Investigation of performance-based contract : modeling and metric formulation

Hassan Mirzahosseini

2014

Hassan Mirzahosseini. (2014). Investigation of performance-based contract : modeling and metric formulation. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/61740>

<https://doi.org/10.32657/10356/61740>

**INVESTIGATION OF PERFORMANCE-BASED CONTRACT:
MODELING AND METRIC FORMULATION**

HASSAN MIRZAHOSSEINIAN

School of Mechanical and Aerospace Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2014

Acknowledgements

All adversities I had in my life have only strengthened me but never stopped me. I learnt through them, and they let me grow wise. I thank my God, making me strong enough to fight against my life's obstacles during past 4 years and helping me to complete my Ph.D. programme.

I would like to thank my supervisor, Professor Rajesh Piplani, for his encouragement, guidance and support throughout my research. It was he who brought this interesting topic to my attention and taught me invaluable research skills to work on it.

I greatly appreciate Professor Tongdan Jin for his valuable advices and guidance that facilitate the progress of my research.

I would like to thank Agency for Science, Technology and Research (A*STAR) for offering me a generous scholarship scheme (SINGA) that enabled me to pursue my Ph.D. at an inspiring institution like Nanyang Technological University.

Last, but not the least, I am deeply indebted to my parents who have supported me throughout my life.

Table of Contents

Acknowledgements	I
Summary	VI
List of Tables.....	VIII
List of Figures	IX
Introduction	1
1.1 Emergence of Performance-based Contracts	1
1.2 Performance-based Contracts: Implementations	4
1.3 Performance-based Contracts: Issues	7
1.3.1 Performance	7
1.3.2 Metric	9
1.3.3 Incentives	10
1.4 Research Objectives and Motivation	10
1.5 Organization of the Thesis	16
Literature Review	18
2.1 Repairable Parts Inventory Management.....	19
2.1.1. METRIC extension models.....	19
2.1.2. Relaxing METRIC assumptions	20
2.1.3. Lateral transshipments	22
2.2 Performance-based Contract.....	25
2.2.1 Performance and the metric	26
2.2.2 Incentives scheme	33
2.2.3 Supplier's risk attitude	36
2.3 Summary	39
Repairable parts inventory system.....	41

3.1 The Inventory System	42
3.1.1 Notation.....	43
3.1.2 The Markov model.....	45
3.2 The Metrics and System Performances.....	46
3.3 Numerical Study	49
3.4 Parametric Analysis	51
3.5 Summary	53
Incentive Design in Performance-based Contracts.....	55
4.1 Performance-based Contract	56
4.1.1 Notation.....	57
4.1.2 Customer's problem.....	58
4.1.3 Supplier's problem.....	59
4.2 Probability Distribution of Inventory Level.....	59
4.3 Incentive Mechanism	61
4.4 Supplier's Costs	62
4.4.1 Base stock production cost.....	63
4.4.2 Service capacity cost.....	63
4.4.3 Linear reliability cost function.....	64
4.4.4 Exponential reliability cost function.....	66
4.5 Numerical Example.....	70
4.6 Summary	75
Moral Hazard Problem and Risk Sharing under PBC	77
5.1 Model of Performance-based Contract	79
5.1.1 Notation.....	79
5.2 Repair Process.....	80
5.3 Average and Variance of Performance Outcome	81
5.4 Risk-averse Supplier and Convex Compensation Model	82
5.5 Supplier's Costs	84
5.5.1 Subsystem design cost	85
5.5.2 Production cost.....	85
5.5.3 Holding cost	85

Table of Contents

5.5.4 Repair activities cost	86
5.5.5 Service capacity cost.....	86
5.6 Reimbursement Policy	87
5.7 Supplier's risk attitude	90
5.8 Supplier's problem.....	91
5.9 Numerical example	92
5.10 Sensitivity analysis.....	98
5.11 Summary	99
Conclusions and Future Research.....	100
6.1 Summary	101
6.2 Future Work	104
6.2.1 Multiple-Phase and Multi-Level Performance Contracting.....	104
6.2.2 Performance Incentives Trade-off	105
6.2.3 Performance Control Measures.....	106
Appendix A.....	107
Appendix B.....	114
Appendix C.....	118
Appendix D.....	124
References	127

Summary

Performance-based Contract (PBC) has developed into a new service strategy in after-market operations that aims to compensate for performance outcomes. While the traditional contracts for after-sales services, such as Fixed price (FP) and Cost plus (C+), only provide insurance or incentive, PBC, by considering the system in a decentralized framework, is designed to serve both the customer and supplier's requirements, simultaneously. In this research, an enhanced inventory system for repairable parts operating under a PBC is modeled. The model provides the supplier and the customer increased flexibility in achieving the target performance by relaxing some assumptions of the previous PBC models and investigating the most frequent enablers (performance, metric and incentive) of successful implementation of PBC.

