with respect to R is:

$$\left.\frac{\partial E[TC]}{\partial R}\right|_{T=1} = 0 \quad \Rightarrow$$

$$\begin{aligned}R^* = {} & \frac{1-\lambda\tau_0}{Dr_0E[X]} + \frac{1}{DQr_0E[X]}\frac{1}{\sqrt{hD^2(Q+1)r_0^2E[X]^2 + 2i(b_0+b_1)}} \times \\ & \Big[\; 2i(b_0+b_1)[Q-D\tau_0]^2 - D^3hr_0^2QE[X][Q-D\tau_0]\big[E[X^2]+2E[X]^2\big] \\ & -2DQiE[X]\Big\{(b_0+b_1)[2Q-2D\tau_0-DQE[X]] + b1[Q-D\tau_0-DQE[X]]\;\Big\} \\ & +D^2hr_0^2E[X]^2\Big[[Q-D\tau_0]^2 + 2D^2Q^2E[X]^2 - Q^3\Big] - 2D^2Q^2E[X]^2r_0^2ib_0\;\Big]^{\frac{1}{2}}\end{aligned} \tag{5.2.5}$$

In the convex investment function for quality improvement (Equation 5.1.16), b_0 represents the cost of eliminating defects, analogous to the r_0/b term of the previous two chapters. It is found by rearranging 5.2.5 as:

$$\begin{aligned}b_0^* = {} & \frac{1}{i}\Big[2ib_1 + D^2r_0^2E[X]^2h(Q+1)\Big] + \\ & \frac{1}{i\Big\{DQr_0E[X]R^{*2} + 2\big[D\tau_0 + DQE[X] - Q\big]R^* + DQr_0E[X]\Big\}}\Big\{ \\ & 2\big[D\tau_0 + DQE[X] - Q\big]\big[D^2hr_0\tau_0E[X] - ib_1\big] - 2DQr_0E[X]ib_1 \\ & +hDr_0E[X]\Big[2Q^2 - D^2\tau_0^2 - 2DQ^2E[X] + D^2E[X]^2\{(Q-1)+(Q+1)r_0\}\Big]\Big\}\end{aligned} \tag{5.2.6}$$

The 'no-invest' boundary occurs at $R^* = 1$, or:

$$b_0 *\big|_{R^*=1} = \frac{1}{i}\left[2ib_1 + D^2 r_0^2 E[X]^2 h(Q+1)\right] + \frac{1}{2i\left\{DQr_0E[X] + \left[D\tau_0 + DQE[X] - Q\right]\right\}} \Big\{ 2\left[D\tau_0 + DQE[X] - Q\right]\left[D^2 hr_0\tau_0E[X] - ib_1\right] - 2DQr_0E[X]ib_1 + hDr_0E[X]\left[2Q^2 - D^2\tau_0^2 - 2DQ^2E[X] + D^2E[X]^2\left\{(Q-1) + (Q+1)r_0\right\}\right] \Big\} \tag{5.2.7}$$

And the 'full-invest' boundary occurs at $R^* = 0$, or:

$$b_0 *\big|_{R^*=0} = \frac{1}{i}\left[2ib_1 + D^2 r_0^2 E[X]^2 h(Q+1)\right] + \frac{1}{iDQr_0E[X]} \Big\{ 2\left[D\tau_0 + DQE[X] - Q\right]\left[D^2 hr_0\tau_0E[X] - ib_1\right] - 2DQr_0E[X]ib_1 + hDr_0E[X]\left[2Q^2 - D^2\tau_0^2 - 2DQ^2E[X] + D^2E[X]^2\left\{(Q-1) + (Q+1)r_0\right\}\right] \Big\} \tag{5.2.8}$$

The 'naive' case boundaries (Equations 5.2.3, 5.2.4, 5.2.7 and 5.2.8) can then be plotted on a decision matrix, as shown in Figure 5.4. It may be noted that this 'naive' case decision matrix is similar to that of Chapter 3 where a linear investment function was used. The idea of the sensitivity of decision matrices to the assumed investment function will be discussed in greater detail in Chapter 7.

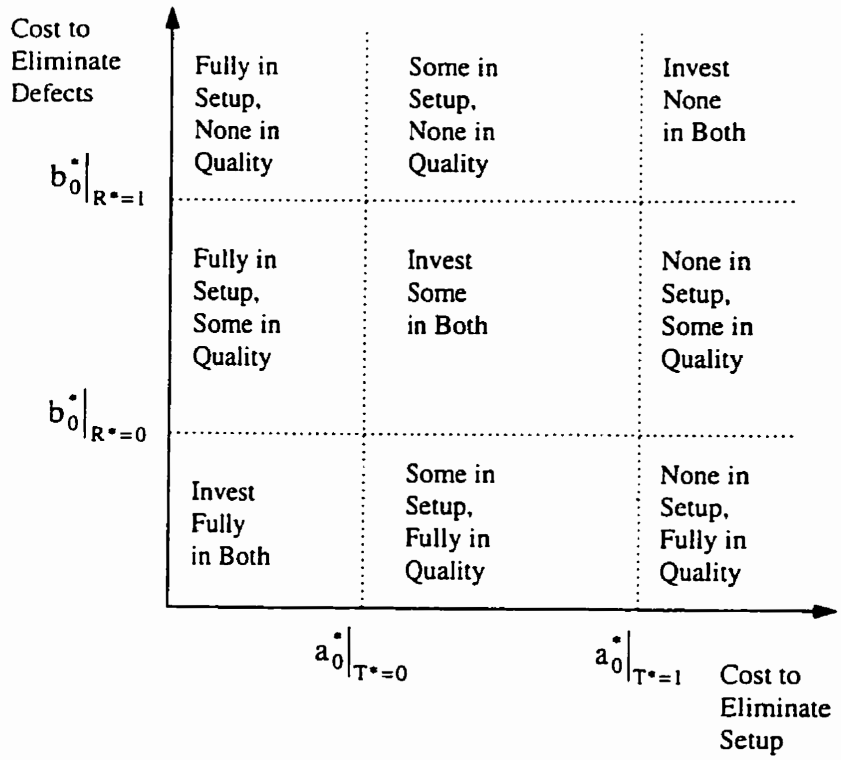


Figure 5.4: Schematic of Decision Matrix for 'Naive' Investment Case.

5.3 'Informed' Investment Decisions

In the 'informed' case, improvements to setup and quality are considered simultaneously. The objective is then to minimize E[TC] as a bivariate function of setup time and defect rate. In this section the convexity of the E[TC] function is examined, then an appropriate optimization method is used to determine where the decision boundaries lie.

5.3.1 Convexity of E[TC]

In minimizing the expected total cost of the system, it is necessary to determine the conditions, if any, under which the E[TC] function is convex, and where locally optimal solutions might exist. The first step taken here will be to test the convexity of the E[TC] function in two variables, setup time (T) and quality level (R). There is some degree of choice of convexity of the investment function, governed by the selection of the values of the a_1/a_0 and b_1/b_0 parameters. Since the object of this chapter is to study the behaviour of the E[TC] system with a convex investment function, these parameters will be set to the level which maximizes the convexity of the investment functions. This level was determined as the lower bound in Equation 5.1.13, i.e., that the investment functions are most convex when a_1/a_0 and b_1/b_0 are set to the value -2. Using these values, the E[TC] function is:

$$E[TC] = hD\left(\frac{\lambda E[S^2]}{2(1-\lambda E[S])} + E[S]\right) + i\left(\frac{a_0(1-T)^2}{(1-\lambda E[S])} + \frac{b_0(1-R)^2}{(1-\lambda E[S])}\right) \tag{5.3.1}$$

The Hessian of this system becomes:

$$H(E[TC]) = \begin{bmatrix} \dfrac{\partial^2 E[TC]}{\partial T^2} & \dfrac{\partial^2 E[TC]}{\partial T \partial R} \\ \dfrac{\partial^2 E[TC]}{\partial T \partial R} & \dfrac{\partial^2 E[TC]}{\partial R^2} \end{bmatrix}$$

Convexity requires the principal minors of this system to be positive, i.e.,

$$\Delta_1 = \frac{\partial^2 E[TC]}{\partial T^2} \quad > 0$$

$$\Delta_2 = \det H(E[TC]) \quad > 0$$

Inserting the E[TC] function , the first principal minor becomes:

$$\Delta_1 = \frac{1}{\rho^3} \left\{ \lambda^2 h D^2 \tau_0^2 \left[(1 + Rr_0)E[X^2] - (1 + R^2 r_0^2)E[X]^2 \right] + 2a_0 i \left[\lambda \tau_0 (1 - T) - \rho \right]^2 \right.$$
$$\left. + \lambda^2 \tau_0^2 \left[hQ + 2b_0 i(1 - R)^2 \right] \right\} \tag{5.3.2}$$

which is seen upon inspection to be strictly positive. The second principal minor, Δ_2, becomes:

$$\det H(E[TC]) = \left[\left\{ 16i^2 a_0 b_0 Q^2 \left(1 - \lambda\left[\tau_0 + (1 + r_0)QE[X]\right]\right)^2 + D^6 h^2 r_0^2 \tau_0^2 \left(\mathrm{Var}[X] - E[X]^2\right)^2 \right. \right.$$
$$\left. + 8D^3 hi \mathrm{Var}[X] \left(D\tau_0 b_0 + r_0^2 Q^2 E[X](1 - \lambda\tau_0) \right) a_0 \right\}$$
$$+ \left\{ 8D^2 hir_0^2 E[X]^2 a_0 Q^2 \left(Q + \left[1 - \lambda(\tau_0 + QE[X])\right] DE[X] - \left[1 - \lambda(\tau_0 + QE[X])\right]^2 \right) \right\}$$
$$+ \left\{ 8D^4 hir_0 \tau_0^2 b_0 \left(E[X^2] - r_0 E[X]^2 \right) \right\}$$
$$\left. + \left\{ 2D^2 h \tau_0^2 \left(4iQb_0 - 2D^2 hiQr_0^2 E[X]^2 - D^4 hr_0^2 E[X^2]^2 \right) \right\} \right] \tag{5.3.3}$$

This equation can also be broken into components for analysis:

$$\det H(E[TC]) = Z_1 + Z_2 + Z_3 + Z_4 \tag{5.3.4}$$

where:

$$Z_1 = \Big\{16i^2 a_0 b_0 Q^2 \big(1 - \lambda[\tau_0 + (1 + r_0)QE[X]]\big)^2 + D^6 h^2 r_0^2 \tau_0^2 \big(Var[X] - E[X]^2\big)^2$$
$$+ 8D^3 hi Var[X]\big(D\tau_0 b_0 + r_0^2 Q^2 E[X](1 - \lambda\tau_0)\big)a_0\Big\}$$
$$Z_2 = \Big\{8D^2 hir_0^2 E[X]^2 a_0 Q^2 \Big(Q + [1 - \lambda(\tau_0 + QE[X])]DE[X] - [1 - \lambda(\tau_0 + QE[X])]^2\Big)\Big\}$$
$$Z_3 = \Big\{8D^4 hir_0 \tau_0^2 b_0 \big(E[X^2] - r_0 E[X]^2\big)\Big\}$$
$$Z_4 = \Big\{2D^2 h\tau_0^2 \big(4iQb_0 - 2D^2 hiQr_0^2 E[X]^2 - D^4 hr_0^2 E[X^2]^2\big)\Big\}$$

The sign of the overall expression is not apparent upon inspection. To determine the sign, each of the components of 5.3.4 will be examined in turn. First, however, to assist in evaluation of the terms in this function, a series of inequalities is derived from the defined bounds on system utilization (i.e., $0 \le \rho < 1$). Each of these inequalities follows from those on the previous lines. Starting with utilization, we have:

$$0 \le \rho < 1 \tag{5.3.5a}$$
$$\Rightarrow 0 < 1 - \rho \le 1$$
$$\Rightarrow 0 < 1 - \lambda E[S] \le 1$$
$$\Rightarrow 0 < 1 - \lambda[T\tau_0 + (1 + Rr_0)QE[X]] \le 1$$
$$\Rightarrow 0 < 1 - \lambda[\tau_0 + (1 + r_0)QE[X]] \le 1 \tag{5.3.5b}$$
$$\Rightarrow 0 < 1 - \lambda[\tau_0 + QE[X]] \le 1 \tag{5.3.5c}$$
$$\Rightarrow 0 < 1 - \lambda\tau_0 \le 1 \tag{5.3.5d}$$

Equation 5.3.5a can also be expressed as:

$$0 \le \lambda[T\tau_0 + (1 + Rr_0)QE[X]] < 1$$
$$\Rightarrow 0 \le D/Q\,[QE[X]] < 1$$
$$\Rightarrow 0 \le DE[X] < 1 \tag{5.3.6}$$

Now the components of 5.3.4 can be examined to determine their signs. The first component, Z_1, was:

$$Z_1 = \left\{16i^2 a_0 b_0 Q^2\left(1-\lambda\left[\tau_0 + (1+r_0)QE[X]\right]\right)^2 + D^6 h^2 r_0^2 \tau_0^2\left(\mathrm{Var}[X] - E[X]^2\right)^2 \right.$$
$$\left. + 8D^3 hi\mathrm{Var}[X]\left(D\tau_0 b_0 + r_0^2 Q^2 E[X](1-\lambda\tau_0)\right)a_0\right\}$$

The first term of Z_1 is always positive by 5.3.5b, while the second is always non-negative. Similarly, the third term is non-negative by 5.3.5c. Overall, Z_1 is always strictly positive.

The second component, Z_2, was:

$$Z_2 = \left\{8D^2 hir_0^2 E[X]^2 a_0 Q^2\left(Q + \left[1-\lambda(\tau_0 + QE[X])\right]DE[X] - \left[1-\lambda(\tau_0 + QE[X])\right]^2\right)\right\}$$

The sign of this component depends upon the relative magnitudes of the terms in the parentheses:

term1: $+Q$

term2: $+\left[1-\lambda(\tau_0 + QE[X])\right]DE[X]$

term3: $-\left[1-\lambda(\tau_0 + QE[X])\right]^2$

From 5.3.5b and 5.3.6, the value of term2 must be in the range [0, 1). By Equation 5.3.5b, term3 must be in the range [-1, 0). In the case of each term assuming its lowest value, term1 = 1 (since $Q \geq 1$), term2 = 0 and term3 = -1, so that $Z_2 = 0$. For any other possible value for each of these three terms, $Z_2 > 0$. Thus, the second component of 5.3.4 is non-negative.

For the terms in component Z_3,

$$Z_3 = \left\{8D^4 hir_0 \tau_0^2 b_0 \left(E[X^2] - r_0 E[X]^2\right)\right\}$$

The non-negativity of variance is used, i.e.,

$$\mathrm{Var}[X] \geq 0$$
$$E[X^2] - E[X]^2 \geq 0$$
$$E[X^2] - r_0 E[X]^2 \geq 0 \qquad \text{(since } 0 \leq r_0 < 1\text{)} \tag{5.3.7}$$

(r_0, the initial defect rate, is constrained to be less than 100%, as well as being non-negative.) Using 5.3.7, it can be seen that Z_3 is non-negative as well.

For the last component, Z_4,

$$Z_4 = \left\{2D^2 h\tau_0^2 \left(4iQb_0 - 2D^2 hiQr_0^2 E[X]^2 - D^4 hr_0^2 E[X^2]^2\right)\right\}$$

consider the definition for the square of the coefficient of variation, cv,

$$cv^2 = \frac{E[X^2]}{E[X]^2} - 1$$
$$\Rightarrow \quad E[X^2] = (cv^2 + 1)E[X]^2 \tag{5.3.8}$$

Then, Z_4 can be written as:

$$2D^2 h\tau_0^2 \left(4iQb_0 - 2D^2 hiQr_0^2 E[X]^2 - (cv^2 + 1)^2 D^4 E[X]^4 hr_0^2\right)$$
$$> \quad 2D^2 h\tau_0^2 \left(4ib_0 - 2hi - (cv^2 + 1)^2 h\right) \tag{5.3.9}$$

since: $0 \le DE[X] < 1$

$0 \le r_0 < 1$

$Q \ge 1$

An intuitive comparison can be made of the magnitudes of the terms in 5.3.9. The interest rate for investments, i, is likely to be of order $O(10^{-1})$ as will the coefficient of variation squared, cv^2. Unit holding cost, h, might be of order $O(10^0)$. The final parameter, b_0, approximates the investment necessary to completely eliminate defects from the system, so would likely be of order $O(10^3)$ or greater. Since this parameter so completely dominates the magnitude of the others, it is reasonable to conclude that 5.3.9, and hence Z_4 will in practice be positive.

Therefore, since all the groups of terms in 5.3.4 can be assumed to be non-negative or positive, it can be concluded that the second principal minor is always positive and E[TC] is always convex in terms of setup time and defect rate.

5.3.2 Bivariate Minimization of E[TC]

Since E[TC] is convex in T and R, any critical points found will represent minimizers of total cost. To locate critical points, the necessary condition applied and optima are found from the simultaneous solution of the system:

$$\frac{\partial E[TC]}{\partial T} = 0$$
$$\frac{\partial E[TC]}{\partial R} = 0$$

Which, after simplification, yields:

$$\left\{\frac{Q}{2}\left[hD^3 r_0 E[X](Q-1) - 2(2ib_0 - hDr_0QE[X])(1-DE[X]) + 2iDr_0E[X](a_0+b_0)\right.\right.$$
$$\left.\left.+hD^2 r_0 E[X^2]\right]\right\}$$
$$+\left\{\frac{D}{2}\left[2hD^2\tau_0 r_0 QE[X]^2 - hD^2\tau_0 r_0 E[X^2] - 2hD\tau_0 r_0 QE[X] + 4i\tau_0 b_0 - 4ir_0QE[X]a_0\right]\right\}T$$
$$-\left\{(1-DE[X])Q\left[hD^2 r_0^2 E[X](Q+1) - 2ib_0\right]\right\}R + \left\{D\tau_0\left[hD^2 r_0^2 E[X](Q+1) - 2ib_0\right]\right\}TR$$
$$\left\{\frac{D}{2} r_0 E[X]\left(hD^2\tau_0^2 + 2iQa_0\right)\right\}T^2 + \left\{\frac{D}{2} r_0 QE[X]\left[hD^2 r_0^2 E[X](Q+1) - 2ib_0\right]\right\}R^2 = 0$$

(5.3.10a)

$$\left\{\frac{hD^3\tau_0}{2}E[X]^2(Q-1) + \frac{hD^3\tau_0}{2}E[X^2] - QE[X]\left(hD^2\tau_0 - 2ia_0\right) + D\tau_0\left(hQ + ia_0 + ib_0\right)\right.$$
$$\left.-2Qia_0\right\}$$
$$+\left\{\left(hD^2\tau_0^2 - 2iQa_0\right)(1-DE[X])\right\}T + \left\{Dr_0E[X]\left(hD^2\tau_0^2 - 2iQa_0\right)\right\}TR$$
$$+\left\{\frac{D}{2}\left[2hD^2\tau_0 r_0 QE[X]^2 + hD^2\tau_0 r_0 E[X^2] - 2hD\tau_0 r_0 QE[X] + 4ir_0QE[X]a_0 - 4i\tau_0 b_0\right]\right\}R$$
$$+\left\{Dr_0E[X]\left(hD^2\tau_0^2 - 2iQa_0\right)\right\}T^2 + \left\{\frac{D\tau_0}{2}\left[hD^2 r_0^2 E[X]^2(Q-1) + 2ib_0\right]\right\}R^2 = 0$$

(5.3.10b)

As Equation 5.3.4 showed, any solution to 5.3.10 will represent a minimizer of E[TC], and an optimal investment strategy. Unfortunately, due to the complexity of this non-linear system, no analytic solution to 5.3.10 has been found.

This result can be used, however, to derive some of the behaviour of the decision matrix. First of all, Equation 5.3.4 showed that the E[TC] function is convex in setup time and defect rate. This would indicate that the decision matrix must have a significant region where the 'some-some' investment decision is optimized, as a result of convexity.

Furthermore, the results from Equations 5.2.3 and 5.2.3 indicate that when there is a high cost of improving quality (i.e., when no investment takes place in quality improvement), three optimal setup investment decisions occur (e.g., invest none, invest some and invest fully), depending on the cost of eliminating setup time. Similarly, for high costs of improving setup, Equations 5.2.7 and 5.2.8 show that three optimal quality investment decisions occur (e.g., invest none, invest some and invest fully).

It may easily be shown that if Equations 5.2.1 through 5.2.4 were re-derived assuming $R = 0$ (i.e., full investment had taken place in quality improvement), setup improvements would still have three optimal decisions: invest none, some and full. As well, assuming $T = 0$ for Equations 5.2.5 through 5.2.8 (i.e., full invest in setup), it could be shown that the three optimal decisions will exist for quality improvement.

Together, these arguments show that the 'informed' decision matrix will contain regions for the nine separate optimal decisions (i.e., 'full-full', 'full-some', ... , 'none-none') represented in the decision matrix for the 'naive' case, although the sizes of the regions will change.

The boundaries for the 'informed' decision matrix can be developed as follows.

1) Setup Boundaries

a) No investment in Quality (i.e., $R^* = 1$)

This is the 'naive' case where the boundaries were given as Equations 5.2.3 and 5.2.4, i.e., No invest boundary

$$a_0^*\Big|_{T^*=1,R^*=1} = \frac{\left(D^2\tau_0^2 h + 2Qia_1\right)}{2Qi} + \frac{1}{4Qi\left[D\tau_0 + Q\{D(1+r_0)E[X]-1\}\right]}\Big\{$$

$$-2Q^2(D(1+r_0)E[X]-1)(D\tau_0 h + ia_1) - D\tau_0\left(D^2\tau_0^2 h + 2Qia_1\right) \tag{5.3.11a}$$

$$+D^3\tau_0 hQ\left[(1+r_0)E[X^2] + \left((Q-1)(1+r_0)^2 + 2(1+r_0) - 2\right)E[X]^2\right]\Big\}$$

Full invest boundary

$$a_0^*\Big|_{T^*=0,R^*=1} = \frac{\left(D^2\tau_0^2 h + 2Qia_1\right)}{2Qi} + \frac{1}{2QiD\tau_0}\Big\{$$

$$-2Q^2(D(1+r_0)E[X]-1)(D\tau_0 h + ia_1) - D\tau_0\left(D^2\tau_0^2 h + 2Qia_1\right) \tag{5.3.11b}$$

$$+D^3\tau_0 hQ\left[(1+r_0)E[X^2] + \left((Q-1)(1+r_0)^2 + 2(1+r_0) - 2\right)E[X]^2\right]\Big\}$$

b) Full investment in Quality (i.e., $R^* = 0$)

These boundaries can be developed by re-deriving Equations 5.2.1 through 5.2.4 under the assumption that $R^* = 0$. Optimal setup time becomes:

$$T^* = \frac{1-DE[X]}{\lambda\tau_0} + \frac{1}{D\tau_0}\sqrt{\frac{Q}{hD^2\tau_0^2 + 2Qi(a_0+a_1)}} \times$$

$$\Big[\; 2Q^2 i(a_0+a_1)[1-DE[X]]^2 - D\tau_0(hD\tau_0 Q + 2D\tau_0 ia_0 + 2Qia_1) \tag{5.3.12}$$

$$+hD^4\tau_0^2\left[E[X]^2 - E[X^2]\right] + 2D^2QE[X]\tau_0 ia_1 \;\Big]^{\frac{1}{2}}$$

No invest boundary

$$a_0^*\Big|_{T^*=1,R^*=0} = \frac{\left(D^2\tau_0^2h + 2Qia_1\right)}{2Qi} + \frac{1}{4Qi\left[Q - D\tau_0 - DQE[X]\right]}\Big\{$$
$$+2Q^2(DE[X] - 1)\left(D\tau_0 h + ia_1\right) + D\tau_0\left(D^2\tau_0^2h + 2Qia_1\right) \qquad (5.3.13a)$$
$$-D^3\tau_0 hQ\left[E[X^2] + (Q-1)E[X]^2\right] + 4QiD\tau_0 b_0 \Big\}$$

Full invest boundary

$$a_0^*\Big|_{T^*=0,R^*=0} = \frac{\left(D^2\tau_0^2h + 2Qia_1\right)}{2Qi} + \frac{1}{2QiD\tau_0}\Big\{ - 2Q^2(DE[X] - 1)\left(D\tau_0 h + ia_1\right)$$
$$-D\tau_0\left(D^2\tau_0^2h + 2Qia_1\right) + D^3\tau_0 hQ\left[E[X^2] + (Q-1)E[X]^2\right] + 2QiD\tau_0 b_0 \Big\} \qquad (5.3.13b)$$

It may be observed that these two boundaries are linear functions of the cost to eliminate defects, b_0. These sections of the boundaries will plot as diagonal lines on the decision matrix.

c) Some investment in Quality (i.e., $0 < R^* < 1$)

This case is more complex since it involves the simultaneous solution to Equations 5.3.10. An analytic solution has not been found, but an algorithm can be used to plot the boundary for given numerical values for the model's parameters.