Under PBC, system availability (a widely used performance measure) is affected by multiple factors such as reliability, spares, fleet size, and service response time. Therefore, this research focuses on reliability improvement and its interaction with the decisions affecting service time, taking into account the fleet size. A principal-agent model is developed to evaluate the impact of the operating fleet size on the incentive mechanism. The results support the hypothesis that a win-win incentive program should take into account the changes in operating fleet size. This finding provides a key insight into many real-world applications where the fleet size may change over the contract horizon due to product introduction, market saturation, or phase-out. The numerical example confirms that the change of the operating fleet size significantly influences the supplier's service decision and the profit margin. Hence, the customer and the supplier both should pro-actively adjust the contract terms to avoid undesirable consequences.

Moral hazard under PBC arises when one of the parties in the contract performs an action that is unobservable to the other party. The customer should design an appropriate contract to counteract the problem of moral hazard because of invisibility of the supplier's response to the contract terms. This study attempts to design an incentive contract based on the supplier's risk attitude. An analytical model is presented to prove the efficacy of considering a convex component in the optimal revenue function under PBC. Then, a numerical analysis is conducted to examine common utility functions (exponential, power, mean-variance), with the objective of finding the optimal revenue function that maximizes the outcome performance under PBC. The findings indicate that a linear incentive contract performs poorly when the supplier's utility belongs to a large class function. On the other hand, the customer should not propose an exponential compensation model if the supplier has a mean-variance risk attitude.

List of Tables

TABLE 1.1 EXAMPLES OF IMPROVEMENT ACHIEVED WITH PBC (VITASEK ET AL. 2006).....	5
TABLE 1.2 PERFORMANCE / METRICS	11
TABLE 3.1 BASIC PARAMETERS OF THE SYSTEM.....	50
TABLE 3.2 SYSTEM PERFORMANCE AND THE METRICS	50
TABLE 3.3 PARAMETER RANGE	51
TABLE 4.1 Φ_1 AND Φ_2 WITH INCREASING NUMBER OF SYSTEMS.....	65
TABLE 4.2. PARAMETER OF HVLC COMPONENTS AND CONTRACT TERMS	71
TABLE 4.3. PARAMETER OF LVHC COMPONENTS AND CONTRACT TERMS	71
TABLE 4.4. COEFFICIENT OF DETERMINATION (R^2)	73
TABLE 5.1 COMPONENT COST PARAMETERS	92
TABLE 5.2 VALUES OF FACTORS	92
TABLE 5.3 THE SUPPLIER'S RESPONSE (CASE 1)	95
TABLE 5.4 THE SUPPLIER'S RESPONSE (CASE 2)	95
TABLE 5.5 SENSITIVITY OF THE OPTIMAL AVAILABILITY TO THE INVENTORY HOLDING COST ..	98
TABLE C.1 SENSITIVITY OF THE OPTIMAL AVAILABILITY TO THE AVERAGE REPAIR COST	122
TABLE C.2 SENSITIVITY OF THE OPTIMAL AVAILABILITY TO K_T	122
TABLE C.3 SENSITIVITY OF THE OPTIMAL AVAILABILITY TO K_p	123
TABLE C.4 SENSITIVITY OF THE OPTIMAL AVAILABILITY TO THE SERVICE CAPACITY COST ...	123
TABLE C.5 SENSITIVITY OF THE OPTIMAL AVAILABILITY TO	123
TABLE D.2 IMPACT OF SERVICE TIME DISTRIBUTION ON AVAILABILITY	125
TABLE D.3 IMPACT OF COMPONENT'S FAILURE DISTRIBUTION ON AVAILABILITY	126

List of Figures

FIG 1.1 SUPPLIER’S MOTIVATION AND CUSTOMER’S ASSURANCE UNDER CONTRACTS.....	4
FIG 1.2 PBC RELIABILITY PLANNING PROCESS FLOW (SMITH 2006).....	9
FIG 2.1 DEPOT REPAIR (GUPTA AND ALBRIGHT 1992)	21
FIG 2.2 CLOSED-LOOP CYCLE FOR REPAIRABLE ITEMS (KIM ET AL. 2007).....	31
FIG 2.3 THE CONCEPT OF BONUS, DEAD AND PENALTY ZONES (SOLS ET AL. 2013).	34
FIG 3.1CLOSED-LOOP INVENTORY SYSTEM	43
FIG 3.2THE TRANSITION DIAGRAM OF THE INVENTORY SYSTEM.	45
FIG 3.3 IMPACT OF SYSTEM PARAMETERS ON AVAILABILITY.....	52
FIG 4.1THE TRANSITION DIAGRAM OF THE INVENTORY SYSTEM WITH CONSTANT SERVICE RATE...	60
FIG 4.2 THE EXPONENTIAL BEHAVIOR OF RELIABILITY COST FUNCTION.....	67
FIG 4.3. FITTING AN EXPONENTIAL RELIABILITY COST FUNCTION	72
FIG 4.4 SUPPLIER’S OPTIMAL RELIABILITY DECISION IN RESPONSE TO THE CONTRACT TERMS	74
FIG 4.5. MINIMUM INCENTIVE RATIO VS. FLEET SIZE	75
FIG 5.1 LINEAR COMPENSATION FUNCTION.....	88
FIG 5.2. EXPONENTIAL COMPENSATION FUNCTION.....	89
FIG 5.3. PIECEWISE-LINEAR-THRESHOLD COMPENSATION FUNCTION	90
FIG 5.4 LINEAR, EXPONENTIAL AND STEP REVENUE FUNCTION FOR THE FIRST SCENARIO	93
FIG 5.5 LINEAR, EXPONENTIAL AND STEP REVENUE FUNCTION FOR THE SECOND SCENARIO	94
FIG 5.6THE IMPACT OF THE DESIGNED REVENUE FUNCTION ON THE AVAILABILITY	97
FIG A.1 IMPACT OF SYSTEM PARAMETERS ON BACKORDERS	109
FIG A.2 IMPACT OF SYSTEM PARAMETERS ON WAITING TIME	110
FIG A.3 IMPACT OF SYSTEM PARAMETERS ON FILL RATE	111
FIG A.4 IMPACT OF SYSTEM PARAMETERS ON READY RATE.....	112
FIG B.1 THE SENSITIVITY OF THE NUMBER OF BACKORDERS TO THE BASE STOCK.....	115
FIG B.2 THE VALUE OF $(A/A-B)$ BASED ON Φ AND S CHANGES.	116
FIG B.3THE VALUE OF $(A/A-B)$ BASED ON Φ AND S CHANGES.....	117
FIG C.1 GLOBALSEARCH STEPS	119
FIG C.2 BASINS OF ATTRACTION	121
FIG D.1 SIMULATION MODEL USING ARENA SOFTWARE.....	124

Chapter 1

Introduction

1.1 Emergence of Performance-based Contracts

It has been estimated that over the life cycle of most systems about 30 percent of all dollars are spent in acquiring the system, while the remaining 70 percent are spent on support (Berkowitz et al. 2005). The after-sales parts and services have become a key source of revenue for firms selling capital-intensive and long-life goods. An Accenture 2003 study of General Motors indicated that \$9 billion in after-sale revenue added \$2 billion to profits whereas profits from the company's \$150 billion in car sales were relatively lower (Dennis and Kambil, 2003). As further evidence, in 2001, IBM derived over \$5 billion of its revenue from maintenance and services (Cohen et al. 2006).

Considering the importance of after-sales service, many studies have been conducted to introduce the most beneficial contract between the supplier (agent) and the customer (principal), as the two main parties involved in after-sales service arrangement. Designing contracts in capital-intensive industries such as aerospace and defense is more complex in comparison to industries such as electronics, as the value of component or product is an integral criterion in specifying the terms of contract and relationships between the principal and the agent. For example a supplier never guarantees a fighter's engine for a fixed payment because of high price and repair cost of the engine.

The traditional after-sales contracts used by the capital goods manufacturers are defined next.

Fixed-price: Under fixed price contracts the supplier accepts to satisfy a required service level in return for a certain payment (Garrett, 2001). Because the parties indicate a specified

service under fixed price type contracts, supplier may suffer losses at the end of period as the risk is fully shifted to the supplier. On the other hand, fixed price type contracts increase the supplier's incentive as the supplier can benefit by reducing the costs.

Cost-plus: Under Cost-plus contracts, the customer reimburses the supplier's expenses based on an estimation of project cost and pays a specified price as profit (Garrett, 2001). Under cost plus contracts the supplier is compensated for permissible costs and a fixed fee is paid as profit by the customer. In this type of contract, there is no motivation for the supplier to control the costs but the risk is now completely shifted to the customer.