Algorithm 5.1: No-invest boundary (i.e., $T^* = 1$)

1) choose R^*, $0 \leq R^* \leq 1$
2) substitute R^* and $T^* = 1$ into Equations 5.3.10a and 5.3.10b. (This yields two equations in two unknowns, a_0 and b_0.)
3) solve simultaneously for $(a_0, b_0)^*$
4) plot point $(a_0, b_0)^*$ on decision matrix
5) repeat from Step 1 for next value of R^*

(In practice, numerical experiments have shown this portion of the boundary to be curved just as the mid-portion of the setup boundary in the 'informed' case of Chapter 4 was curved, e.g., such as is shown in Figure 5.5.)

These results can be used to develop the no-invest and full-invest boundaries for setup time reduction over the range of cost to eliminate defects of zero to infinite. The general shape of these boundaries is shown in Figure 5.5.

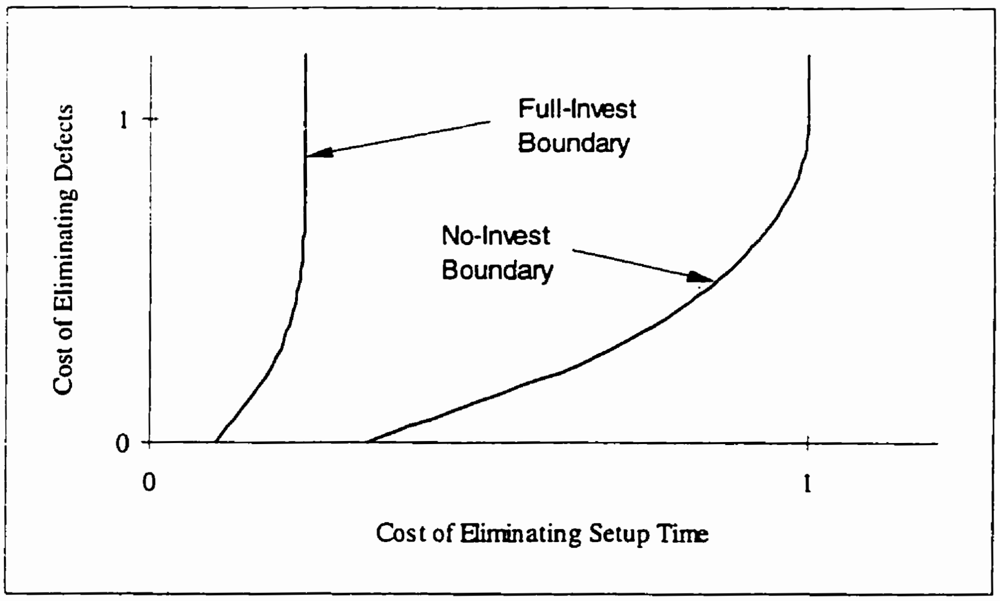


Figure 5.5: 'Informed' Case Setup Decision Boundaries

2) Quality Boundaries

The decision boundaries for quality investment are developed in a manner similar to that used for the setup boundaries. Three cases are delineated, namely, 'no', 'full' and 'some' investment in quality improvement, and are given below:

a) No investment in Setup (i.e., T* = 1)

This is the 'naive' case where the boundaries were given as Equations 5.2.7 and 5.2.8, i.e.,

No invest boundary

$$b_0^* \Big|_{R^*=1, T^*=1} = \frac{1}{i}\left[2ib_1 + D^2 r_0^2 E[X]^2 h(Q+1)\right] + \frac{1}{2i\left\{DQr_0E[X] + \left[D\tau_0 + DQE[X] - Q\right]\right\}}\Big\{ 2\left[D\tau_0 + DQE[X] - Q\right]\left[D^2 h r_0 \tau_0 E[X] - ib_1\right] - 2DQr_0E[X]ib_1 + hDr_0E[X]\left[2Q^2 - D^2\tau_0^2 - 2DQ^2E[X] + D^2E[X]^2\left\{(Q-1) + (Q+1)r_0\right\}\right]\Big\} \tag{5.3.14a}$$

Full invest boundary

$$b_0^* \Big|_{R^*=0, T^*=1} = \frac{1}{i}\left[2ib_1 + D^2 r_0^2 E[X]^2 h(Q+1)\right] + \frac{1}{iDQr_0E[X]}\Big\{ 2\left[D\tau_0 + DQE[X] - Q\right]\left[D^2 h r_0 \tau_0 E[X] - ib_1\right] - 2DQr_0E[X]ib_1 + hDr_0E[X]\left[2Q^2 - D^2\tau_0^2 - 2DQ^2E[X] + D^2E[X]^2\left\{(Q-1) + (Q+1)r_0\right\}\right]\Big\} \tag{5.3.14b}$$

b) Full investment in Setup (i.e., T* = 0)

These boundaries can be developed by re-deriving Equations 5.2.5 through 5.2.8 under the assumption that T* = 0. Optimal setup time becomes

$$R^* = \frac{1}{Dr_0E[X]} + \frac{Q}{Dr_0E[X]}\frac{1}{\sqrt{hD^2(Q+1)r_0^2E[X]^2 + 2i(b_0+b_1)}} \times$$
$$\Bigg[2i(b_0+b_1) - D^3hr_0^2E[X]\left[E[X^2] + 2E[X]^2\right]$$
$$-2DiE[X]\left\{(b_0+b_1)[2-DE[X]] + b1[1-DE[X]]\right\}$$
$$+D^2hr_0^2E[X]^2\left[1+2D^2E[X]^2 - Q\right] - 2D^2E[X]^2r_0^2ib_0 \Bigg]^{\frac{1}{2}} \tag{5.3.15}$$

<u>No invest boundary</u>

$$b_0{}^*\Big|_{R^*=1,T^*=0} = \frac{1}{i}\left[2ib_1 + D^2r_0^2E[X]^2h(Q+1)\right] +$$
$$\frac{1}{2iQ\{1-D(1+r_0)E[X]\}}\Big\{ 2QiDr_0E[X]a_0 + 2Qib_1[1-D(1+r_0)E[X]] \tag{5.3.16a}$$
$$-hDr_0E[X]\left[2Q^2 - 2DQ^2E[X] + D^2E[X]^2\{(Q-1)+(Q+1)r_0\}\right]\Big\}$$

<u>Full invest boundary</u>

$$b_0{}^*\Big|_{R^*=0,T^*=0} = \frac{1}{i}\left[2ib_1 + D^2r_0^2E[X]^2h(Q+1)\right] + \frac{1}{iDQr_0E[X]}\Big\{$$
$$2QiDr_0E[X]a_0 + 2Qib_1[1-DE[X]] - 2DQr_0E[X]ib_1 \tag{5.3.16b}$$
$$+hDr_0E[X]\left[2Q^2 - 2DQ^2E[X] + D^2E[X]^2\{(Q-1)+(Q+1)r_0\}\right]\Big\}$$

It may be observed that these two boundaries are linear functions of the cost to eliminate setup time, a_0. These sections of the boundaries will plot as diagonal lines on the decision matrix.

c) Some investment in Setup (i.e., $0 < T^* < 1$)

This case is more complex since it involves the simultaneous solution to Equations 5.3.10. An analytic solution has not been found, but an algorithm can be used to plot the boundary for given numerical values for the model's parameters.

<u>Algorithm 5.2</u>: No-invest boundary (i.e., $T^* = 1$)

1) choose T^*, $0 \leq T^* \leq 1$
2) substitute T^* and $R^* = 1$ into Equations 5.3.10a and 5.3.10b. (This yields two equations in two unknowns, a_0 and b_0.)
3) solve simultaneously for $(a_0, b_0)^*$
4) plot point $(a_0, b_0)^*$ on decision matrix
5) repeat from Step 1 for next value of T^*

Using these results a decision matrix for quality can be drawn and is shown in Figure 5.6. Combining Figures 5.5 and 5.6 leads to the complete 'informed' case decision matrix, given as Figure 5.7. The most striking difference between this decision matrix and Figure 3.11 from Chapter 3's 'informed' case decision matrix is that a large 'some-some' optimal decision region exists here but did not exist in the model using the linear investment function. This discrepancy suggests that the form of the investment function plays a large role in the optimal investment decisions in a system.

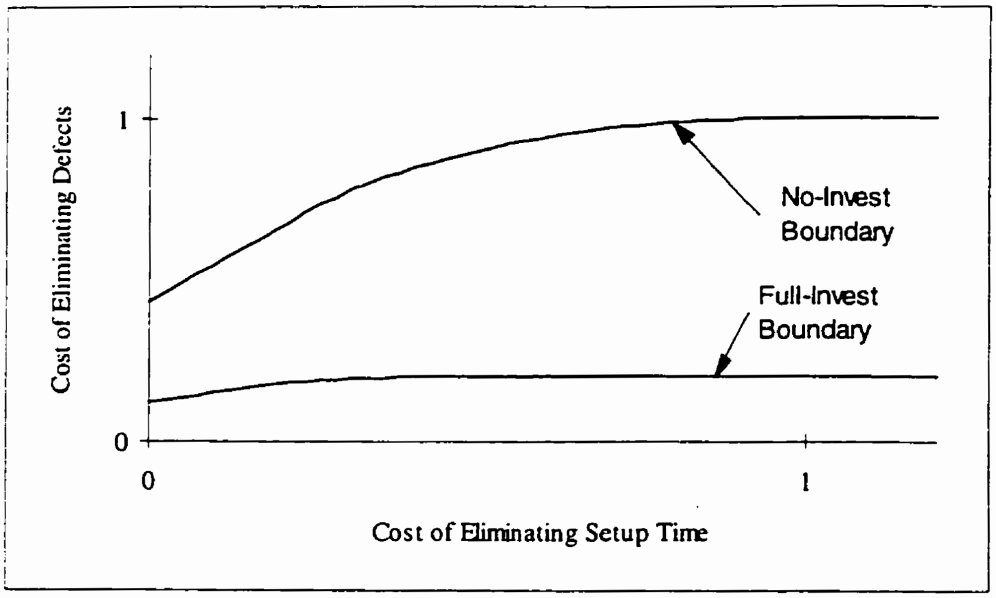


Figure 5.6: 'Informed' Case Quality Decision Boundaries

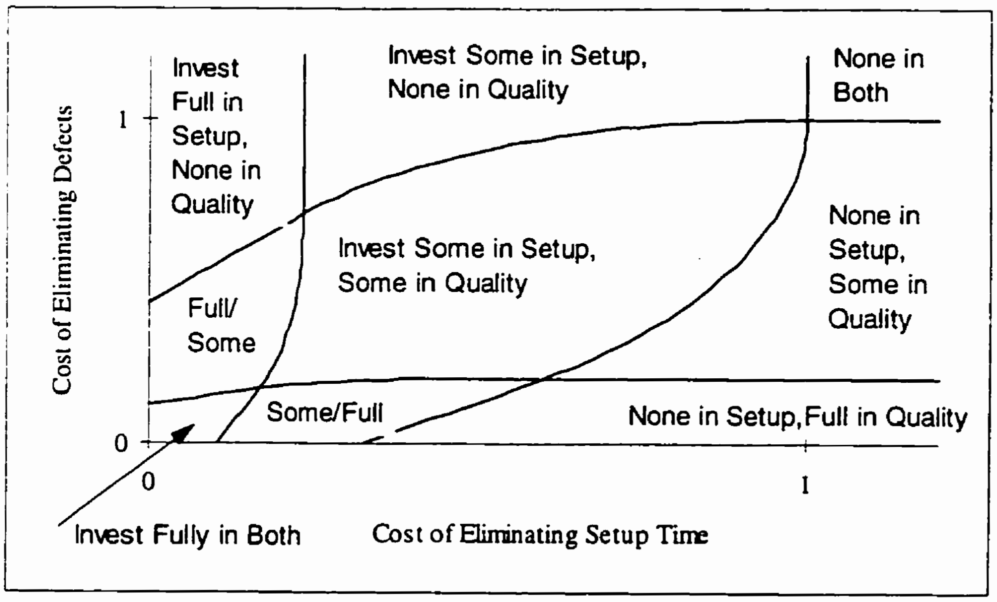


Figure 5.7: Completed Decision Matrix for 'Informed' Case

5.4: Decision Error and Sensitivity Analysis

5.4.1 Decision Error

Analytic determination of an expression for decision error in this version of the model has not been possible due to the complexity of the decision matrix in the 'informed' case. As in the last chapter, comparisons will be made based upon numerical examples. Example problem parameters are all taken from Table 3.2.

Figure 5.8 shows how the setup boundaries change from the 'naive' to 'informed' cases. Due to the convexity of the investment function, this decision matrix contains a relatively large region for 'some' investment. In the "some setup over-investment" region, the optimal investment changes from investing 'some' to investing 'none' as interactions are considered. The "full setup over-investment" region has the optimal investment change from investing 'fully' to investing 'some'. In each case the effect of ignoring interactions is that over-investment is possible. No regions on this matrix would lead to under-investment.

Quality decisions behave similarly, and are shown in Figure 5.9. Again there are two regions of substantial size on the decision matrix where ignoring interactions will lead to over-investment. Combining Figures 5.8 and 5.9 would show that there are significant fractions of the overall decision matrix where over-investment errors can take place, but not over all the matrix. This is on contrast to the more extreme behaviours of the models from Chapters 3 and 4 where the majority of the decision matrix area would be prone to decision error. This result also suggests that the form of the investment function is very important in decision making.

Frequencies of over-investment were calculated using the sample parameter values in Table 3.2, and are presented in Table 5.1. As in previous chapters, the magnitudes of potential over-investments are significant.

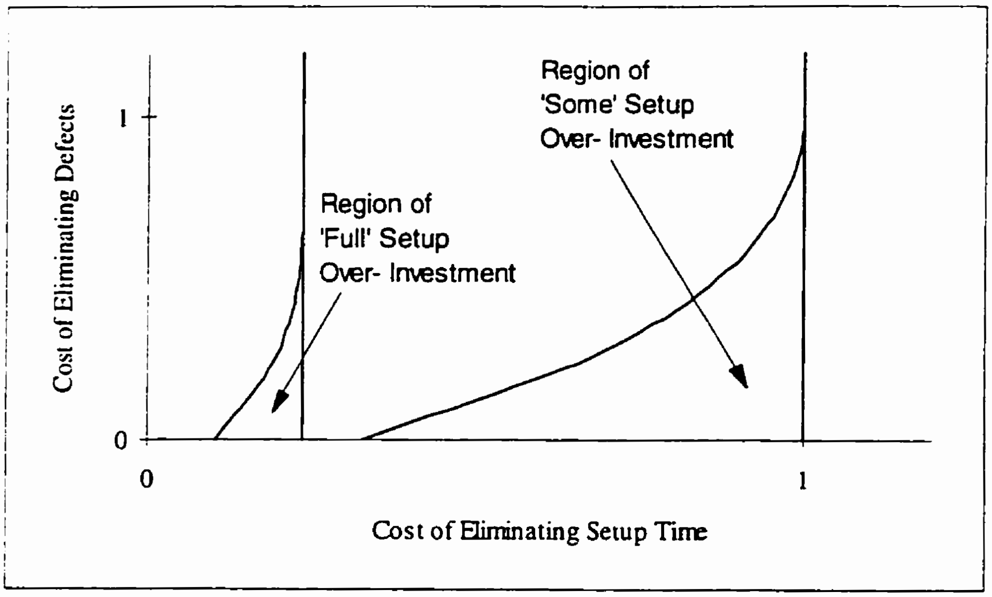


Figure 5.8: Differences in Setup Investment Decisions, 'Naive' to 'Informed' Cases

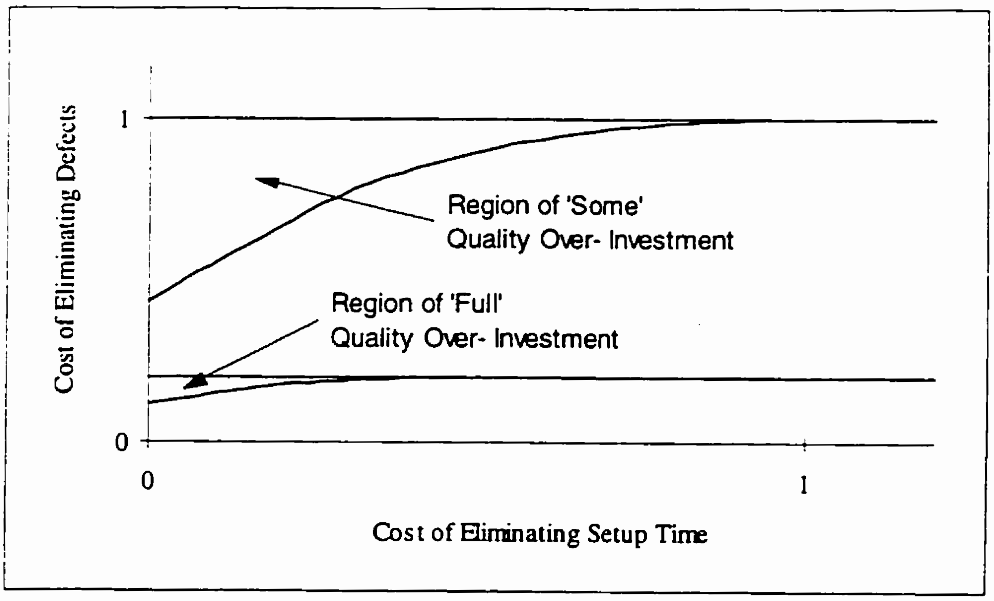


Figure 5.9: Differences in Quality Investment Decisions, 'Naive' to 'Informed' Cases

Table 5.1: Frequency of Over-investment Levels

Over-Investment Level (percentage of optimal investment)	Frequency of Observances (n = 2601)
< 5%	18%
5 - 50%	34%
50 - 100%	30%
100 - 150%	7%
150 - 200%	3%
200 - 250%	2%
250 - 300%	1%
300 - 350%	1%
350 - 400%	1%
400 - 450%	0%

5.4.2 Sensitivity Analysis

Distributions of over-investment have been calculated for the high and low parameter values given in Table 3.2 to evaluate the sensitivity of decision error to each of the input parameters in the model.

The sensitivities for the four parameters which affect system utilization, Demand, Batch Size, Initial Defect Rate and Initial Setup Time, are graphed as Figures 5.10, 5.11, 5.12 and 5.13, respectively. Consistent with the results from Chapter 3, the level of each of these parameters which increases system utilization (i.e., high demand, low batch size, high defect rate, high setup time) also shifts the distributions of over-investments upward. This is the same result as was found in Chapter 3 and corresponds to the analytic results derived there, that is, as utilization increases, so does Decision Error.