The profitability of support services, especially for systems with long operational life cycles, encourages both private and public sectors to pay attention to supportability which is a core performance measure of the logistics system. Accordingly, performance based supportability has become an established sustainment strategy that aims for the best industry practices implementation in a cost-effective manner (Rogers, 1997). Based on the fact that support services are outsourced the sustainment contract is reshaping into a novel approach termed performance based contract (PBC) that defines an agreement between the two contracting parties (i.e., customer and supplier). The contract between RENFE (the Spanish railway state company) and CAF (the manufacturer of the original TRD trains) was a clear case of PBC implementation put in place to manage REFE's reliability for 14 years (Sols et al. 2013).

The U.S. Department of Defense (DoD) has adopted Continuous Process Improvement (CPI) to improve the performance of its systems measured by readiness, reliability and process cycle times (Nowicki, 2008). In DoD, continuous process improvement consists of the application of a wide range of methods, such as Lean, Six Sigma, and Theory of Constraints (TOC). In accordance with CPI Transformation

Guidebook, one of the key elements of CPI is to concentrate on structure of the arrangement which facilitates meeting the system's outcome (DoD March, 2006).

The DOD mandated implementation of the PBC concept and improved business systems, including metrics, which eliminate non-value-added steps, and improve the total readiness of major weapons over life-time. As Berkowitz et al. (2005) note, "PBC is an integrated acquisition and sustainment strategy for enhancing weapon system capability and readiness where the contractual mechanisms include long-term relationships and appropriately structured incentives with service providers, both organic and non-organic, to support the end user's objectives".

While the traditional after-sales contracts in capital-intensive industries, such as Fixed-price and Cost-plus, are based on buying spare parts, repair action, equipment, transactional goods or services, the PBCs focus on buying the outcome performance. For example, under PBC, maintenance and servicing of the engine is not paid according to the cost of the spare parts used, or effort expended in repairs, but how many hours the customer obtains power from the engine (Ng et al. 2009). Under PBC, the supplier has more flexibility and more opportunity for creativity and innovation as the customer only specifies the requirements, not the means of achieving them. The customer, on the other hand, can obtain the best performance by reimbursing the supplier based on the real outcome. The PBC enhances the supplier's ability by shifting to him the responsibility to control the service elements, such as packaging, handling, storage, transportation and warehousing, required to successfully support the system.

Sols et al. (2013) compare the differences in the customer's assurance relative to achievement of effectiveness goals and the supplier's motivation to meet the objective.

Figure (1.1) demonstrates how PBC is superior to traditional contracts (i.e., Fixed-price and Cost-plus) for both the customer and the supplier.

Fig 1.1 Supplier's motivation and customer's assurance under contracts

As it can be seen, under PBC the customer's assurance and supplier's motivation are high as the payment is tied to realized outcome.

1.2 Performance-based Contracts: Implementations

Defense sector: Investigation of PBC implementation in the defense sector was launched in 1998 where a team of 60, including the office of the Secretary of Defense and Logistics Agency, evaluated the shifting process from traditional contracts to PBC. They set a target of 50% migration to PBCs by the end of 2005 in Army, Air and Navy U.S force (Rievley, 2001).

One of the earliest PBC implementation dates back to 1998 when Lockheed Martin offered a system to DoD for supporting F-117 Nighthawk, which tied its compensation to the fighter's performance outcome. While there were about 100 acquisition programs supported

by PBC by October 2005 (Phillips, 2005), this number increased to more than 200 PBCs in place across the DoD by July 2006 (Geary, 2006). Table (1.1) illustrates some examples of improved performance under PBC implementation.

Table 1.1 Examples of improvement achieved with PBC (Vitasek et al. 2006)

Program	Availability		Response Time	
	Pre-PBC	Post-PBC	Pre-PBC	Post-PBC
F-14 Lantern	73%	90%	56.9 days	5 days
ARC-210			22.8 days	5 days
H-60 Avionics	71%	85%	52.7 days	8 days
F/A-18 SMS	65%	98%	42.6 days	2 days
Tires	70%	85%	28.9 days	2 days
APU	65%	90%	35 days	6.5 days
F/A-18E/F	73%	85%		

Health care system: PBC is becoming a dominant logistics support strategy. Application of PBC is not limited to the commercial and defense industries but the health care system is benefiting from PBC implementation as well. Los Angeles Times (January 26, 1999) reported that the total reimbursement of employees' prescription drugs has been increased by 13.8% within one year (1997-1998). Due to rising drug costs, both governments and health service providers encounter serious challenges. To overcome these problems, Health Maintenance Organizations (HMO) and Governments began to put some restrictive policies on their coverage and drug reimbursement that reduce costs but adversely influence the pharmaceutical companies' profit. For example, "Amgen's stock price dropped by 12% on August 12, 1997 when the Federal Health Care Finance Administration (FHCFA) proposed changing the reimbursement policy of Epogen, a drug for boosting red blood cell counts for patients with chronic kidney disease"(So and Tang, 2000).