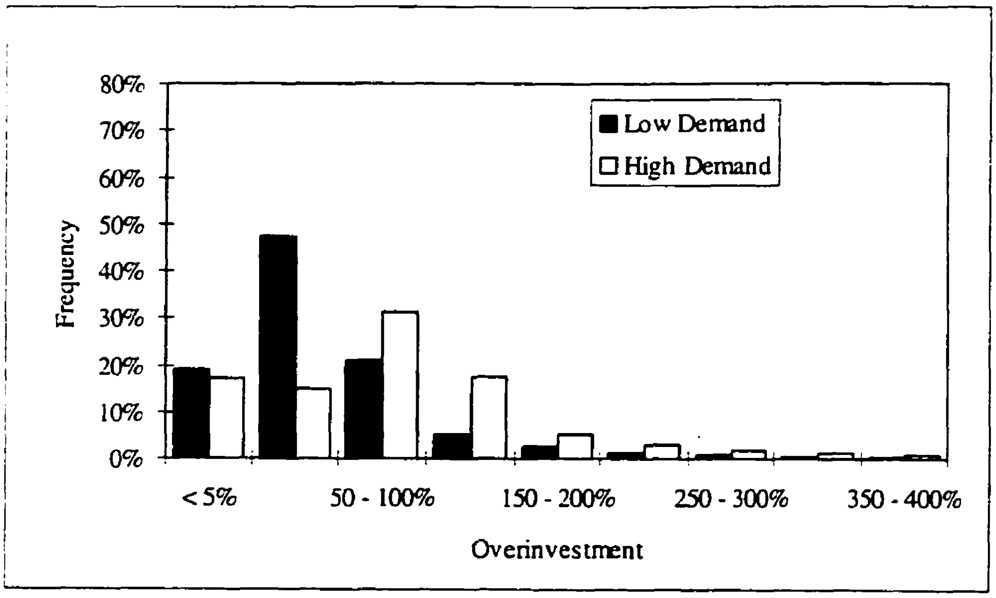


Figure 5.10: Distribution of Over-Investment as a Function of Demand

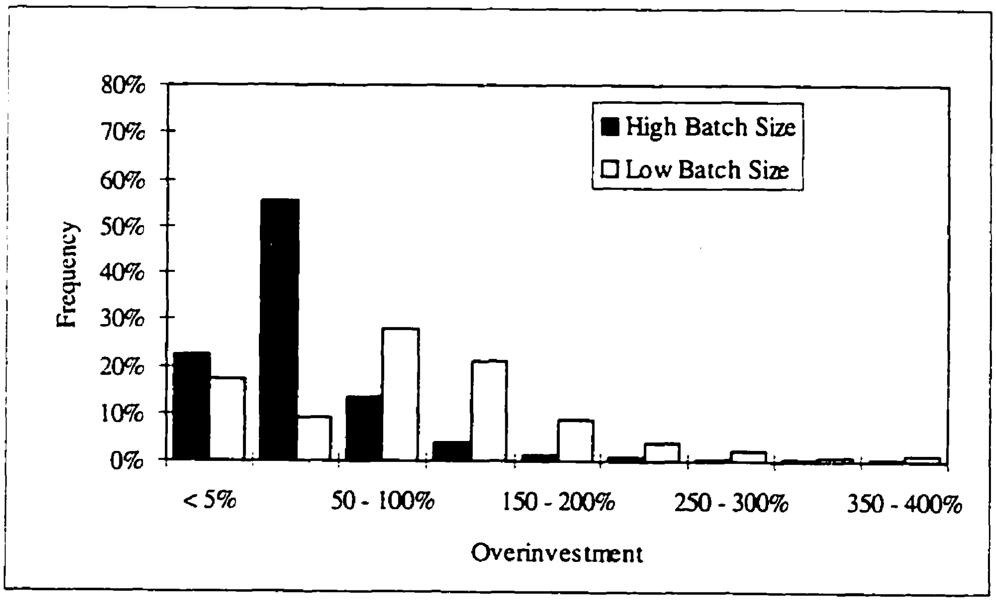


Figure 5.11: Distribution of Over-Investment as a Function of Batch Size

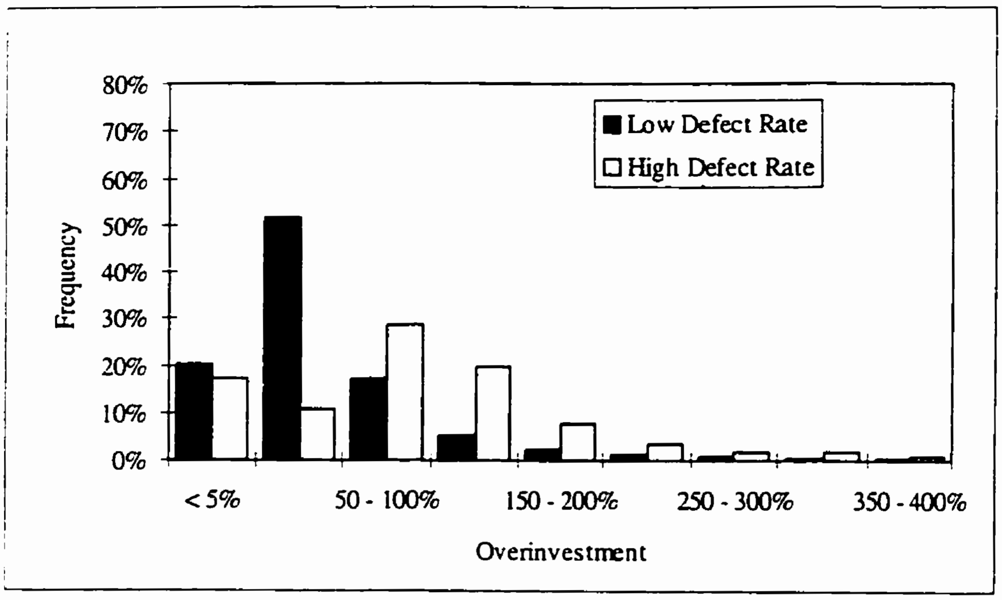


Figure 5.12: Distribution of Over-Investment as a Function of Initial Defect Rate

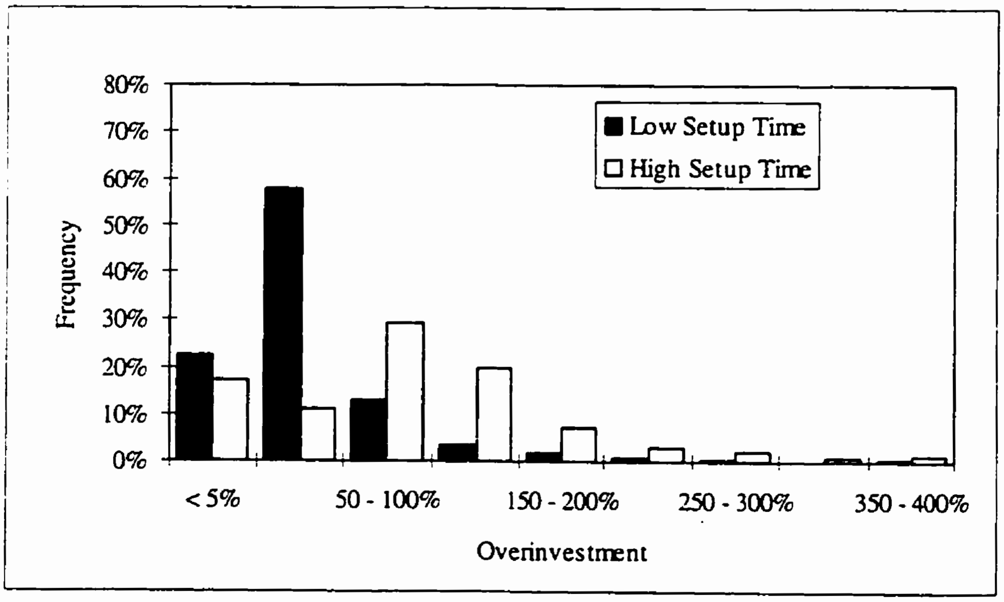


Figure 5.13: Distribution of Over-Investment as a Function of Initial Setup Time

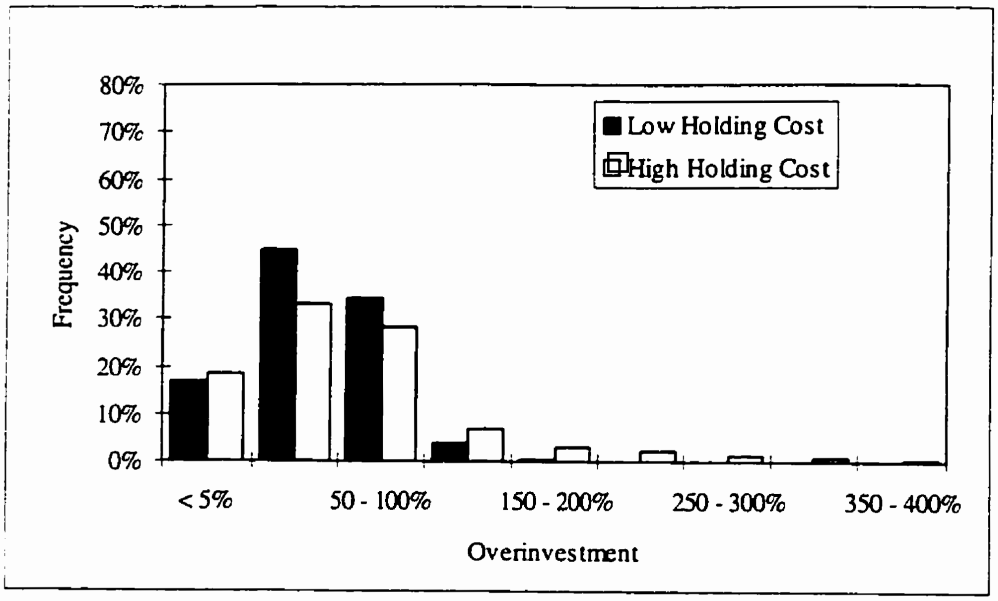


Figure 5.14: Distribution of Over-Investment as a Function of Holding Cost Rate

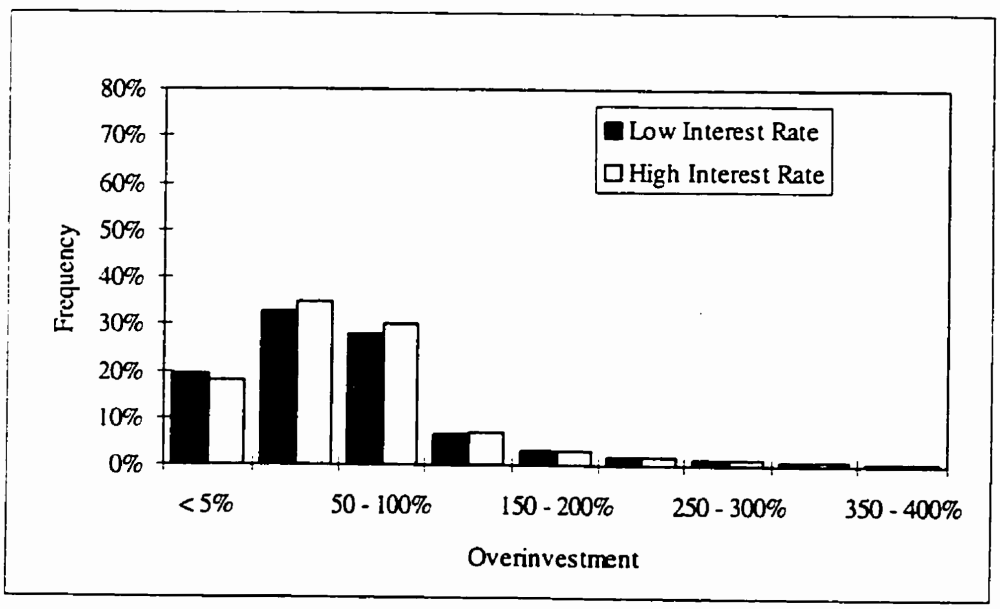


Figure 5.15: Distribution of Over-Investment as a Function of Interest Rate

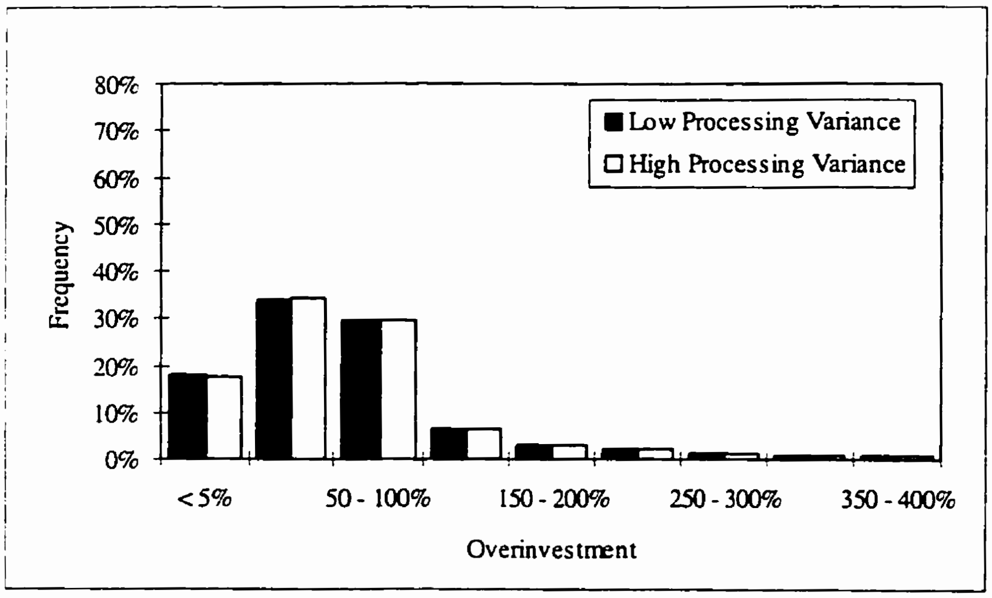


Figure 5.16: Distribution of Over-Investment as a Function of Processing Time Variance

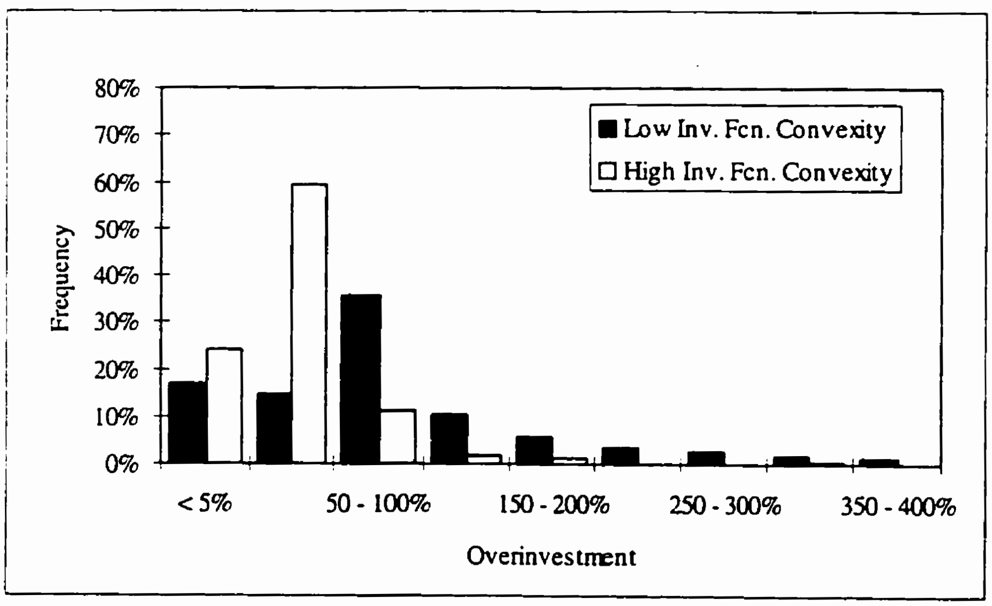


Figure 5.17: Distribution of Over-Investment as a Function of Investment Function Convexity

Figures 5.14 and 5.15 show the sensitivities for holding cost and interest rates, respectively. These sensitivities are similar to those in the models of Chapters 3 and 4, and for the same reasons. Similarly, there is essentially no sensitivity to processing time coefficient of variation (Figure 5.16) as the batch service time in this model is insensitive to the unit processing time variance for relatively large batch sizes, as in Chapter 3.

Sensitivity to investment function convexity (Figure 5.17), however, is quite significant. The explanation can be found in the decision matrices for low and high levels of investment function convexity, shown in Figures 5.18 and 5.19, respectively. As the investment function becomes more convex, so does the E[TC] function. The potential for an optimal investment decision being 'some-some' increases with the increasing E[TC] convexity, resulting in the area of the 'some-some' region increasing to occupy a greater fraction of the decision matrix. As this happens, the curvature of the decision boundaries decreases and regions where decision error occurs shrink. The net result is a decrease in the distribution of over-investment frequencies, as shown in Figure 5.17.

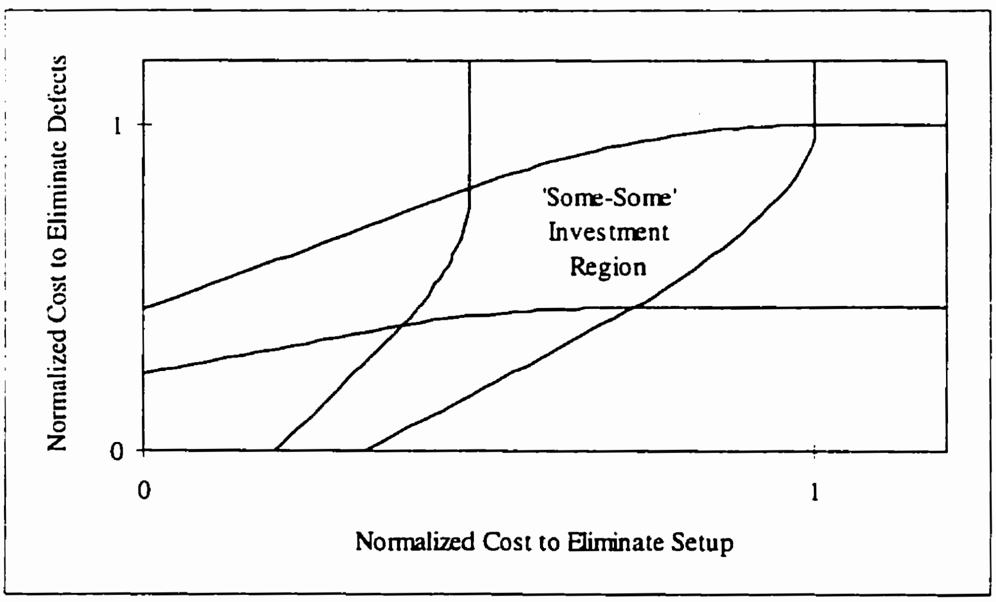


Figure 5.18: 'Informed' Decision Matrix for Low Investment Function Convexity

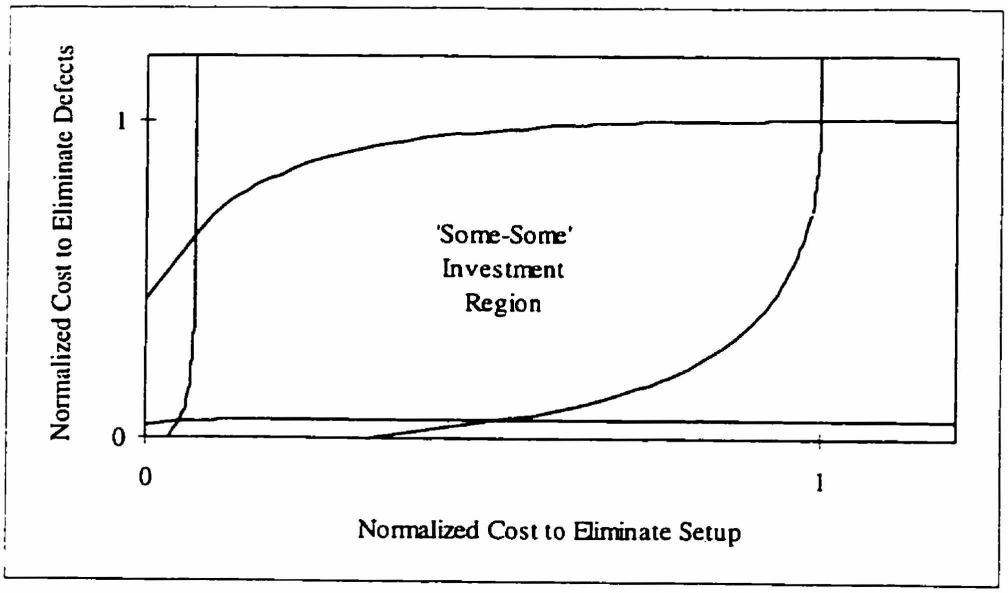


Figure 5.19: 'Informed' Decision Matrix for High Investment Function Convexity

5.5: Discussion

In this chapter, the basic model of Chapter 3 was extended to include convex, rather than linear, investment functions. The first section of this chapter discussed the types of convex investment functions found in the literature, why there is no compelling reason to choose any particular form of a function and ultimately developed a convex function mathematically suited to the queuing-based model used in this research.

The system with a convex investment function showed many of the same behaviours that were exhibited in the model of Chapter 3. Optimal investment levels for setup and quality could be 'none', 'some' or 'full' just as in Chapter 3. The most striking differences between the results of the two chapters is that the 'informed' case boundaries curved smoothly, rather than being piece-wise linear, and, the 'informed' decision matrix showed a region of 'some-some' investment as optimal.

The curved vs. piece-wise linear boundaries no doubt are due to a greater degree of convexity in the E[TC] function resulting from the convex investment function. The E[TC] function of Chapter 3 was strictly non-convex (in two variables), forcing extreme-point optima. That non-convexity explains the decision boundaries being 'kinked' as decisions change. The E[TC] function in this chapter, on the other hand, has a range in which it is convex, so that as problem parameters change, the decision boundary can follow a smoothly curving path.

The second behaviour seen, the existence of the 'some-some' optimal decision region confirms the speculation in Chapter 3 that the non-intuitive lack of such an optimal decision in that chapter was due to the linear investment function.

This result, along with the sensitivity to the convexity of the investment function shown in Figure 5.17 suggest that the form of the investment function strongly influences the output of the model. The role of the investment function will be discussed in greater detail in Chapter 7.

Chapter 6: Variable Batch Size, Convex Investment Function

In this, the last analysis chapter, a model is considered in which the convex investment function from Chapter 5 is used and batch size is assumed to be an optimization variable. Due to the complexity of this model, general analytic solutions have not been obtained. Instead, results come from numerical examples, which show behaviour similar to that found in the models of previous chapters. As in the previous chapters, optimal decisions are compared between a 'naive' case (ignoring potential interaction) and an 'informed' case (allowing potential interactions).

6.1 'Naive' Investment Decisions

The 'naive' case assumes that optimal investment levels will be determined for setup and quality improvements independently of each other. As in Chapter 4, because of the strong empirical link between setup reduction and batch size, batch size is optimized with setup time.

The two optimization problems in this case are to 1) optimize investment in setup time reduction along with batch size, and , 2) optimize investment in quality improvement. These problems are studied in turn.

6.1.1 Setup and Batch Optimization

The objective in this sub-case is to minimize total costs simultaneously with respect to setup time (T) and batch size (Q). Before examining the necessary conditions, the convexity of the system can be determined through the sufficiency condition. To manage the complexity of the expressions, investment function convexity parameters were selected maximum convexity, i.e., $a_1 = -2\ a_0$ and $b_1 = -2\ b_0$. Then, for convexity, E[TC] must satisfy:

i) Positive first principal minor of Hessian, Δ_1

$$\Delta_1 = \frac{\partial^2 E[TC]}{\partial T^2} > 0 \tag{6.1.1a}$$

ii) Non-negative second principal minor of Hessian, Δ_2

$$\Delta_2 = \begin{vmatrix} \dfrac{\partial^2 E[TC]}{\partial T^2} & \dfrac{\partial^2 E[TC]}{\partial T \partial Q} \\ \dfrac{\partial^2 E[TC]}{\partial T \partial Q} & \dfrac{\partial^2 E[TC]}{\partial Q^2} \end{vmatrix} \geq 0 \tag{6.1.1b}$$

These two conditions become, for the model of this chapter,

i)

$$\begin{aligned} \Delta_1 = \frac{1}{\rho^3} \Big\{ & \lambda^2 h D^2 \tau_0^2 \left[(1 + Rr_0) E[X^2] - (1 + R^2 r_0^2) E[X]^2 \right] + 2a_0 i \left[\lambda \tau_0 (1 - T) - \rho \right]^2 \\ & + \lambda^2 \tau_0^2 \left[hQ + 2b_0 i (1 - R)^2 \right] \Big\} \end{aligned} \tag{6.1.2a}$$

and,

ii)

$$\begin{aligned} \Delta_2 = & -D^6 \tau_0^2 h^2 \left[(1 - Rr_0) E[X^2] - (1 + R^2 r_0^2) E[X]^2 \right]^2 \\ & -4D^3 \tau_0 h i \left[(1 - Rr_0) E[X^2] - (1 + R^2 r_0^2) E[X]^2 \right] \\ & \times \left[D\tau_0 b_0 (1 - R)^2 + D\tau_0 a_0 (1 - 3T^2) - 2TQa_0 \rho \right] \\ & +4D\tau_0 i \Big[-D\tau_0 i \left\{ b_0 (1 - R)^2 - a_0 \right\}^2 + 4TQa_0 i\rho \left\{ b_0 (1 - R)^2 + a_0 (1 - T)^2 \right\} \\ & \quad D\tau_0 a_0 \left\{ 6T^2 i b_0 (1 - R)^2 + T^2 a_0 i (6 - 8T + 3T^2) + 2T^2 Qh \right\} \Big] \end{aligned} \tag{6.1.2b}$$

While each of the terms of 6.1.2a is positive, so that Δ_1 is positive, the sign of 6.1.2b is much more difficult to determine analytically from this expression. Instead, a numerical evaluation has been performed. A Monte Carlo simulation was developed according to Algorithm 6.1, using parameter values selected uniformly over the ranges given in Table 6.1.

Algorithm 6.1: Monte Carlo Simulation for E[TC] Convexity Check

1. Sample values for parameters D, τ_0, r_0, h, i, cv, a_0, b_0 according to uniform distribution with endpoints given in Table 6.1
2. Calculate upper and lower bounds for a_1/a_0 and b_1/b_0 from Equations 5.1.15 and 5.1.16 (bounds on investment function convexity)
   2.1. If upper and lower bounds for either a_1/a_0 or b_1/b_0 overlap (i.e., investment function is infeasible), go to Step 1
   2.2. Sample a_1/a_0 and b_1/b_0 uniformly over their feasible ranges
3. Calculate utilization. If $\rho \geq 1$, go to Step 1
4. Calculate values for minors from Equations 6.1.2a and 6.1.2b
5. Repeat from Step 1 until required number of points are obtained.

Table 6.1 Monte Carlo Simulation Parameter Ranges

Parameter	Lower Limit	Upper Limit
D	20,000	200,000
τ_0	0	0.01
r_0	0	0.5
h	0.1	10
i	0.01	0.5
cv	1	2
a_0	1,000	1,000,000
b_0	1,000	1,000,000
a_1/a_0	-2	from Equation 5.1.15
b_1/b_0	-2	from Equation 5.1.16

The simulation was run until 5000 points had been obtained (this number was arbitrarily chosen). A histogram of the frequency of the results for the first principal minor is given as Figure 6.1. Most values, > 81% in this example, are positive, indicating that in a significant number of cases the E[TC] function is convex with respect to setup time alone, and it has local minima for values of T between 0 and 1 (i.e., where 'some' investment is optimal). (This result differs from Equation 6.1.2a since the Monte Carlo simulation also sampled for the investment function convexity parameters a_i and b_i.)

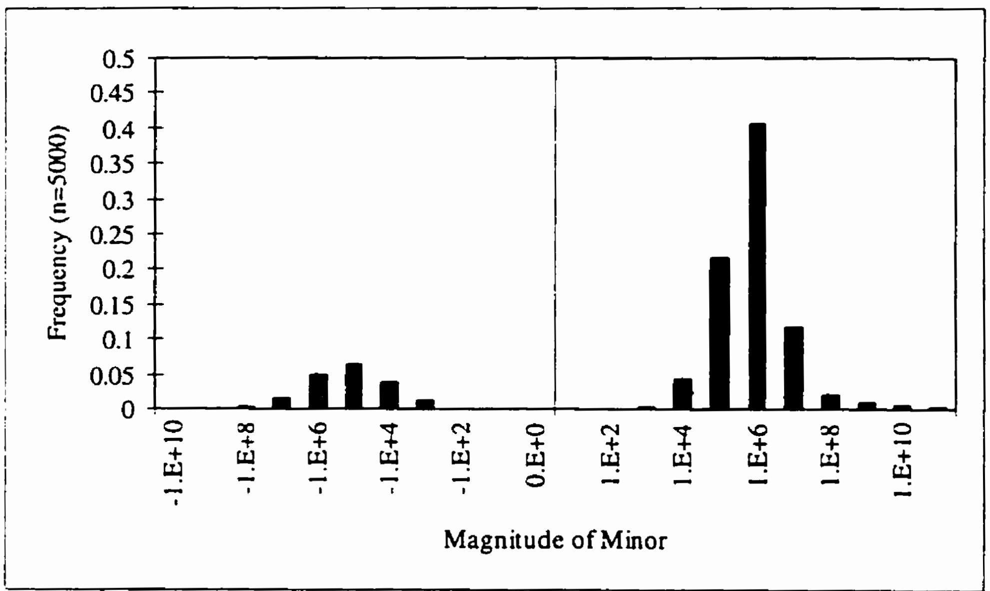


Figure 6.1: Frequencies of First Principal Minor Observations from Monte Carlo Simulation

The results for the second principal minor are given in Figure 6.2. The distribution in this case is primarily non-positive (1.5% of values being positive), indicating E[TC] as a function of setup time and batch sizes is convex only over a small region of the decision space, but can be convex. This result suggests that a 'some' investment decision can be optimal in the (T, Q) system, but will be relatively small compared to the 'full' investment region.

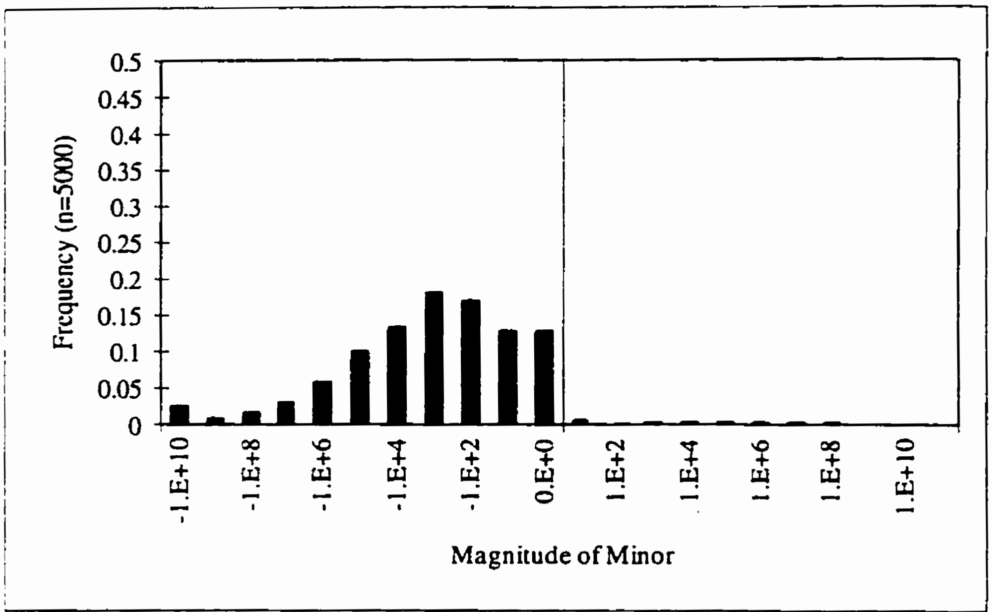


Figure 6.2: Frequencies of Second Principal Minor Observations from Monte Carlo Simulation

Indeed, a numerically generated decision matrix, Figure 6.3, using the example parameter values given in Table 3.2 confirms this prediction. There is a region of 'some' investment, but it is relatively narrow compared to the 'full' investment region, in contrast, say, to the sizes of the corresponding regions in the 'naive' case decision matrices of Figures 5.4 or 3.4. (The system in Chapter 4 was strictly non-convex and there was no 'some' investment region.) The axes in Figure 6.3, and subsequent decision matrices, are normalized such that 'naive' case no-investment boundaries fall at a normalized value of one on the respective axes.

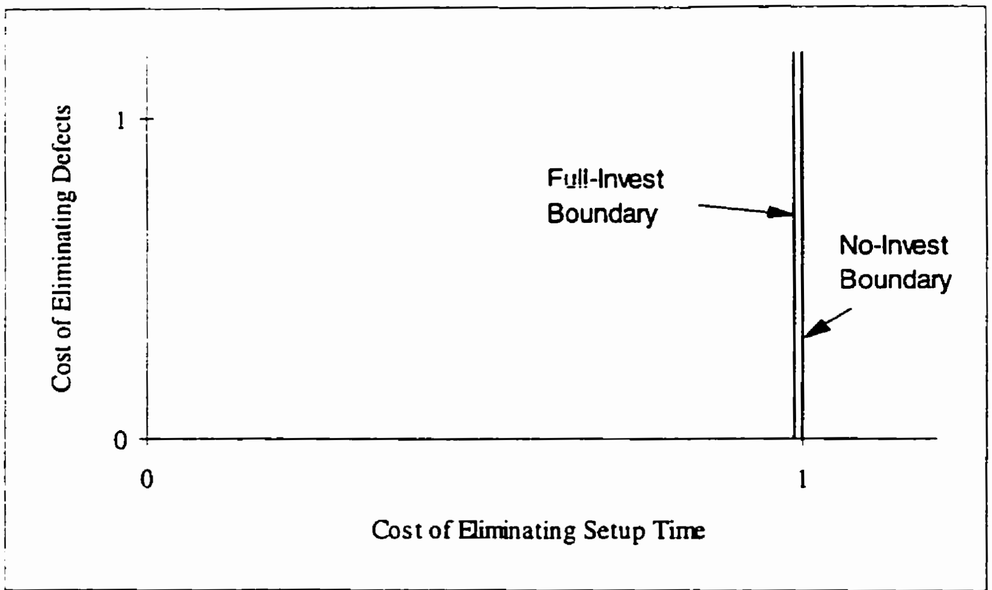


Figure 6.3: Normalized Decision Matrix for Optimal Investments in Setup, 'Naive' Case (costs are normalized such that the no-invest boundary falls at cost = 1).

6.1.2 Quality Optimization

Determining the optimal investment in quality improvement is done just as in Chapter 5. Investment boundaries, from Equations 5.2.7 and 5.2.8, are:

The 'no-invest' boundary occurs at R* = 1, or:

$$
\begin{aligned}
b_0^* \big|_{R^*=1} = {} & \frac{1}{i}\left[2ib_1 + D^2 r_0^2 E[X]^2 h(Q+1)\right] + \\
& \frac{1}{2i\left\{DQr_0E[X] + \left[D\tau_0 + DQE[X] - Q\right]\right\}} \Big\{ \\
& 2\left[D\tau_0 + DQE[X] - Q\right]\left[D^2 h r_0 \tau_0 E[X] - ib_1\right] - 2DQr_0E[X]ib_1 \\
& + hDr_0E[X]\left[2Q^2 - D^2\tau_0^2 - 2DQ^2E[X] + D^2E[X]^2\left\{(Q-1) + (Q+1)r_0\right\}\right] \Big\}
\end{aligned}
\tag{6.1.3}
$$

And the 'full-invest' boundary occurs at R* = 0, or:

$$b_0^*\Big|_{R^*=0} = \frac{1}{i}\left[2ib_1 + D^2 r_0^2 E[X]^2 h(Q+1)\right] + \frac{1}{iDQr_0E[X]}\Big\{$$

$$2\left[D\tau_0 + DQE[X] - Q\right]\left[D^2 h r_0 \tau_0 E[X] - ib_1\right] - 2DQr_0E[X]ib_1 \qquad (6.1.4)$$

$$+hDr_0E[X]\left[2Q^2 - D^2\tau_0^2 - 2DQ^2E[X] + D^2E[X]^2\{(Q-1) + (Q+1)r_0\}\right]\Big\}$$

Using numerical values from Table 3.2, these boundaries are plotted in Figure 6.4. Figure 6.5 completes the decision matrix for the 'naive' case by overlapping Figures 6.3 and 6.4 to the define regions of optimal investment for both setup and quality. This 'naive' case decision matrix is similar to those of Chapters 3 and 5, with the primary difference being the three regions involving 'some' investment in setup reduction are relatively narrow in this system.

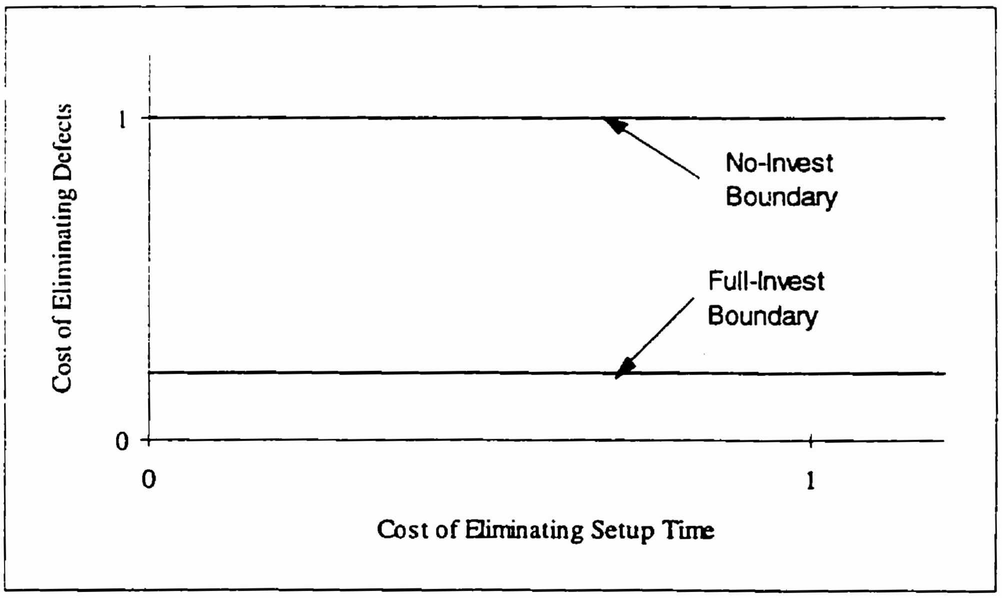


Figure 6.4: Normalized Decision Matrix for Optimal Investments in Quality, 'Naive' Case.

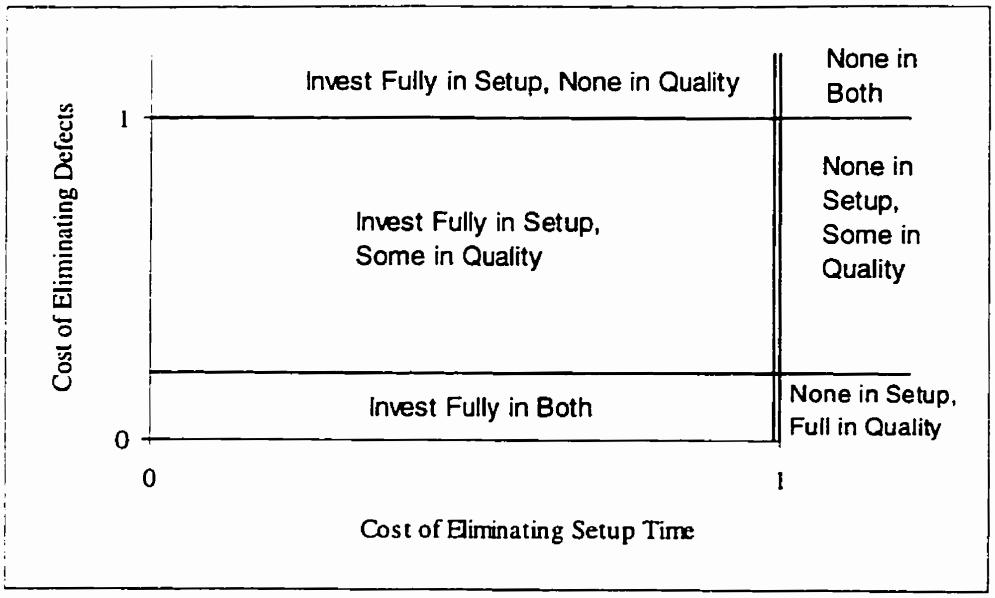


Figure 6.5: Completed Decision Matrix for 'Naive' Case.

6.2 'Informed' Case

In the 'informed' case, decisions are optimized simultaneously over the three variables, T, R and Q. Examining the sufficiency condition first, E[TC] is convex if Equations 6.1.1a and 6.1.1b are satisfied and the third principal minor, Δ_3, is non-negative, i.e.,

$$\Delta_3 = \begin{vmatrix} \frac{\partial^2 E[TC]}{\partial T^2} & \frac{\partial^2 E[TC]}{\partial T \partial Q} & \frac{\partial^2 E[TC]}{\partial T \partial R} \\ \frac{\partial^2 E[TC]}{\partial T \partial Q} & \frac{\partial^2 E[TC]}{\partial Q^2} & \frac{\partial^2 E[TC]}{\partial Q \partial R} \\ \frac{\partial^2 E[TC]}{\partial T \partial R} & \frac{\partial^2 E[TC]}{\partial Q \partial R} & \frac{\partial^2 E[TC]}{\partial R^2} \end{vmatrix} \geq 0 \qquad (6.2.1)$$

Evaluating this determinant analytically for sign is impractical; written out, 6.2.1 takes up 19 pages. Again, a Monte Carlo simulation was performed according to Algorithm 6.1, modifying step 4 to calculate the value of 6.2.1. The results of this simulation are summarized in Figure 6.6.

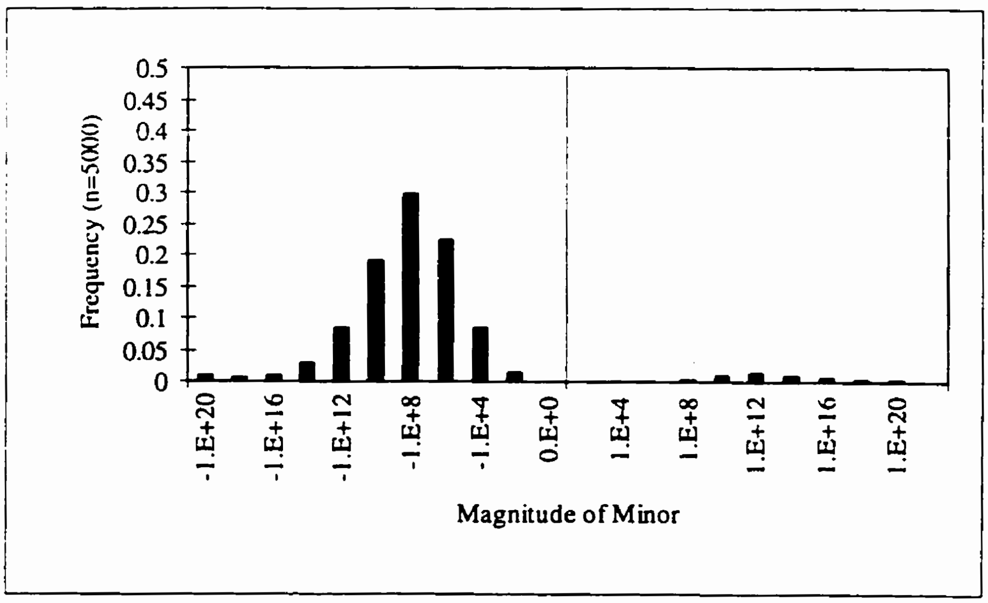


Figure 6.6: Frequencies of Third Principal Minor Observations from Monte Carlo Simulation

This principal minor is also negative the vast majority of the time (>94% in this example), indicating non-convexity of the E[TC] most of the time. However, the other two principal minors must be considered when determining convexity. In examining the 5000 cases from the Monte Carlo simulation, all three principal minors were simultaneously positive only twice (i.e., 0.04%). This result suggests that an optimal 'some-some' decision can exist, but is rare.

Similar to the 'naive' case, the decision boundaries can be calculated numerically given example parameter levels. Figure 6.7 shows the decision boundaries for setup investment while Figure 6.8 shows the boundaries for quality investment decisions. Overlaying Figures 6.7 and 6.8 yields the completed decision matrix for the 'informed' case, as given in Figure 6.9. This matrix shows the 'some-some' decision region occupies a narrow strip in the center of the matrix, consistent with the low frequency of E[TC] convexity predicted by the Monte Carlo simulation.

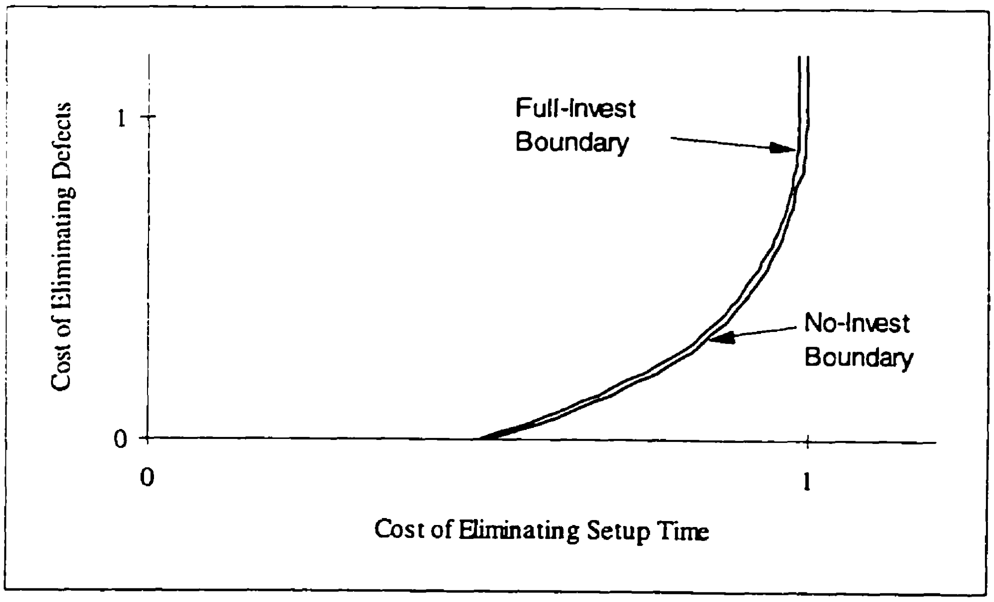


Figure 6.7: Setup Decision Boundaries for 'Informed' Case.

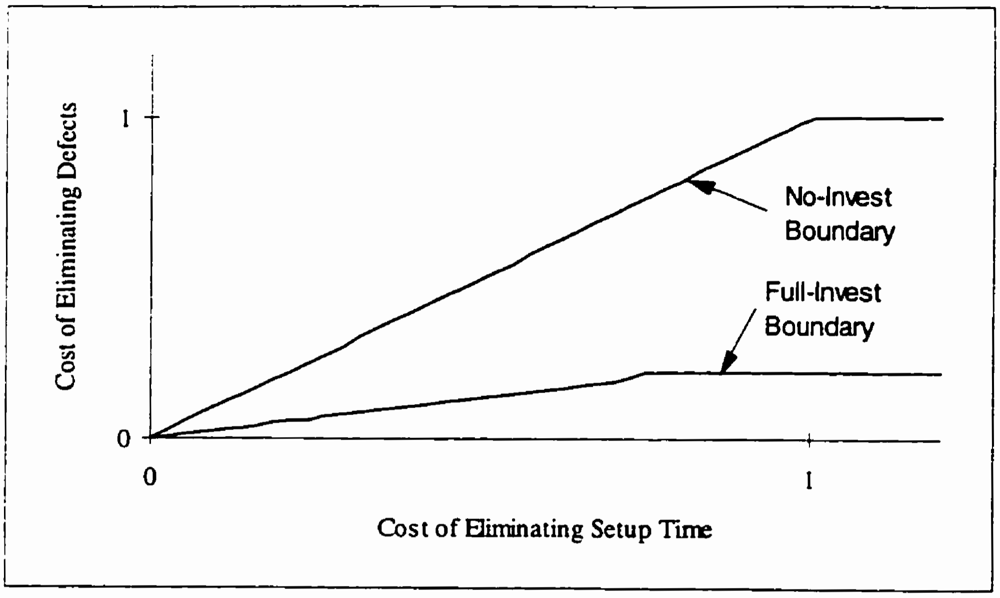


Figure 6.8: Quality Decision Boundaries for 'Informed' Case.

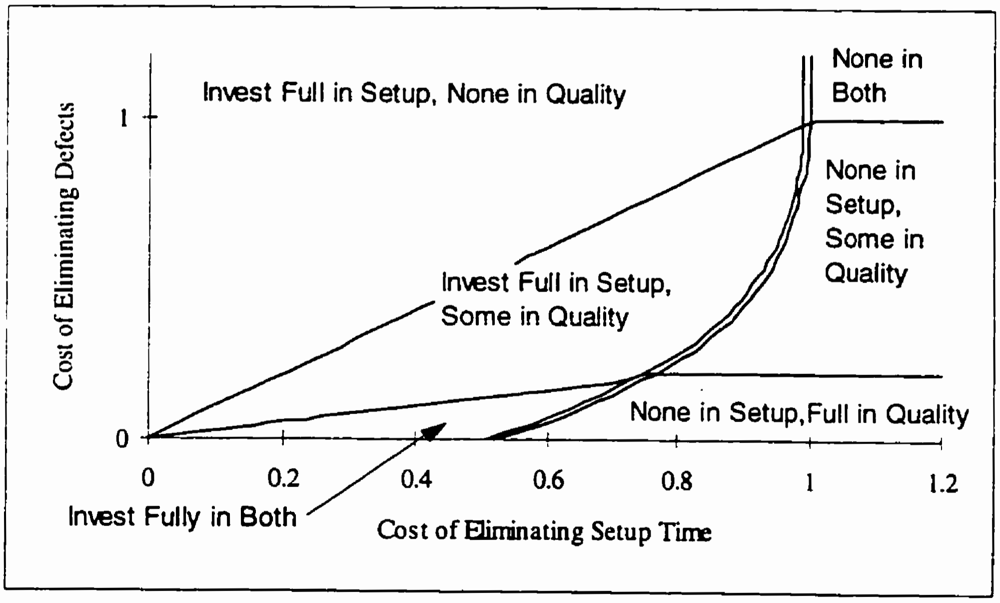


Figure 6.9: Completed Decision Matrix for 'Informed' Case.

Elements of the shape of the decision matrix in Figure 6.9 are also seen in the 'informed' case decision matrices of the previous chapters. For instance, on the right of the matrix, the three regions for investing 'none', 'some', and 'fully' in quality are consistent with the regions predicted analytically with other versions of the model. The "J" shape of the setup boundaries is consistent with the results of Chapter 4 where batch size was a variable but the linear investment function was used. Since one would expect that using the convex investment function would lead to the E[TC] function being more convex, intuitively, a region of 'some' investment in setup would be expected, as is found between the two setup boundaries. So, although this decision matrix was generated numerically for a single example, its shape is expected to be the generic shape for the variable batch/convex investment function model.

6.3 Decision Error and Sensitivity

6.3.1 Decision Error

Due to the complexity of the solution to this model, analytic determination of an expression for the 'decision error' has not been possible. Instead, a numerical evaluation for the example case from Table 3.2 has been performed. Using these example parameters, optimal investment decisions have been calculated as functions of the costs to improve setup and improve quality, under both the assumption of no interactions ('naive' case) and allowing interactions ('informed' case). These numerical results were then used to construct the respective decision matrices to permit the comparisons between cases in this section.

The difference in optimal investment decisions due to interactions can be illustrated by comparing the 'naive' and 'informed' decision matrices. To assist in making this comparison, the case of optimal setup decisions alone is considered first, followed by the case of optimal quality decisions alone. When combined, these two cases form the complete decision matrix.

Figure 6.10 shows the setup investment decision boundaries for both the 'naive' case from Figure 6.3 and the 'informed' case from Figure 6.7 overlaid on the same graph. The region between the straight ('naive' case) and curved ('informed' case) boundaries represents where the optimal investment decision changes. Since the 'informed' case boundaries always fall on or to the left (i.e., at lower costs of improvement) of the 'naive' case boundaries, ignoring interactions always results in equal- or over-investments, never under-investment.

Quality decision boundaries are compared in the same way. In Figure 6.11, the 'naive' and 'informed' boundaries for quality investment decisions from Figures 6.4 and 6.8, respectively, are overlaid.

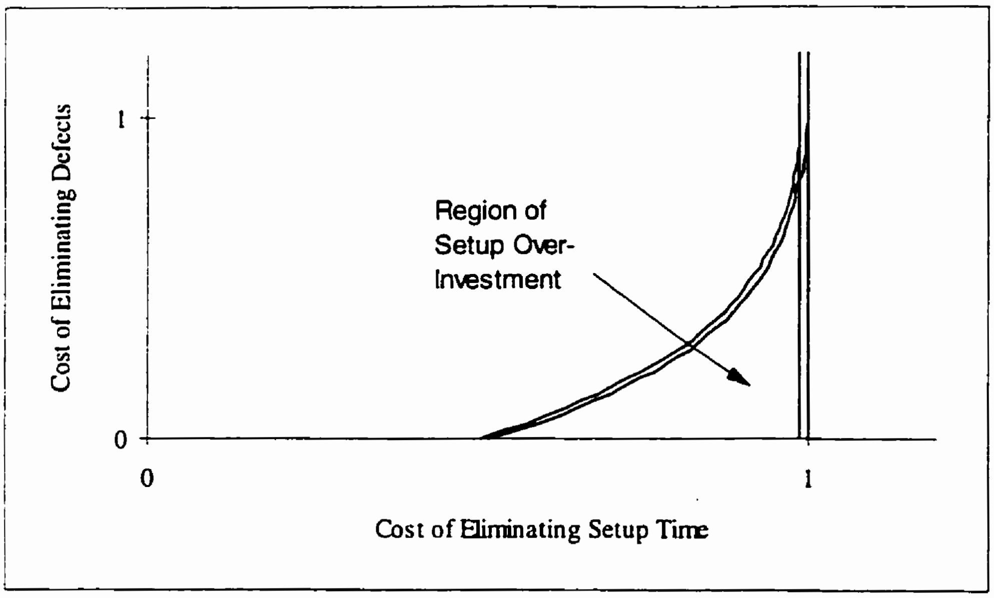


Figure 6.10: Regions of Setup Over-Investment, 'Naive' to 'Informed' Case

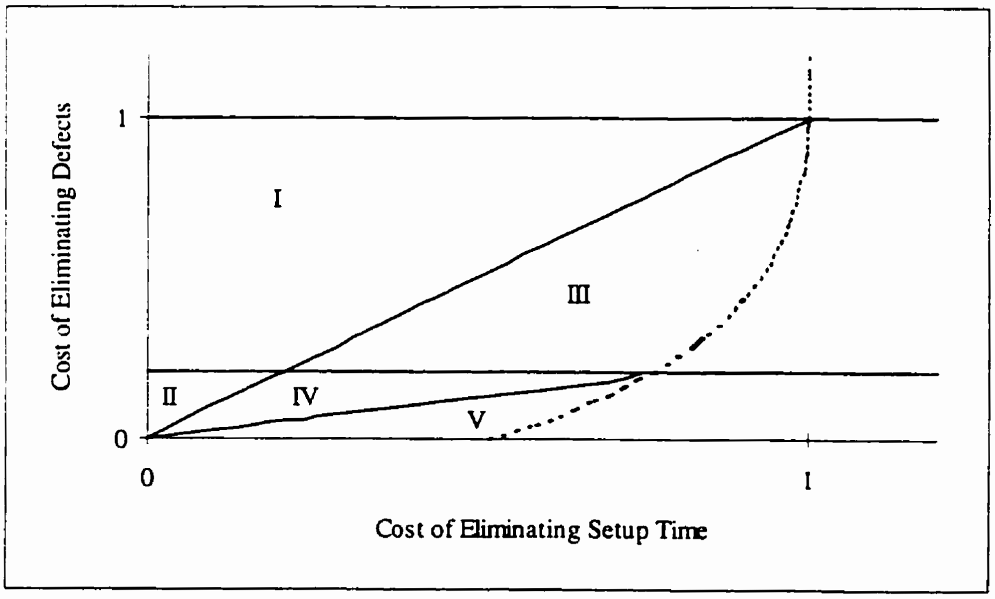


Figure 6.11: Regions of Quality Over-Investment, 'Naive' to 'Informed' Case

(Note: Roman numerals indicate regions described in the text.)

To assist in the description of the changes in optimal investment levels, five sub-regions are shown in Figure 6.11, numbered I through V (the dotted line represents the setup no-invest boundary, which is where 'informed' case quality investment decisions start to deviate from 'naive' case decisions). The five sub-regions can be described as:

Sub-Region	'Naive' Case Investment	'Informed' Case Investment
I	'Some'	'None'
II	'Full'	'None'
III	'Some'	'Some', but less than 'Naive' Case
IV	'Full'	'Some'
V	'Full'	'Full'

In four of these five sub-regions, ignoring interactions leads to higher investment levels then when they are considered, while equivalent investments result in the fifth sub-region. Again, the result of ignoring interactions is making investments of the same or greater magnitude than when they are considered.

The magnitudes of potential over-investments have been tabulated for this example and are summarized in Table 6.1. As in previous chapters, these values are calculated by comparing optimal investment levels predicted by the two cases over the area of the decision matrix where the 'naive' strategy called for investing 'some' or 'fully' in both improvement practices. This form of the model, with convex investment functions and variable batch size, also shows significant levels of over-investment are possible.

Table 6.1: Frequency of Over-investment Levels

Over-Investment Level (percentage of optimal investment)	Frequency of Observances (n = 2601)
< 5%	19%
5 - 50%	26%
50 - 100%	28%
100 - 150%	6%
150 - 200%	3%
200 - 250%	2%
250 - 300%	3%
300 - 350%	4%
350 - 400%	2%
400 - 450%	1%

6.3.2 Sensitivity Analysis

Similar to the models in the previous chapters, the decision error from ignoring interactions with this version of the model was examined for sensitivity to the various parameters in the model. The first parameter studied was demand. High and low levels were as given in Table 3.2. The resulting distribution of over-investment are given in Figure 6.12.

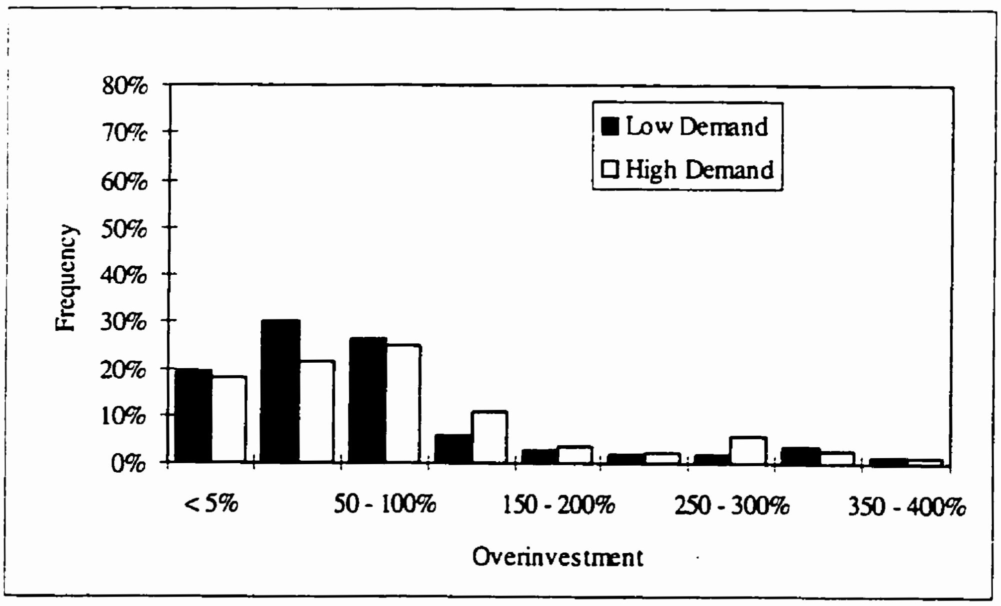


Figure 6.12: Distribution of Over-investment as a Function of Demand

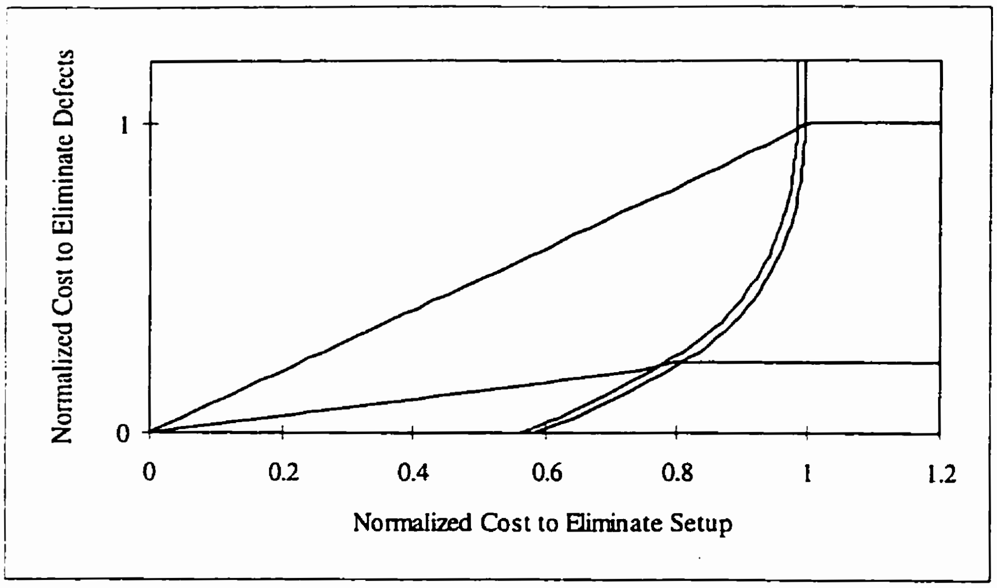


Figure 6.13: 'Informed' Decision Matrix for Low Demand

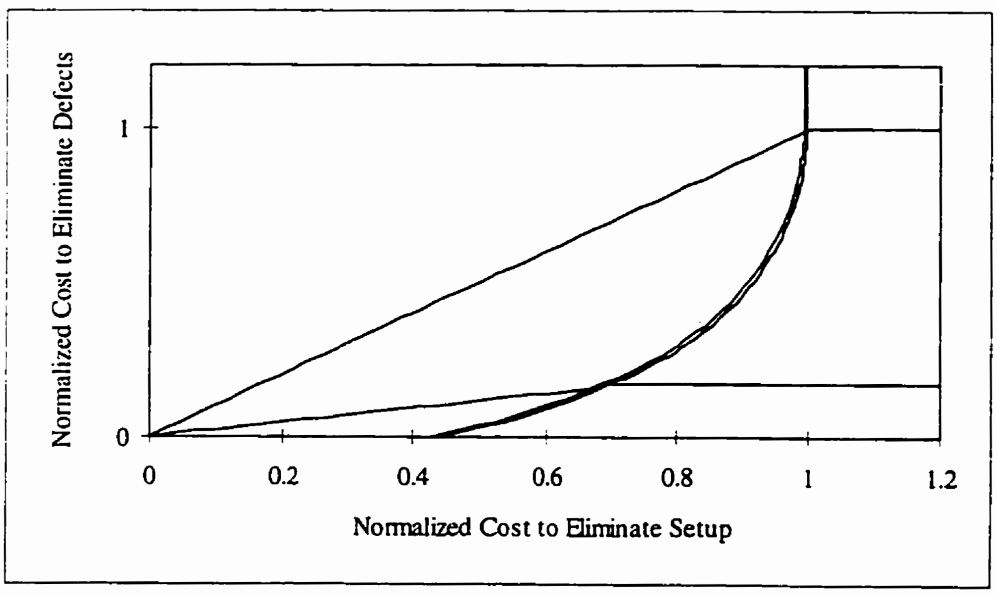


Figure 6.14: 'Informed' Decision Matrix for High Demand

From Figure 6.12, there does seem to be sensitivity to demand, with increasing demand shifting the distribution of over-investment to higher levels. To investigate the reason behind this shift, the decision matrices for the 'informed' cases with low and high demand are given as Figures 6.13 and 6.14, respectively. Three primary differences are seen in these figures. First, the curvature of the setup boundaries (e.g., the "J" curves) increases with demand, and the boundaries intersect the axis at a location of lower cost of making improvements.

The second difference is that the 'full-invest' boundary for quality investment moves down, to a position of lower cost of making improvements, just as the boundary for setup had done.

The third difference between the matrices is that the regions for 'some-invest' in setup have decreased in size. This would suggest the E[TC] function becomes less convex as Demand increases.

To help see how these changes affect the decision error, the 'naive' and 'informed' case boundaries for setup are plotted in Figure 6.15 for both low and high Demand levels. Since the area enclosed by these curves represents the region where setup over-investment takes place, the size of the region corresponds to a likelihood of over-investment in setup. Figure 6.15 shows that as Demand goes from low (dotted line) to high (solid line), the size of this region increases.

The corresponding graph has been drawn for quality improvements, and is given as Figure 6.16. Since, when normalized, the quality 'no-invest' boundary goes from the points (0,0) to (1,1) regardless of Demand level, only the 'full-invest' boundaries have been drawn. The region where decision error occurs is that triangular area bounded by the horizontal and diagonal components of the decision boundaries, and by the quality cost axis. In this instance, going from low to high Demand has led to a decrease in the area of the region where quality decision error occurs.

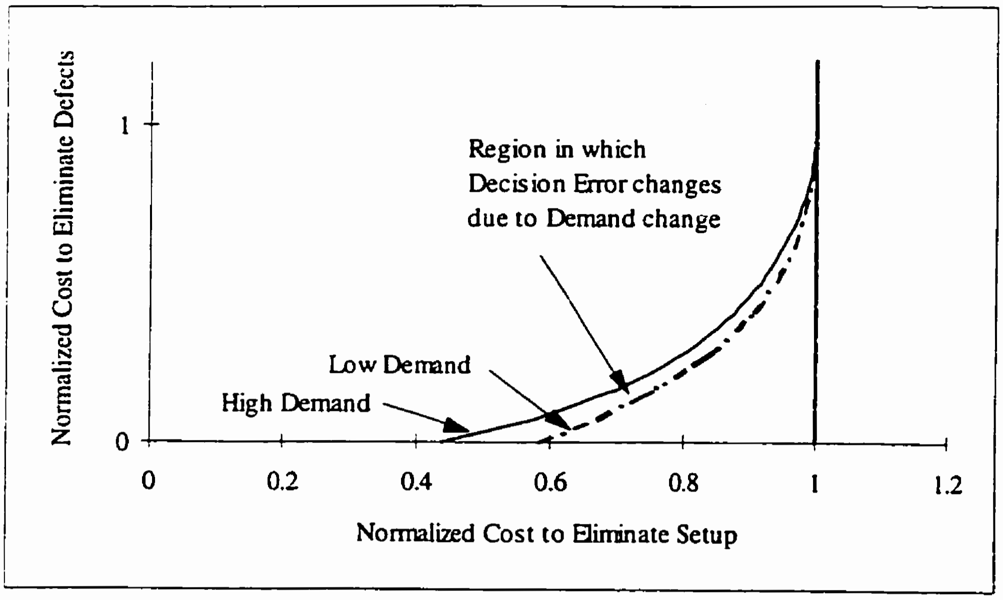


Figure 6.15: Differences in Setup Boundaries, Low to High Demand

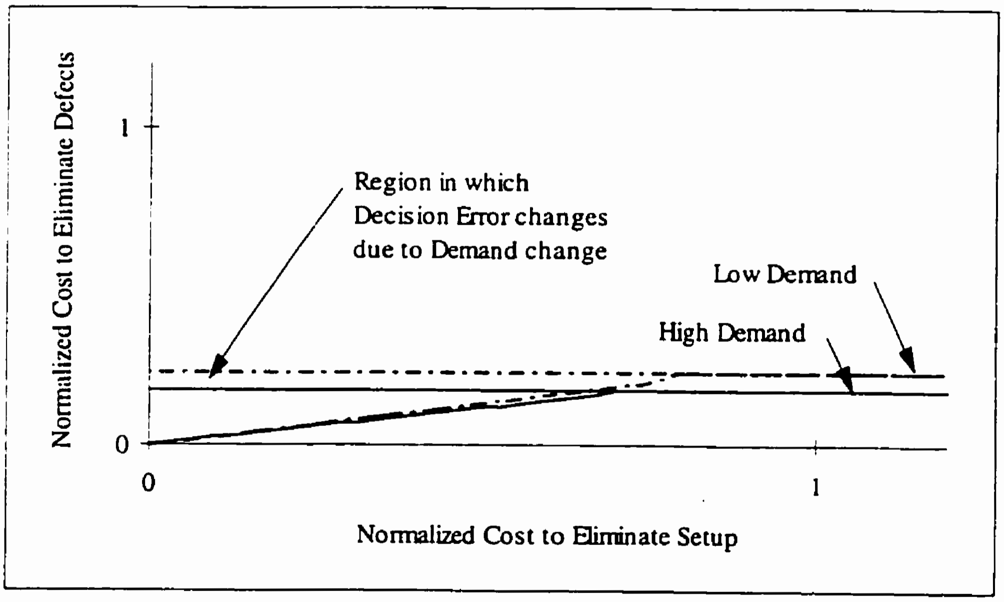


Figure 6.16: Differences in Quality Boundaries, Low to High Demand

Since an increase in demand simultaneously increases the size of the region where setup decision error occurs, but decreases the size of the region where quality decision error occurs, no clear conclusion can be drawn on the sensitivity of decision error to demand. It would appear that the sensitivity would depend on the relative changes in the sizes of these areas, which is not known in general from the examples calculated here.

The distribution of decision errors for the next parameter, initial defect rate, r_0, is given in Figure 6.17. This distribution shows a more significant shift from lower levels of over-investment to higher levels as r_0 goes from low to high. To examine the possible reason for this, the 'informed' decision matrices for the low and high initial defect rate cases are presented as Figures 6.18 and 6.19, respectively. These decision matrices show the same behaviour as those for changes in demand; the curvature of the setup boundary increases, the 'full' investment boundary for quality drops and the 'some' investment regions for setup become more narrow. These behaviors imply the same changes to the areas of the decision error as when demand was varied: the size of the region of setup over-investment increases while the size of the region of quality over-investment decreases. No general conclusion can be drawn about the sensitivities of the decision error to the initial defect rate.

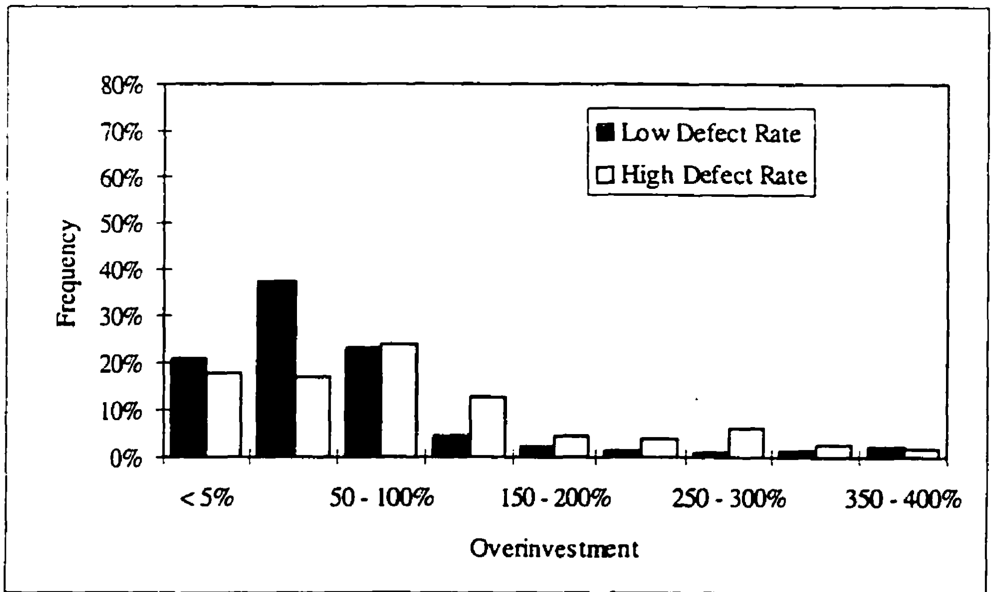


Figure 6.17: Distribution of Over-investment as a Function of Defect Rate

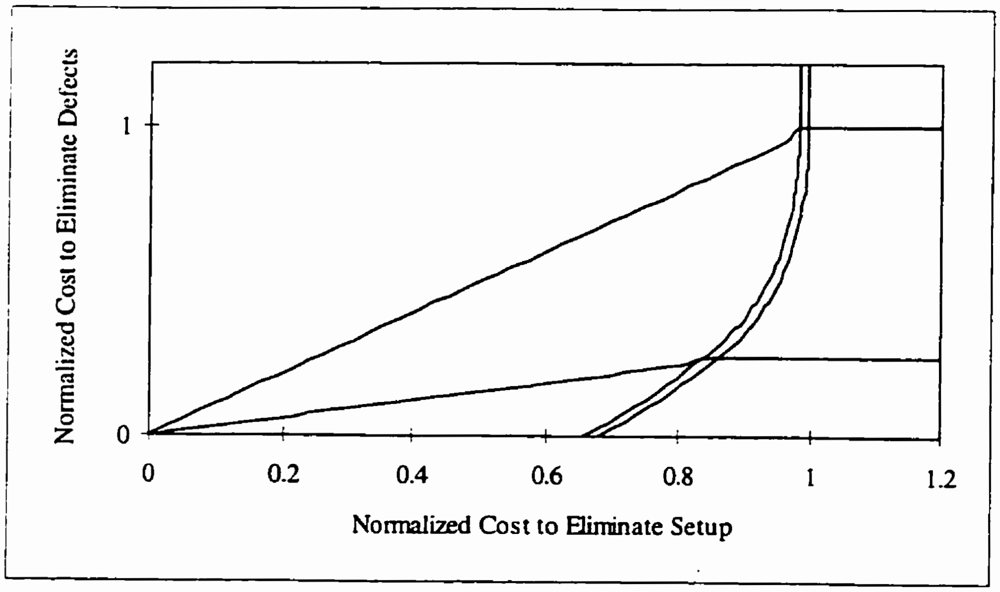


Figure 6.18: 'Informed' Decision Matrix for Low Initial Defect Rate

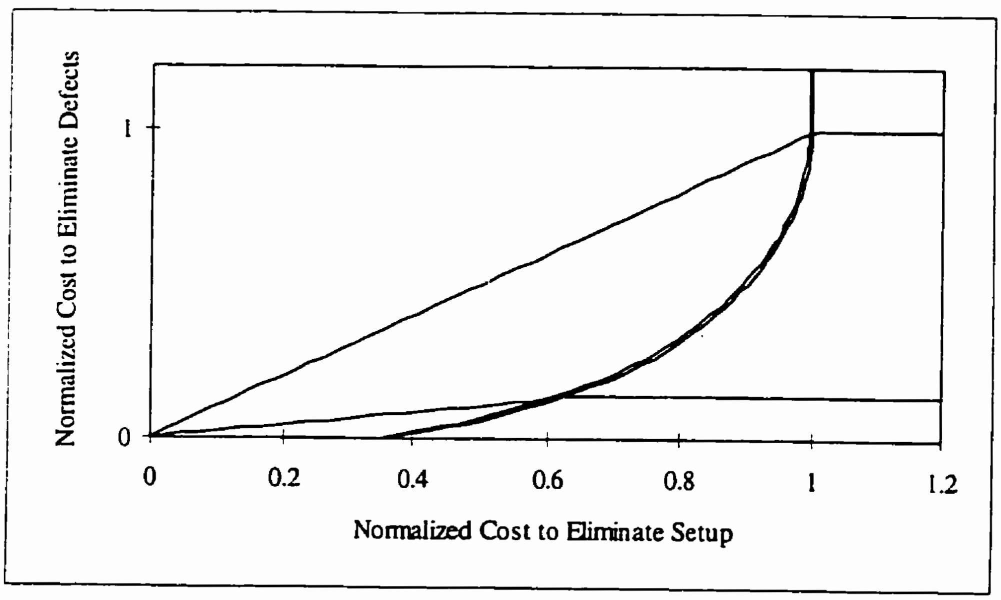


Figure 6.19: 'Informed' Decision Matrix for High Initial Defect Rate

The sensitivity of the over-investment distributions to changes in initial setup time, τ_0, is shown in Figure 6.20. Essentially no sensitivity is seen. The decision matrices for low and high initial setup time are given as Figures 6.21 and 6.22, respectively, and appear to be very similar, if not identical. It is hypothesized that the reason for this indifference is that investment in setup is almost binary: either no investment or full investment occurs. Since these decision matrices are normalized at the point where investment starts to take place, and since over-investment is calculated in relative terms, the scaling performed to create these measures factors out the sensitivity to initial setup times.

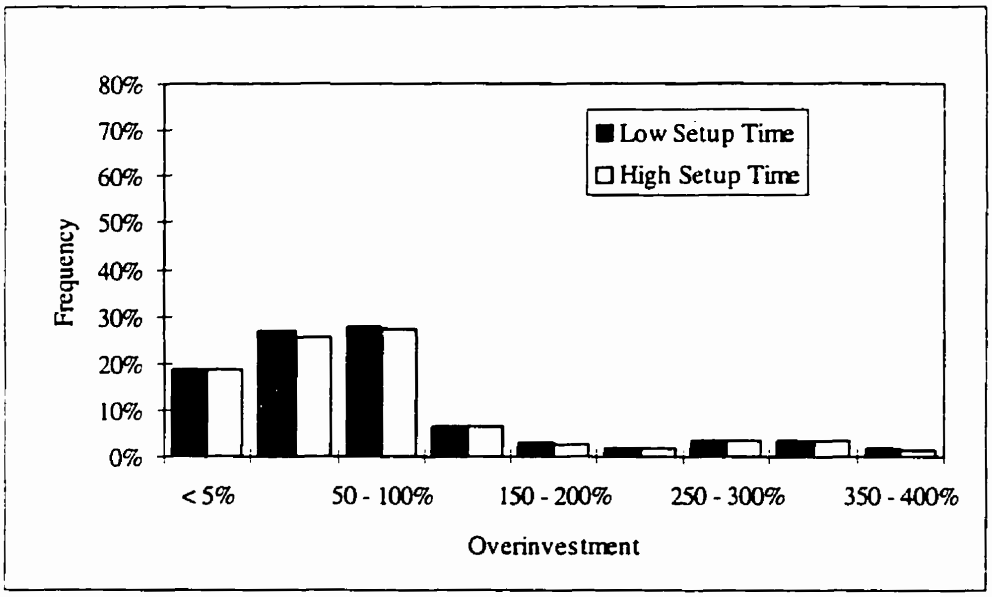


Figure 6.20: Distribution of Over-investment as a Function of Setup Time

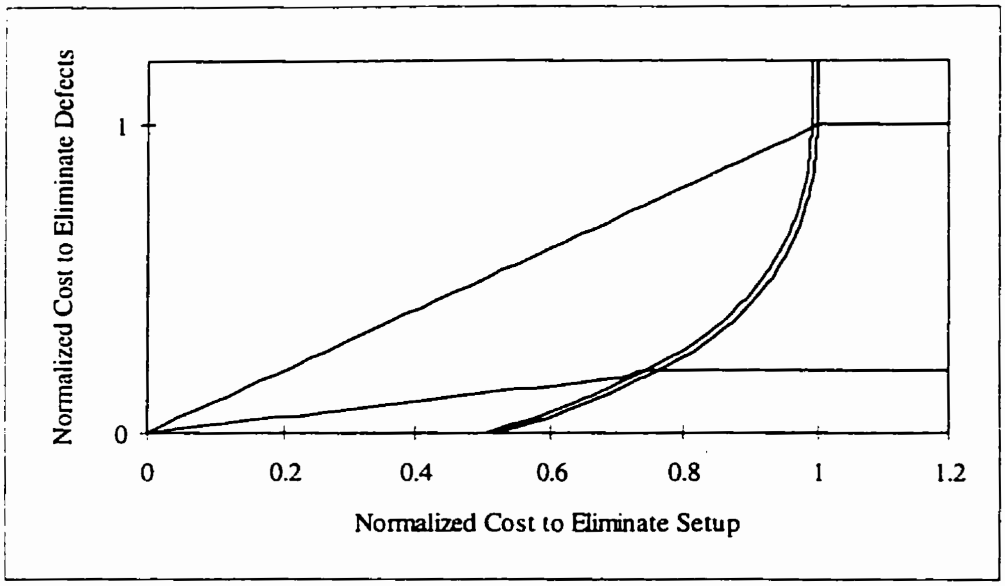


Figure 6.21: 'Informed' Decision Matrix for Low Initial Setup Time

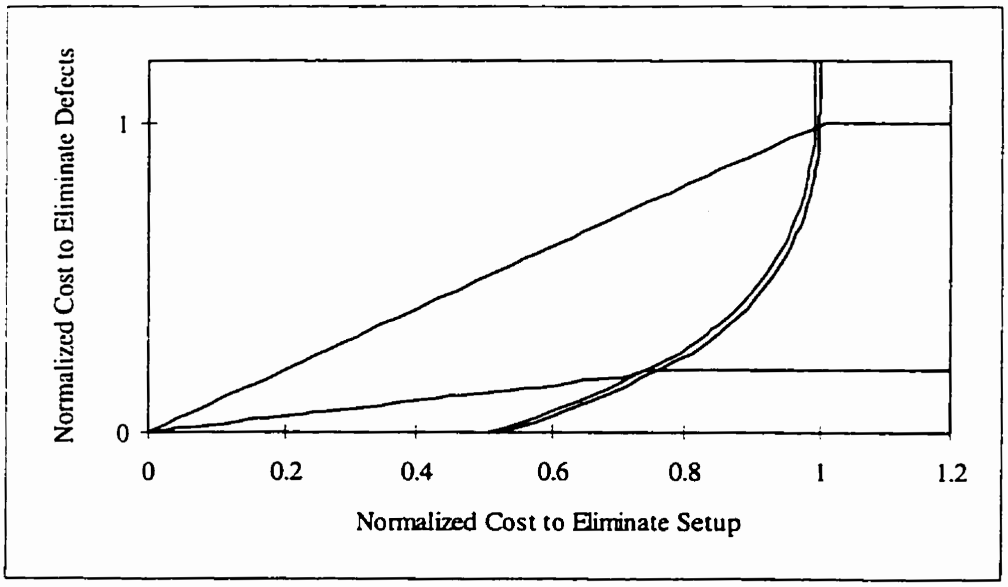


Figure 6.22: 'Informed' Decision Matrix for High Initial Setup Time

Figures 6.23 and 6.24 give the distributions of over-investment for low and high holding cost rates and interest rates, respectively. The (relatively small) sensitivities seen are the same as those for the other forms of the model in previous chapters: the sensitivities are in opposite directions with lower holding costs leading to lower average over-investments.

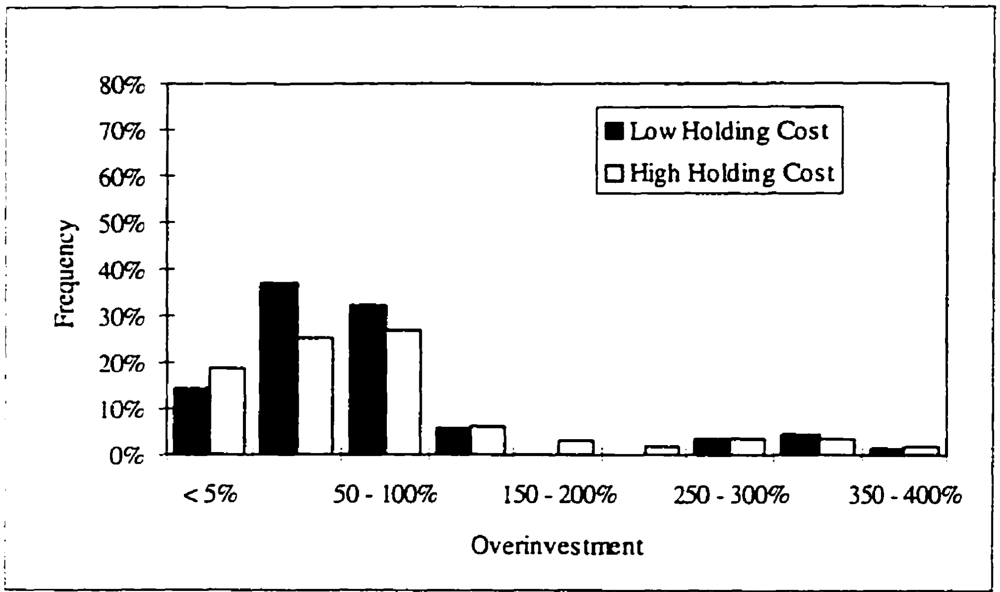


Figure 6.23: Distribution of Over-investment as a Function of Holding Cost Rate

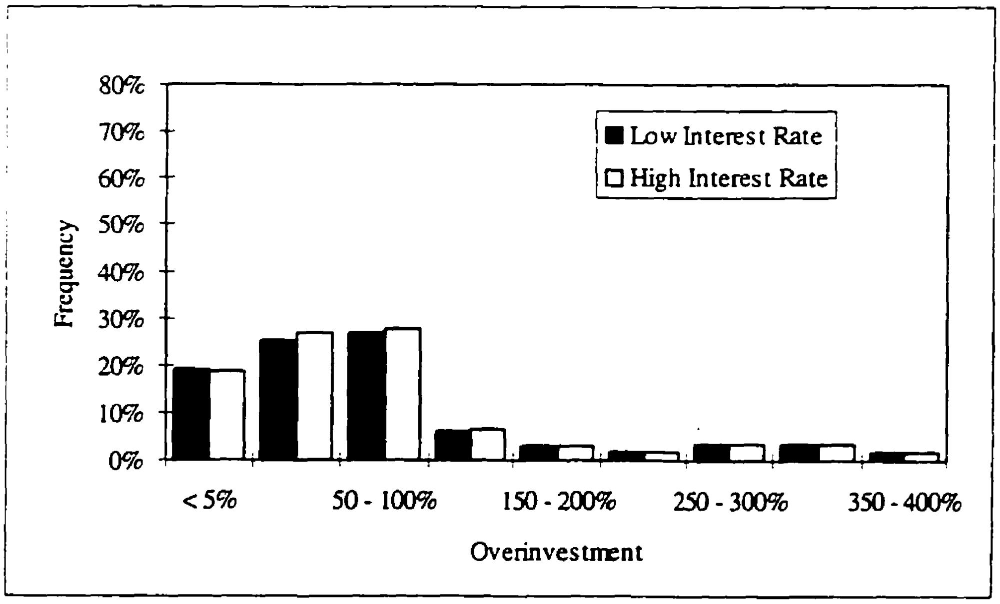


Figure 6.24: Distribution of Over-investment as a Function of Interest Rate

The sensitivity with respect to processing time coefficient of variation is given in Figure 6.25. Just as in the other chapters, batch service time coefficient of variation is relatively insensitive to this parameter.

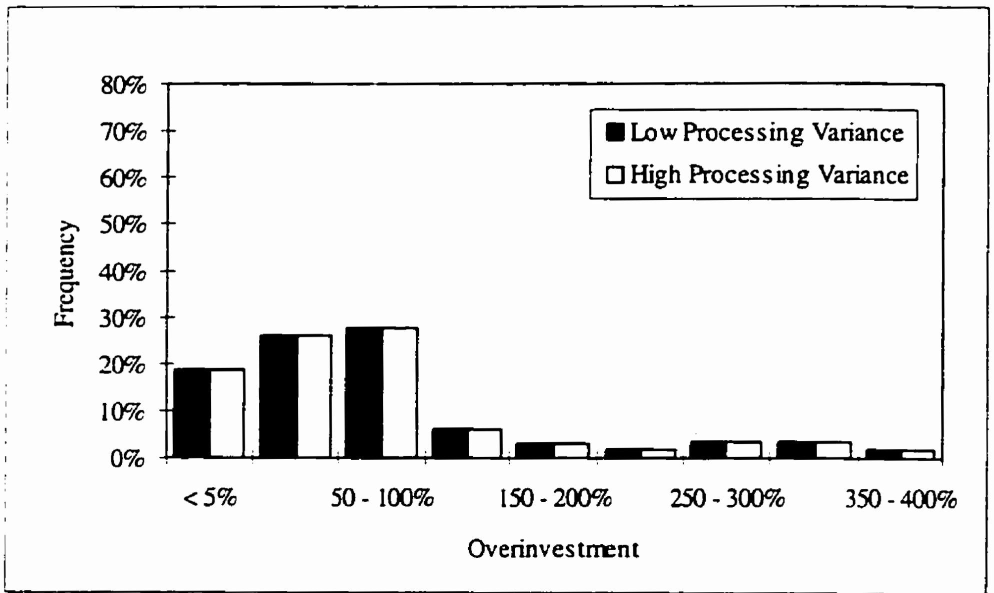


Figure 6.25: Distribution of Over-investment as a Function of Processing Time Variance

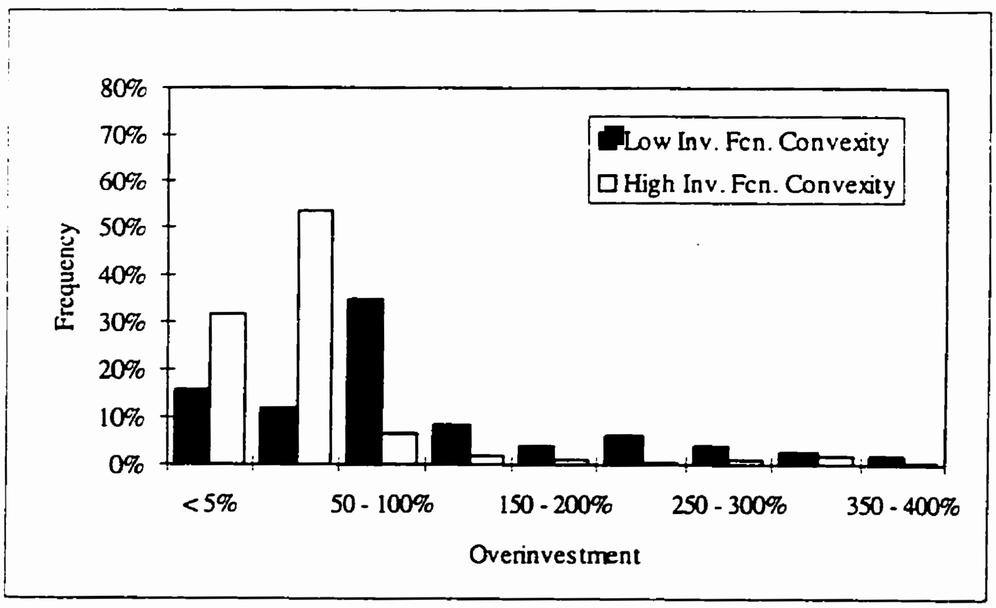


Figure 6.26: Distribution of Over-investment as a Function of Investment Function Convexity

The last parameter to be considered is the convexity of the investment functions. Figure 6.26 shows the distributions of over-investment as the investment function convexity was varied from low to high. From this graph, there is significant sensitivity to this parameter.

To help understand why, the 'informed' decision matrices are given as Figures 6.27 and 6.28. These figures show significant differences, with the most dramatic being the size of the 'some-some' investment region. In the low convexity case this region appears to have collapsed completely, while in the high convexity case it occupies the largest area of the decision matrix. This behaviour is expected. As the convexity of the investment function decreases, the function will eventually become linear, which is the case studied in Chapter 4. In that form of the model, investment in setup reduction was either to invest enough to reduce batch size to one unit, or no investment was optimal, i.e., there was essentially no 'some' investment decision for setup. This behaviour is represented in Figure 6.27 by the lack of regions for 'some' investment in setup. A difference between Figure 6.27 and Figure 4.6 for the 'informed' decision matrix from the model using a linear investment function is that the quality boundaries in the 'invest-fully' in setup region were much lower than they are here. It can be noted, however, that the quality boundaries in Figure 6.27 are discontinuous across the setup decision boundary, and that in the limit as the investment function became linear, could tend towards the decision matrix developed in Chapter 4.

In Figure 6.28, the 'some-some' decision region has expanded to fill much of the decision matrix. This result can be explained by considering what would happen if the investment function became very convex. In such a limit, the cost of making the last increment of improvement to setup or quality would reach infinity, so the decision to 'invest-fully' in setup or quality would never be optimal. The decision matrix would then include only two decisions: 'invest-none' in practice, or 'invest-some', and the 'some-some' region would fill the area from the origin to the (1,1) point on these normalized graphs.

Considering the sizes of the areas on these figures where over-investment takes place, it can be seen that as investment function convexity increases, the size of the regions where over-investment takes place decrease. This trend suggests that the distribution of over-investments shown in Figure 6.26 is explained by the expected sensitivity of the system to changes in investment function convexity.

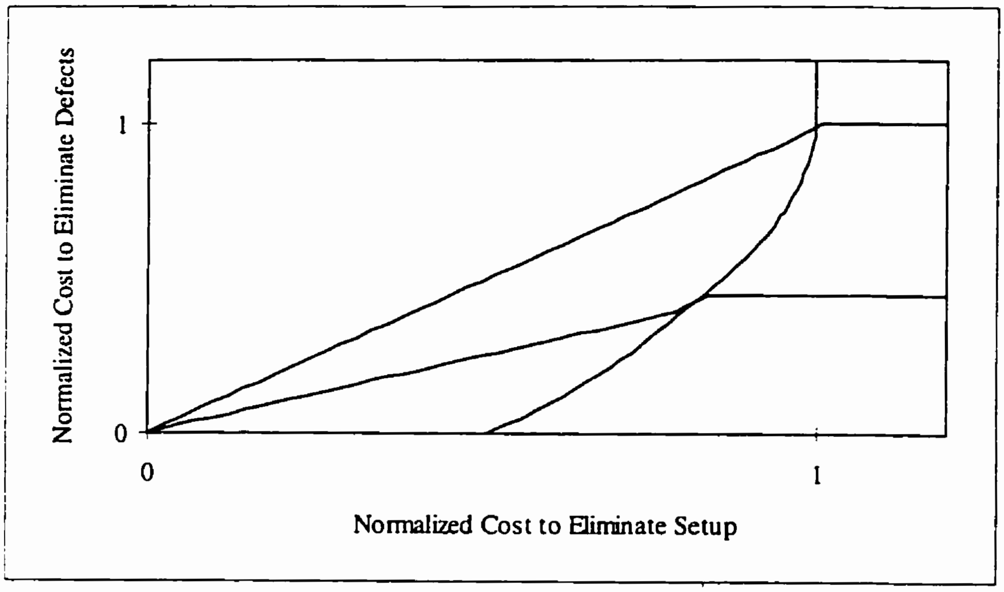


Figure 6.27: 'Informed' Decision Matrix for Low Investment Function Convexity

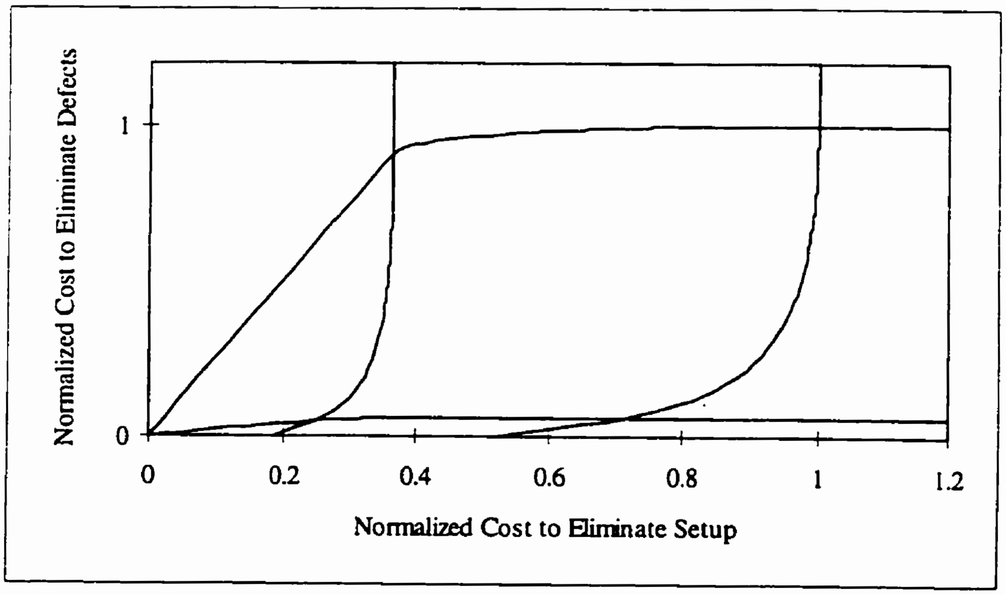


Figure 6.28: 'Informed' Decision Matrix for High Investment Function Convexity

6.4: Discussion

This chapter examined the most complex case in this dissertation: a model with variable batch sizes and a convex investment function. Due to the complexity of the system, few analytic results were obtained and numerical examples had to be relied upon.

Once again, a 'naive' case in which optimal investments in setup time (with batch size as a secondary optimization variable) were determined independently of optimum investments in quality improvement, was compared to an 'informed' case in which the three decision variables were optimized simultaneously. Just as in the previous three chapters, setup and quality improvements were found to interact, with the interaction leading to a significant chance of over-investment if it is ignored.

The behaviour of the 'informed' decision matrix is consistent with that seen in the previous chapters. There is a primary, "J"-shaped decision boundary through the matrix, similar to that found in Chapter 4 where batch sizes were variable, but the investment function was linear. The behaviour of the optimum investment in setup reduction similarly tended towards extreme points, with optimum decisions to invest either 'fully' or 'none', but with a narrow region where 'some' investment was called for. This behaviour was again due to the influence of the variable batch size. As setup time reduction takes place, optimum batch size drops which provides additional benefit for more setup reduction.

The relationship between optimal investments in setup and quality, however, were not as extreme as that found in Chapter 4. In that chapter, if investment took place in setup reduction, there was essentially no investment called for in quality improvement. In this chapter, there was investment in quality in the region where the optimal decision is to invest fully in setup reduction, although the size of the sub-regions for quality investment were smaller than the comparable regions in Chapter 5 (convex investment function, fixed batch size). This difference is a result of the use of the convex investment function in the model of this chapter. The convex investment function leads to the E[TC] function being convex under some circumstances. When this happens, 'some' investment in setup and quality is optimal.

Also similar to Chapter 4 were the results of the sensitivity analysis. Decision Error had only limited sensitivities to the problem parameters, with the exception of the convexity of the investment function, to which it was very sensitive. Examination of the decision matrices under conditions of low and high investment function convexity showed dramatic changes in relative areas of the various regions. As convexity increased, so did the sizes of the areas representing 'some' investment as optimal. (In the other direction, as convexity decreased, the decision matrix approaches that of Chapter 4, in which the linear investment function was used.) A consequence of this shift in area sizes is that the prospect of the decision maker over-investing if interactions are ignored decreases since the areas of the decision matrix where Decision Error occurs shrink as convexity of the investment functions increase.

These results further underscore a conclusion that the shape of the investment function used is very important to the investment decision making process, and has a very large impact upon the prospect for over-investing and the magnitudes of the over-investments.

The final chapter, Chapter 7, summarizes the results of this dissertation, draws conclusions and makes recommendations for future research.

Chapter 7 Discussion, Conclusions and Further Research

This dissertation has studied the role of interactions between two simultaneously implemented improvement practices, namely setup reduction and quality improvement, on optimal investment decisions. A model of a manufacturing cell with WIP inventories was created, and four versions of the model were examined in Chapters 3 through 6, leading to a number of results. This chapter summarizes the results of this research, draws conclusions and offers some directions for future research.

7.1 Discussion

To examine the effects of possible interactions on investment decisions, a new model was developed based on an M/G/1 queuing model with setup time and quality level as the primary decision variables. In contrast to the previous EOQ-based models, which based benefits from improvements on savings of finished goods cycle-stock inventories, this queuing-based model estimated expected levels of WIP inventories and costs. WIP is expected to provide a more representative reflection of overall manufacturing inventory costs than finished goods cycle stock [e.g., Boucher, 1984; Primrose, 1992]. The development of this model was presented in Chapter 3.

Two key assumptions made in the model deal with the treatment of the batch size variable and the assumed relationship between levels of investments and the resulting levels of improvements to the system. Because these assumptions were expected to have a significant impact upon the model predictions, different variations of the model were created with different treatments of these factors. Two types of batch size were considered: batch size as a fixed parameter in the model and batch size as a decision variable. Similarly, two types of investment function were considered: linear (i.e., constant returns to scale) and convex (i.e., decreasing returns to scale). These two factors of two levels each resulted in four possible combinations, each of which has been studied through a different variant of the model. The four model variants were each developed and analysed in separate Chapters:

- Fixed batch size, linear investment function (Chapter 3).
- Variable batch size, linear investment function (Chapter 4).
- Fixed batch size, convex investment function (Chapter 5).
- Variable batch size, convex investment function (Chapter 6).

To assist in presentation of the model results, new terms were advanced. For instance, it was found that investment decisions for each improvement practice could be classified into three categories: a) 'invest none', in which no investment takes place, b) 'invest fully', in which a maximum possible investment is made, or, c) 'invest some', for optimum investments between the previous two categories. It was found that the optimal investment decision for a given improvement practice can be expressed as a function of cost to fully improve that practice, which is a quantity easily determined from the investment function. Using this cost, the two interfaces between the three categories (i.e., between 'invest none' and 'invest some', and between 'invest some' and 'invest fully' decisions), termed the 'decision boundaries', can be described as a function of the other problem parameters. These boundaries represent the locations where the optimal investment decisions for a practice change as the cost of improving that practice increases or decreases.

As the objective of this research was to examine the effects of possible interactions on investment decisions, comparisons were made between optimal decisions predicted for two cases. The first case, the 'naive' case, determined optimal investment decisions based on the assumption that the two practices were independent of each other. This case did not include the effects of any interactions in predicted decisions. The second case, the 'informed' case, optimized the model simultaneously over the two practices. With simultaneous optimization, any interactive effects were included in optimal decisions. Differences between the optimal investment predictions of each case are then assumed to be the result of interaction.

Since there were two improvement practices, these categories of optimal decisions can be plotted on a two-dimensional graph, called the 'decision matrix.' By plotting the decision boundaries for each practice against the costs of making improvements to each practice, a decision matrix is developed. The decision matrix shows all optimal investment decisions as regions on the graph (e.g., 'invest none' in setup and 'invest some' in quality, etc.). By comparing the sizes and shapes of the corresponding regions between decision matrices for different models, changes to the structure of optimal investment decisions can be quickly determined.

Differences in the areas of the various regions between the decision matrices are the result of interactive behaviour. The interpretation of theses differences is that one investment decision is optimal under a given set of circumstances if interactions are ignored (the 'naive' case), but a different investment is optimal if interactions are considered (the 'informed' case). This discrepancy in optimal decisions is termed the 'decision error', and the different model variants of each analysis chapter have been used to develop relationships useful for examining this concept of 'decision error.'

The mathematical complexity of each model variation increases as one goes down the list given above. In the first case, fixed batch size and linear investment function, optimal decisions and decision errors were determined analytically, while in the last case, variable batch size and convex investment function, most of the results were obtained through numerical means. For instance, in Chapter 3, relationships for the location of the decision boundaries on the decision matrix were developed analytically. Formulas were also developed for 'boundary shift', that is, how much each decision boundary moves to a location of lower costs when investments are made in the other practice (i.e., as a consequence of interactive behaviour). These formulas show that this is the generic behaviour for the system with a linear investment function and fixed batch size, regardless of parameter settings. Figure 7.1 illustrates the concept of 'boundary shift' for the 'no invest' boundary on setup reduction.

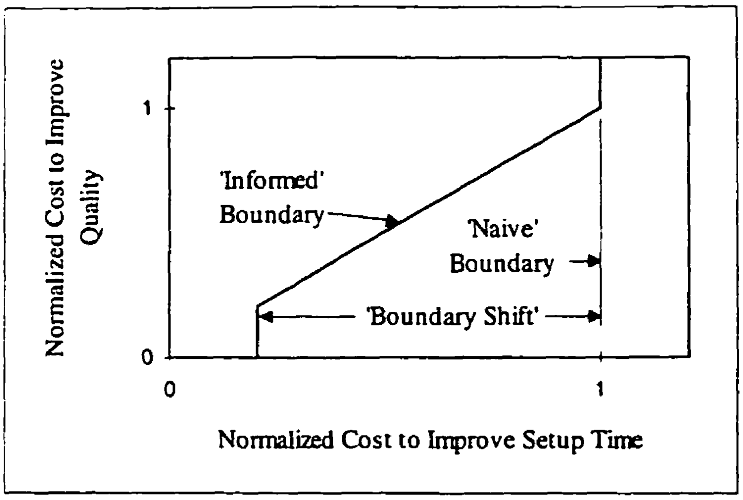
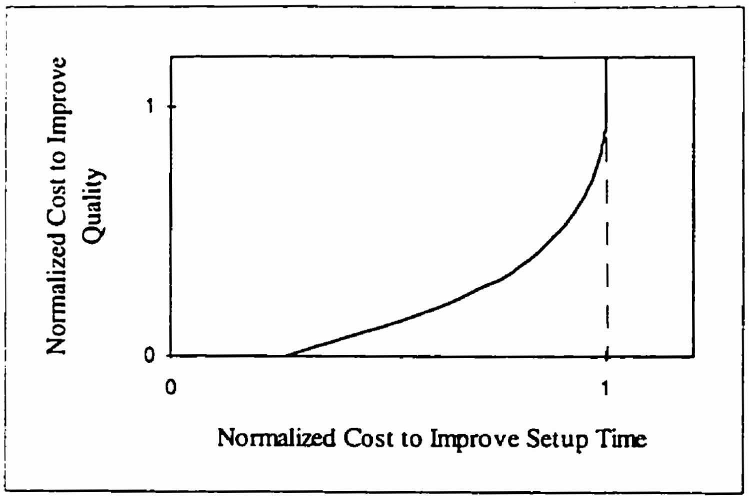
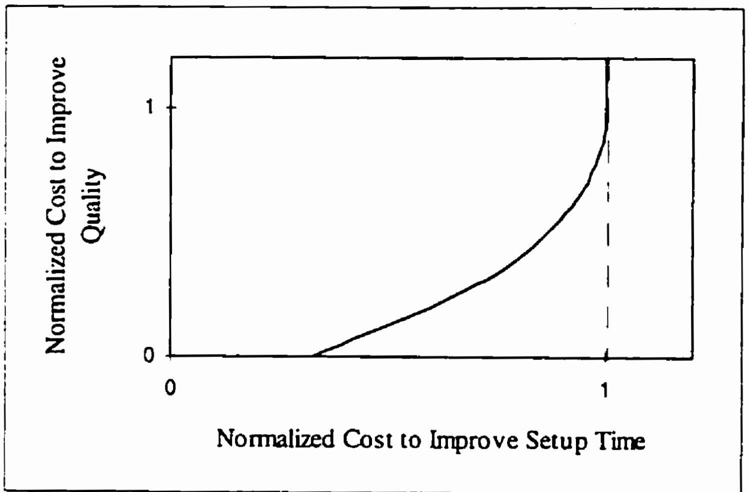


Figure 7.1: 'Boundary Shift' for Setup 'No Invest' Decision Boundary
(for linear investment function/fixed batch size model of Chapter 3)

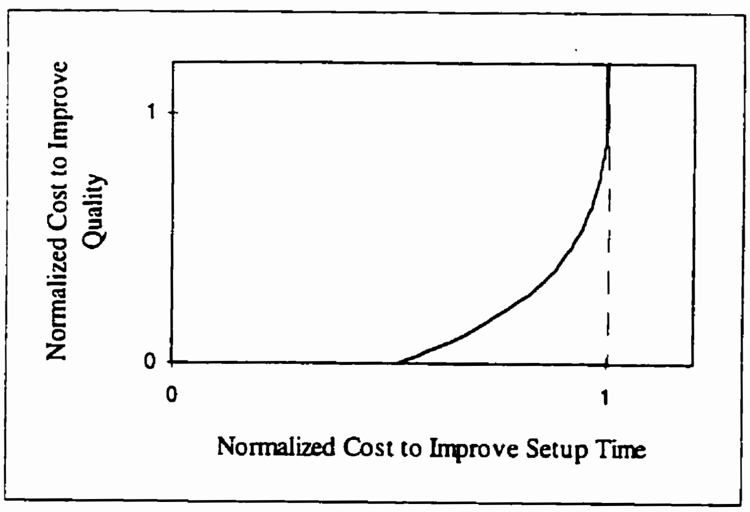
In the model variations of the remaining chapters, it became necessary to develop increasing portions of model results though numerical examples. While numerical examples do not provide general results, patterns consistent with those developed analytically in Chapter 3 were seen throughout Chapters 4 through 6. For example, consider the concept of 'boundary shift'. In Chapter 3, analytic results showed that the decision boundaries for one practice shift toward lower cost values as the cost of improving the other practice decreases. In Chapters 4 and 5, closed-form solutions were developed for the extreme ends of the decision boundaries which also showed this behaviour, although the center section of the boundaries had to be plotted numerically. In Chapter 6, only numerically-derived boundaries could be obtained, which also showed this behaviour in all examples studied. The similarity across the findings for each case provides evidence that this 'boundary shift' is a generic behaviour. Illustrations of the 'boundary shift' for the setup reduction 'no invest' boundary are given in Figures 7.2a ,b and c for the models of Chapters 4, 5 and 6, respectively.



(a) Linear Invest/Variable Batch



(b) Convex Invest/Fixed Batch



(c) Convex Invest/Variable Batch

Figure 7.2 'No Invest' Setup Reduction Boundaries Showing 'Boundary Shift'

Indeed, an intuitive explanation can also be given for 'boundary shift'. Consider one practice, setup reduction. If the cost of improving the other practice, quality improvement, is very high, no improvement will be made in the quality improvement. All the 'waste' in the system (i.e., WIP holding costs for this model) can be used to justify investment in setup reduction, and the investment boundaries for setup reduction will fall at certain costs of improving setup (i.e., along the horizontal axis of the decision matrix). Now suppose that the cost of improving quality was reduced to the point where the optimal decision is to 'invest fully' and eliminate all defects. The time the server formerly spent reworking defective units is now freed, decreasing system utilization and reducing queue length and WIP costs. In this state the system contains less 'waste' in the form of WIP inventories, so setup reduction will only be justified at lower costs than when quality was not improved. The investment boundaries for setup will shift to lower cost of improvement levels as the cost of improvements in quality decreases. Repeating this argument with the practices reversed will lead to the same conclusion for the shift of quality improvement boundaries. Thus, the 'boundary shift' predicted analytically and numerically in Chapters 3 through 6 is the expected general result for this model.

'Boundary shift' is a useful intermediate result for predicting implications for management decisions. At relatively high costs of improvements, the decision boundaries for the 'naive' and 'informed' cases coincide. As the costs decrease, more improvements will take place in each practice, and result in 'boundary shift' in the 'informed' case. (The 'naive' case assumes improvement practices are independent of each other so determination of a boundary position for one practice will not consider any investments in the other practice and no 'boundary shift' will occur.) The analytic results of Chapters 3 through 5 show that 'boundary shift' always occurs in the same direction, that is, towards positions of lower costs of improvements, while the numerical examples in Chapter 6 also support this finding. Thus, it is logical to conclude the 'informed' case will always call for optimum investment levels of the same or smaller amounts than the 'naive' case. Regions on the decision matrix where the 'informed' (with 'boundary shifts') and 'naive' (without 'boundary shifts') decision boundaries have diverged represent the regions where decision error occurs, i.e., where the optimal investment decision predicted by the 'naive' and 'informed' strategies differ. The net result of 'boundary shift' is that if a project falls in one of these regions, a manager using the 'naive' strategy and ignoring potential interactions will over-invest in improvements compared to a manager using the 'informed' strategy.

The level of over-investment in improvement projects when using the 'naive' strategy can be quite significant. For instance, for the model variant of each Chapter, the levels of over-investments were calculated and frequencies were determined and categorized. They are presented in Table 7.1. Over-investment was calculated by finding the difference between 'naive' and 'informed' case optimum investment levels and normalizing with respect to the 'informed' case level. Levels of over-investment were calculated at discrete points uniformly distributed over the decision matrix for the region where the 'naive' strategy calls for investments in both practices, and frequencies of these values were determined for the categories presented in Table 7.1. Since over-investment is defined as a relative measure, it should be kept in mind that a small level of 'informed' case optimal investment in some instances leads to a dramatic level of over-investment (e.g., in the hundreds of percent), although in absolute terms the over-investment is much more modest. Nevertheless, not only does each variation of the model predict a great potential for over-investment when the 'naive' strategy is used, but the levels of over-investment are quite substantial as well. The most frequent levels of over-investment seen in these examples were between 5 and 50%, or between 50 and 100%.

The potential for over-investment was affected by the form of the model, as well as by parameter levels, but was quite substantial in every example. For instance, in the example frequencies of over-investment of Table 7.1, if over-investment levels above 5% are considered significant, there was better than an 80% potential in each form of the model to over-invest if interactions are ignored.

Table 7.1: Relative Levels of Over-Investment from Model Variation of Each Chapter (from numerical example using parameters of Table 3.2)

Relative Over-Investment (% of 'Informed' Case Optimum)	Linear/Fixed Batch (Chapter 3)	Linear/Variable Batch (Chapter 4)	Convex/Fixed Batch (Chapter 5)	Convex/Variable Batch (Chapter 6)
< 5%	19%	18%	18%	19%
5 - 50%	25%	31%	34%	26%
50 - 100%	38%	16%	30%	28%
100 - 150%	7%	5%	7%	6%
150 - 200%	2%	3%	3%	3%
200 - 250%	1%	3%	2%	2%
250 - 300%	1%	4%	1%	3%
300 -350%	0%	5%	1%	4%
350 - 400%	0%	2%	1%	2%
400 - 450%	1%	1%	0%	1%

Sensitivity analyses were also performed. In general, changes to parameters which increase system utilization (i.e., increasing demand, initial defect rates and setup times or decreasing batch sizes), also increase the potential for making over-investments, along with shifting the frequencies of the levels of over-investments upwards. Holding costs and interest rates had opposite effects on over-investment, with a higher holding cost rate or a lower interest rate leading to increased over-investment frequencies. The two model variants with a convex investment function were very sensitive to changes in the investment function convexity, with greater convexity leading to smaller levels over-investment. The final parameter studied in the sensitivity analyses was unit processing time coefficient of variation, for which no sensitivity in model output was found. This result came as something of a surprise since a stochastic service time model was derived in Chapter 3 specifically to capture the effects of processing time variance. The explanation for this lack of sensitivity has to do with the summing of independent and identically distributed unit processing times acting to reduce the batch service time coefficient of variation for batches made up of large numbers of individual units. While the resulting batch service time variances were less than might have been expected at the outset of this research, this difference is not expected to affect the general conclusions drawn from the model.

In systems where batch size was treated as a variable, there was found to be a very strong linkage between batch size and setup time. In the majority of instances the expected total cost function was non-convex with respect to the batch and setup variables. Consequently, the minimum total cost values generally occurred at extreme points of these two decision variables. For the cases with a linear investment function, or a convex investment function of smaller levels of convexity, if any investment in setup reduction was justified, essentially full investment in setup reduction is optimum, with the resulting optimum batch size being reduced to one unit. In these cases, investment in setup reduction tended to be 'all or nothing', with a corresponding quantum change in optimal batch sizes. Quality and batch size, on the other hand, did not show this extreme behaviour, and continuous changes in variable values were seen across regions of 'some' investment in quality improvement.

Ultimately, the results of this dissertation show that the shape (i.e., the relationship between the level of investment made and the resulting level of improvement) of the investment functions has a serious impact upon the optimum investment decisions in simultaneous improvement projects. As was mentioned in the literature review, very little data was found regarding the form of empirical investment functions. In light of these research results, this lack of knowledge may have a significant unforeseen impact on investment decisions.

7.2 Conclusions

Based on the foregoing summary of this dissertation, the following conclusions have been drawn from this research:

1) The two improvement practices considered in the model of this research, setup time reduction and quality improvement, can interact under a wide variety of circumstances to affect the economic performance of the system. This interaction can lead to over-investments of significant magnitude if not taken into account when investment decisions are made.

2) The form of the investment function strongly affects optimal decisions and the potential for decision error. When using a linear investment function, a 'some-some' investment decision was shown to never be optimal. With a convex investment function, distributions of over-investment frequencies were found to be very sensitive to changes in the investment function convexity parameter. Greater convexity of the investment function led to the 'some-some' optimal decision region on the decision matrix increasing in size, and the potential for over-investment decreasing. Since very little is known empirically about the form of investment functions, decision makers face a significant risk of making incorrect decisions because of the uncertainty surrounding investment functions and their effect on optimal decisions.

3) The risk of this decision error was also affected by system utilization, with greater levels of decision error associated with higher utilizations. Distributions of over-investment frequencies were sensitive to the system parameters which affected utilization, namely, product demand, batch size, initial setup time and quality level, although systems with variable batch sizes were found to be less sensitive to these parameters than systems with fixed batch sizes. As these parameters changed such that utilization increased, the potential for over-investment increases. Thus, a decision maker studying a more fully utilized system faces a greater risk of decision error.

4) No instances were predicted or found in which under-investment occurs, that is, where the consequence of ignoring interactions is investing less than in the case where interactions are allowed for. Ignoring interactions always leads to investments of equal or greater magnitude than when interactions are considered.

5) Optimal batch size and setup times are strongly linked. Setup reduction led to dramatic decreases in optimal batch size, to an optimal batch size of a single unit, while quality improvement leads to more modest decreases. This is in contrast to the results for previous EOQ-based models which, in general, never predicted that single-unit batch sizes were optimal. The fact that these two types of models make such different predictions raises validity issues which must be kept in mind by decision makers.

6) This phenomenon of decision error and risk of significant levels of over-investment was robust to all the variations of the model studied here. It occurred whether batch size was fixed or variable, or whether a linear or convex investment function was used. This result suggests that this type of decision error may be quite wide-spread.

7) Lot-size and investment/improvement models of production systems can be implemented using queuing models such as the M/G/1 queue. Use of queuing models provides a richer modeling environment compared to the EOQ/EPQ models which have been widely used previously since a queuing-based model can capture features of a production system such as process variability and WIP inventories.

7.3 Model Limitations

As with any model, those used in this research made a number of simplifying assumptions which place implicit limits upon how model results may validly be used. These limitations are briefly discussed here.

1) Physical Model

a) Setup Time. Modeling of setup time, that is, assuming each batch entering service requires a setup of fixed length, is consistent with that of most previous researchers mentioned in the literature review. This model does not capture randomness of setup times or differences in setup times between different products, which may be seen in actual production systems.

b) Quality. There is a considerable divergence in the modeling of quality among previous models found in the literature. This suggests a similarly broad treatment of defect handling practices in industry. The assumptions used for quality in the model of this dissertation may not represent accurately a wide cross-section of industrial practice, but is expected to include the basic trend that as quality improves, fewer system resources must be expended.

c) Service Time. The service time model has been discussed in Section 3.6. While it probably does not accurately capture service time variation seen in typical production systems, its use still leads to queuing behaviour so that results will be qualitatively valid.

d) Single Server/Single Product. Only a single server was modeled. This limitation precludes the model from addressing behaviours resulting from interaction with upstream and downstream operations normally found in a production system. Similarly, only a single product was assumed, which permitted arbitrary selection of batch sizes in Chapters 4 and 6.

2) Economic Model

a) Production Costs. Only WIP inventory holding costs were considered as relevant production costs. While WIP cost has been identified by Primrose [1992] as being the most important benefit from setup and quality improvement, it is by no means the only one. Including additional benefits (cost reductions) in a cost model is likely to show greater investments are justified, although the interactive trends found here are unlikely to change.

b) Investment-Improvement Functions. One of the major conclusions of this research is that the form of the investment functions plays a very significant role in model predictions. Obviously, valid model output depends upon the validity of the investment function used.

c) Investment Amortization. This model, just as most of the EOQ-based models found in the literature, used a very simple financial justification model for capital investment. As most firms will use more formal methods to evaluate investment projects, the point at which an investment is deemed acceptable will probably shift when using such methods. Since both setup and quality projects would be subjected to the same evaluation methods, the observations found here for investment decisions are expected to hold.

3) Analytic vs. Numerical Results.

Many of the behaviours found in this research were shown through closed-form solutions of the various forms of the basic model. For instance, 'boundary shift' was shown to be a general result in Chapters 3, 4 and 5 through analytic solutions. Since 'boundary shift' led to the concept of decision error as a result of ignoring potential interactions, conclusions regarding decision error are valid regardless of parameter values (excepting, of course, for trivial values).

Results for the model of Chapter 6, however, had to be obtained through numerical examples due to the mathematical complexity of that model. While results of numerical examples are functions of the particular parameter values used, all results in Chapter 6 showed the same trends as those found in the analytically-solved models of the previous chapters. This outcome does not prove the results seen in Chapter 6 are generic, but it does provide evidence of that.

All numerical examples in this dissertation were based on example parameter values given in Table 3.2. These values were adapted from examples found in the literature (for example, from Porteus [1986b], Spence and Porteus [1987] and Kim et al. [1992]), as well being representative of the discrete products manufacturing systems the author has worked in.

In summary, the ability to make quantitative predictions with this model, for example, determining optimal investment levels in a given manufacturing improvement project, is limited because of the number of simplifying assumptions used. However, the qualitative conclusions drawn, such as the risk of decision error and the importance of investment function form, are expected to be valid and generalizable for many manufacturing systems, while recognizing the model limitations described above.

7.4 Future Research

The model developed in this research has significant potential for being extended and further evolved. These are some possibilities for future research:

1) Empirical Validation. There are seen to be three main opportunities to apply empirical validation to this research, namely, to the investment function form, to the process variance modeling and to the economic model formulation, in addition to an overall validation of the model conclusions. Each of these opportunities is discussed in turn.

a) Investment Function Form. This study found that the form of the investment function had a very significant effect on optimal decisions and potential decision errors. Quite a number of investment function forms have been proposed in the literature (e.g., see Table 2.2), although very little literature was found which provides empirical data for costs of improving these practices. Validating the investment functions is a necessary step in validating the overall model.

b) Process Variance Modeling. As was mentioned in the Discussion, the results of each form of the model used in this study were found to be insensitive to the coefficient of variation of unit processing times. This result was not expected, and in hindsight can be explained by the damping of the variation of service time in large batches of work. Empirical evidence would suggest that processing time variance for batches of work in many production systems is significant, and it appears the processing time model in this study does not adequately capture processing time variance. Developing and validating an improved processing time model based on empirical data is another opportunity to further this research.

c) Economic Model Formulation. The economic (cost) model formulation used here was quite simple. All costs were due to one-time capital investments, amortized over an infinite time period and resulting in immediate improvements to the system. Benefits were due only to reductions in mean WIP inventories, with an assumption of no other benefits such as labour or material savings, increased system capacity or flexibility, etc., due to the improvements made. The extent of these assumptions may mean the model does not adequately represent typical industrial practice. The effect of these economic assumptions on model predictions should be investigated.

2) **Non-Capital Improvements.** To expand upon one of the points in the previous recommendation, there is evidence to suggest that setup times in some situations can be reduced by 50% or more "with very little cost" [Leschke, 1996]. If the assumption is made that these types of non-capital induced improvements occur through a learning process, then use of a learning model, such as those discussed in the Literature Review, combined with a capital investment function may provide an opportunity to explore a richer model of the improvement practice decision process.

3) **Multi-Item/Multi-Stage Models.** This research assumed a single, uniform product was made in a single cell, that total annual demand could be divided into equally-sized batches, that setup times, defect rates and processing times are uniform across all batches and, in two cases, that batch sizes can be arbitrarily set. These assumptions prohibit asking questions such as how improvement efforts should be allocated over various products, or how improvement efforts should be allocated to each cell in a production system. Developing models capable of answering these questions would be expected to lead to significant managerial insights.

4) **Alternative Improvement Practices.** In this study a model was developed which incorporated two improvement practices so that the existence and effects of interactions could be investigated. The two practices used, setup and quality improvement, were chosen because they were found to be the most frequently used practices in industry. However, they are by no means the only improvement practices available, or employed. Determining if other practices exhibit similar behaviour and whether the level or type of decision error changes as more practices are simultaneously implemented would also provide useful managerial insights.

5) Comparisons to Results of Other Models. A significant literature has developed for EOQ-based models with setup and/or quality improvements (e.g., see Table 2.1). These models have been used to make predictions of optimal batch sizes, investment decisions and resource allocation strategies, among other things. The results of these EOQ models may be compared to those of the queuing-based model developed here. For instance, Kim [1990, p85] found with an EPQ-based model that optimum investment levels decrease as system utilization increases, while in a queuing-based model, increasing utilization leads to increasing WIP costs which justify greater levels of investments. Similarly, optimum batch sizes predicted by EOQ-based models are unlikely to be reduced to a single unit, while the models in this study frequently predicted a unit optimal batch. These, and any other, discrepancies between results of each type of model should be investigated.

6) Alternate Quality/Setup Improvement Models. The literature review showed that quality improvements have been modeled in a number of ways, for instance, with assumptions about rework versus scrapping of defectives, that rework is processed by the primary or secondary server, or that the probability of a process going 'out-of-control' is affected by the length of the production run or the care taken in the setup operation. The choice of quality model is expected to have serious implications for the predicted optimal investment decisions. For example, Porteus' [1986b] treatment of quality assumed that after a setup, the process was 'in-control', making good product. With each unit processed, there is a fixed probability of the process going 'out-of-control', and producing defects. Smaller batch sizes will lead to better average quality in such a model. Suppose that quality were modeled in this way with the linear investment function/variable batch size model of Chapter 4. If setup improvement is made, batch size is reduced to one unit, which would always be 'in-control' (i.e., good). All defects would be eliminated without any investment in quality improvement. Obviously, the modeling of quality, and likely setup, play an important role in predicting optimal investment decisions.

The contribution of this research has been to provide a greater understanding of investment decisions for simultaneous improvements to existing manufacturing systems. The significance of the consequences of making investment decisions without considering interactive effects between practices was also demonstrated, as was the value of using information about interactions. Relevant gaps in existing knowledge on this subject have been identified and recommendations were made for future study. In the process, this research has also developed new mathematical models of manufacturing systems for the purpose of studying the effects of making improvements to the systems.

References

Agility Forum, *Next-Generation Manufacturing; A Framework for Action.* Volume I: Summary Report, Next-Generation Manufacturing Project, Bethlehem, PA, (1997).

Bakerjian, R. *Continuous Improvement: Tool and Manufacturing Engineer's Handbook, Volume 7.* Society of Manufacturing Engineers, Dearborn MI, (1993).

Banerjee, A., Pyreddy, V. R. and Kim, S. L. *Investment Policy for Multiple Product Setup Reduction Under Budgetary and Capacity Constraints.* International Journal of Production Economics, 45(1996), pp 321-327.

Barad, M. and Bennett, G. *Optimal Yield Improvement in Multi-Stage Manufacturing Systems.* European Journal of Operational Research, 95(1996), pp 549-565.

Billesbach, T. J. *A Study of the Implementation of Just-In-Time in the United States.* Production and Inventory Management Journal, Third Quarter (1991), pp 1-4.

Billington, P. J. *The Classic Economic Production Quantity Model with Setup Cost as a Function of Capital Expenditure.* Decision Sciences, 18(1987), pp 25-42.

Boucher, T. O. *Lot Sizing in Group Technology Production Systems.* International Journal of Production Research, 22(1984), pp 85-93.

Buzacott, J. A. and Shanthikumar, J. G. *Stochastic Models of Manufacturing Systems.* Prentice-Hall, Englewood Cliffs NJ, (1993).

Chakravarty, A. K and Shtub, A. *New Technology Investments in Multistage Production Systems.* Decisions Sciences, 16(1985), pp 248-264.

Cheng, T. C. E. *An Economic Production Quantity Model with Flexibility and Reliability Considerations.* European Journal of Operations Research, 39(1989), pp 174-179.

Cheng, T. C. E. *EPQ with Process Capability and Quality Assurance Considerations.* Journal of Operations Research Society, 42(1991a), pp 713-720.

Cheng, T. C. E. *An Economic Order Quantity Model with Demand Dependent Unit Production Cost and Imperfect Production Processes.* IIE Transactions, 23(March 1991b), pp 23-28.

Cheng, T. C. E. *An EOQ Model with Learning Effect on Setups.* Production and Inventory Management Journal, First Quarter (1991c), pp 83-84.

Cheng, T. C. E. *An Economic Manufacturing Quantity Model with Learning Effect.* International Journal of Production Economics, 33 (1994), pp 257-264.

Cordon, C. *Quality Defaults and Work-In-Process Inventory.* European Journal of Operational Research, 80(1995), pp 240-251.

Cubberly, W. H. and Bakerjian, R. *Tool and Manufacturing Engineers Handbook: Desk Edition.* Society of Manufacturing Engineers, Dearborn MI, (1989).

Diaby, M. *Optimal Setup Time Reduction for a Single Product with Dynamic Demands.* European Journal of Operational Research, 85(1995), pp 532-540.

Fader, C. *An Economic Analysis Of The Impact Of Just-In-Time On Manufacturing Firms.* Ph.D. Thesis, Department of Management Sciences, University of Waterloo, (1992).

Federgruen, A. and Katalan, Z. *The Impact of Setup Times on the Performance of Multiclass Service and Production Systems.* Operations Research, 44(1996), pp 989-1001.

Fine, C. H. and Porteus, E. L. *Dynamic Process Improvement.* Operations Research, 37(1989), pp 580-591.

Flynn, B. B., Sakakibara, S. and Schroeder, R. G. *Relationship Between JIT and TQM: Practices and Performance.* Academy of Management Journal, 38(1995), pp 1325-1360.

Fraser, N. M., Bernhardt, I. and Jewkes, E. M. *Engineering Economics in Canada.* Prentice Hall Canada, (1997).

Freeland, J. R., Leschke, J. P. and Weiss, E. N. *Guidelines for Setup-Cost Reduction Programs to Achieve Zero Inventory.* Journal of Operations Management, 9(1990), pp 85-100.

Gallego, G. and Moon, I. *The Effect of Externalizing Setups in the Economic Lot Scheduling Problem.* Operations Research, 40(1992), pp 614-619.

Gallego, G. and Moon, I. *Strategic Investment to Reduce Setup Times in the Economic Lot Scheduling Problem.* Naval Research Logistics, 42(1995), pp 773-790.

Gilbert, J. P. *The State of JIT Implementation and Development in the USA.* International Journal of Production Research, v 28 (1990), n 4, pp 1099-1109.

Golhar, D. Y. and Stamm, C. L. *The Just-In-Time Philosophy: A Literature Review.* International Journal of Production Research, v 29 (1991), n 4, pp 657-676.

Goyal, S. K. and Deshmukh, S. G. *A Critique of the Literature on Just-In-Time Manufacturing.* International Journal of Operations and Production Management, v 12 (1992), n 1, pp 18-28.

Goyal, S. K. and Gunasekaran, A. *Effect of Dynamic Process Quality Control on the Economics of Production.* International Journal of Operations and Production Management, 10(1990), pp 69-77.

Hall, R. W. *Zero Inventories.* Dow Jones - Irwin, Homewood, Illinois, (1983).

Harmon, R. L. *Reinventing the Factory : Productivity Breakthroughs in Manufacturing Today.* Free Press, New York, (1990).

Hayes, R. H. and Wheelwright, S. C. *Restoring Our Competitive Edge.* John Wiley & Sons, New York, (1984).

Hilton, P. D. and Gill, G. K. *Achieving Agility: Lessons from the Leaders.* Manufacturing Review, v 7 (1994), n 2, pp 172-179.

Hong, J. D., Hayya, J. C. and Kim, S. L. *JIT Purchasing and Setup Reduction in an Integrated Inventory Model.* International Journal of Production Research, 30(1992), pp 255-266.

Hong, J. D. and Hayya, J. C. *An (s, Q) Policy with Setup Reduction in a Single-Stage Batch Manufacturing System.* Computers & Operations Research, 19(1992), pp 9-19.

Hong, J. D., Xu, S. H. and Hayya, J. C. *Process Quality Improvement and Setup Time Reduction in Dynamic Lot-Sizing.* International Journal of Production Research, 31(1993), pp 2693-2708.

Hong, J. D. *Effects of WIP Inventories on Integrated Production Lot Sizing in Cellular Manufacturing.* International Journal of Production Research, 32(1994), pp 1001-1012.

Hong, J. D. and Hayya, J. C. *Joint Investment in Quality Improvement and Setup Reduction.* Computers in Operations Research, 22(1995), pp 567-574.

Hong, J. D., Kim, S. L. and Hayya, J. C. *Dynamic Setup Reduction in Production Lot Sizing with Nonconstant Deterministic Demand.* European Journal of Operational Research, 90(1996), pp 182-196.

Hwang, H., Kim, D. B. and Kim, Y. D. *Multiproduct Economic Lot Size Models with Investment Costs for Setup Reduction and Quality Improvement.* International Journal of Production Research, 31(1993), pp 691-703.

Inman, R. A. *The Impact of Lot-Size Reduction on Quality.* Production and Inventory Management Journal, First Quarter (1994), pp 38-42.

Karmarkar, U. S. *Lot Sizes, Lead Times and In-Process Inventories.* Management Science, 33(1987), pp 409-418.

Karwan, K. R., Mazzola, J. B. and Morey, R. C. *Production Lot Sizing Under Setup and Worker Learning.* Naval Research Logistics, 35(1988), pp 159-175.

Keller, G and Noori, H. *Justifying New Technology Acquisition Through its Impact on the Cost of Running an Inventory Policy.* IIE Transactions, 20(1988), pp 284-291.

Kim, S. L. *Setup Cost Reduction Models and Synergistic Effects*. Ph.D. Dissertation, Pennsylvania State University, (1990).

Kim, S. L. and Arinze, B. *A Knowledge Based Decision Support System for Setup Reduction*. Decision Sciences, 23(1992), pp 1389-1405.

Kim, S. L., Hayya, J. C. and Hong, J. D. *Setup Reduction in the Economic Production Quantity Model*. Decision Sciences, 23(1992), pp 500-508.

Lambrecht, M. R., Chen, S. and Vandaele, N. J. *A Lot Sizing Model with Queuing Delays: The Issue of Safety Time*. European Journal of Operational Research, 89(1996), pp 269-276.

Lee, W. J., Daesoo, K. and Cabot, A. V. *Optimal Demand Rate, Lot Sizing and Process Reliability Improvement Decisions*. IIE Transactions, 28(1996), pp 941-952.

Leschke, J. P. *An Empirical Study of the Setup-Reduction Process*. Production and Operations Management, 5(1996), pp 121-131.

Liu, J. J. and Yang, P. *Optimal Lot-Sizing in an Imperfect Production System with Homogeneous Reworkable Jobs*. European Journal of Operational Research, 91(1996), pp 517-527.

Lubben, R. T. *Just-in-Time Manufacturing : An Aggressive Manufacturing Strategy*. McGraw-Hill, New York (1988).

Mekler, V. A. *Setup Cost Reduction in the Dynamic Lot-Size Model*. Journal of Operations Management, 11(1993), pp 35-43.

Min, K. J. and Chen, C. K. *A Competitive Inventory Model with Options to Reduce Setup and Inventory Holding Costs*. Computers in Operations Research, 22(1995), pp 503-514.

Monden, Y. *Toyota Production System : An Integrated Approach To Just-in-Time*. 2nd ed., Industrial Engineering and Management Press, Norcross, GA, (1993).

Montgomery, J. C. and Levine, L. O. *The Transition to Agile Manufacturing; Staying Flexible for Competitive Advantage*. ASQC Quality Press, Milwaukee, Wisconsin, (1996).

Moon, I. *Multiproduct Economic Lot Size Models with Investment Costs for Setup Reduction and Quality Improvement: Review and Extensions*. International Journal of Production Research, 32(1994), pp 2795-2801.

Nasri, F., Affisco, J. F. and Paknejad, M. J. *Setup Cost Reduction in an Inventory Model with Finite Range Stochastic Lead Times*. International Journal of Production Research, 28(1990), pp 199-212.

Ocana, C. and Zemel, E. *Learning from Mistakes: A Note on Just-In-Time Systems*. Operations Research, 44(1996), pp 206-214.

Peressini, A. L., Sullivan, F. E., Uhl Jr., J. J. *The Mathematics of Nonlinear Programming*. Springer-Verlag, New York, (1988).

Porter, M. E. *Canada at the Crossroads: The Reality of a New Competitive Environment*. Supply and Services Canada, (1991).

Porteus, E. L. *Investing in Reduced Setups in the EOQ Model*. Management Science, 31(1985), pp 998-1010.

Porteus, E. L. *Investing in New Parameter Values in the Discounted EOQ Model*. Naval Research Logistics Quarterly, 33(1986a), pp 39-48.

Porteus, E. L. *Optimal Lot Sizing, Process Quality Improvement and Setup Cost Reduction*. Operations Research, 34(1986b), pp 137-144.

Primrose, P. L. *Evaluating the Introduction of JIT*. International Journal of Production Economics, 27(1992), pp 9-22.

Rachamadugu, R. and Tan, C. L. *Policies for Lot Sizing with Setup Learning*. International Journal of Production Economics, 48(1997), pp 157-165.

Replogle, S. H. *The Strategic Use of Smaller Lot Sizes through a New EOQ Model.* Production and Inventory Management Journal, Third Quarter (1988), pp 41-44.

Rogers, D. F. *Optimal Strategies for Investment in Setup Cost Reductions in a Just-In-Time Environment.* In Gulledge, T.R. and Litteral, L.A. (eds), "Lecture Notes in Economics and Mathematical Systems", Springer Verlag, New York, (1989), pp 298-310.

Rosenblatt, M. J. and Lee, H. L. *Economic Production Cycles with Imperfect Production Processes.* IIE Transactions, March 1986, pp 48-55.

Sarker, B. R. and Coates, E. R. *Manufacturing Setup Cost Reduction under Variable Lead Times and Finite Opportunities for Investment.* International Journal of Production Economics, 49(1997), pp 237-247.

Schonberger, R. *Japanese Manufacturing Techniques.* Free Press, New York, (1982).

Schonberger, R. *World Class Manufacturing : The Lessons of Simplicity Applied.* Free Press, New York, (1986).

Shingo, S. *The Shingo Production Management System: Improving Process Functions.* Productivity Press, Cambridge, MA, (1992).

Shingo, S. *A Study of the Toyota Production System From an Industrial Engineering Viewpoint.* Productivity Press, Cambridge, MA, (1989).

Soesilo, D. and Min, K. J. *An Inventory Model with Variable Levels of Quality Attributes via Geometric Programming.* International Journal of Systems Science, 27(1996), pp 379-386.

Sohal, A. S., Ramsay, L., Samson, D. *JIT Manufacturing: Industry Analysis and a Methodology for Implementation.* International Journal of Operations and Production Management, 13(1993), pp 22-56.

Spence, A. M. and Porteus, E. L. *Setup Reduction and Increased Effective Capacity.* Management Science, 33(1987), pp 1291-1301.

Sung, C. S. and Lee, W. S. *Setup Cost Reduction in a Dynamic Lot Size Model with Multiple Finite Production Rates.* Engineering Optimization, 24(1995), pp 19-37.

Tielemans, P. F. J. and Kuik, R. *An Exploration of Models that Minimize Leadtime Through Batching of Arrived Orders.* European Journal of Operational Research, 95(1996), pp 374-389.

Tijms, H. C. *Stochastic Models: An Algorithmic Approach.* Wiley, New York, (1994).

Trevino, J., Hurley, B. J. and Friedrich, W. *A Mathematical Model for the Economic Justification of Setup Time Reduction.* International Journal of Production Research, 31(1993), pp 191-202.

Trigeiro, W. W. *The Effect of Setup Time on Production Lot Sizes.* Production and Inventory Management, Third Quarter (1987), pp 50-52.

Warnecke, H. J. *The Fractile Company, A Revolution in Corporate Culture.* Springer Verlag, Berlin, (1993).

White, R. E. *An Empirical Assessment of JIT in U.S. Manufacturers.* Production and Inventory Management Journal, Second Quarter 1993, pp 38-42.

Womack, J. P, Jones, D. T. and Roos, D. *The Machine Which Changed the World.* Harper Collins, New York, (1990).

Yang, J. and Deane, R. H. *Setup Time Reduction and Competitive Advantage in a Closed Manufacturing Cell.* European Journal of Operational Research, 69(1993), pp 413-423.

Zangwill, W. I. *From EOQ Towards ZI.* Management Science, 33(1987), pp 1209-1223.