The conflicting consequences of such policies motivate both HMOs and Medicare to shift toward a new service strategy that benefits all parties (i.e., patients, hospital, clinic and pharmaceutical). Performance contracts as a new service strategy in health care systems are replacing traditional restrictive policies to control drug costs. PBC aims to develop reimbursement policies of the health care providers that reduce costs and improve quality of the service, simultaneously.

Public transport system: The first application of incentive-based performance contracts in public transport operators back to early 2000 where an innovative payment mechanism was used for three bus operators in the Hordaland County. The Hordaland payment mechanism has a two-stage process where the first stage aims to maximize a social surplus function by specifying bus revenue-km, fare levels and bus capacities. The second stage tries to maximize the operator's profit by computing rates for revenue-km and fare subsidies. In this approach, the subsidy is arranged to balance the avoided external costs of running empty vehicles and the income of increased service frequency (Fearnley et al. 2004).

Solving the intercity traffic problem in Norwegian rail transport is an example of PBC successful implementation in industry sector which was launched in 2003 between the Ministry of Transport and Communication (MOT) and the Norwegian State Railways (NBS). They designed an incentive mechanism based on demand and production to reach the social goals of the Ministry of Transport with maximizing the Norwegian State Rail's profit. The payment was made in accordance with the train kilometer and seat kilometer production leading to increased service frequency and reduced crowding. To capture the risk of ignoring punctuality, they also penalize NBS for train cancellations and delays.

1.3 Performance-based Contracts: Issues

PBC asks the supplier “what to achieve” based on its own knowledge, instead of “what to do” according to the customer experiences. Devries (2004) studied the relationship between successful implementation of PBC and the barriers and enablers to its implementation. A survey was done to collect data from 26 PBC managers from different industries. Findings indicated that performance, metrics and incentives were the most frequent enablers of PBCs and enhancing them can lead to their successful implementation.

1.3.1 Performance

Performance goal is one of key items of a PBC that should be based around a well understood metric. The transparency between the customer and supplier is essential to figuring out an appropriate performance metric. An example of appropriate performance definition can be found in PBC made between a Spanish aircraft manufacturer (EADS CASA) and South American countries (Chile, Paraguay, Ecuador and Brazil) to support CN-235 and C-212 aircrafts (Sols et al. 2013). Both parties agreed to choose the fleet availability that is monitored daily by both the customer and the supplier. They also considered the following four states of the aircraft which helped them to reach a common definition of availability during inspection:

Operational: The aircraft is ready to fly;

Scheduled maintenance: The aircraft is on the ground but the situation is acceptable as it is a planned maintenance.

Down: The aircraft is unavailable due to component failure and is awaiting maintenance;

Maintenance: The aircraft is not available due to a failure by the user, but is undergoing maintenance.

By employing a PBC the customer acquires an outcome rather than a specific service. Program manager (PM) specifies the system performance based on the objectives of logistics support and documents it in a Performance Based Agreement (PBA). The PBA, which is the foundation for the PBC program, is used to conduct the Business Case Analysis (BCA) and is the basis for the Request for Proposal (RFP) (Landreth et al. 2005). Defense Acquisition Guidebook defines the performance, as a key goal of PBC, in terms of the system objectives as follows:

1. Operational Availability: The percent of time a weapon system is available for a mission or is able to sustain operations tempo.
2. Operational Reliability: The measure of a weapon system in meeting a mission's success objectives (percent of objectives met, by weapon system). Depending on the weapon system, a mission objective could be a sortie, tour, launch, destination reached, or capability.
3. Cost per Unit Usage: The total operating costs divided by the appropriate unit of measurement for a given weapon system. Depending on the weapon system, the measurement unit could be flight hour, steaming hour, launch, miles driven, etc.
4. Logistics Footprint: The government/contractor size or "presence" of logistics support required to deploy, sustain, and move a weapon system. Measurable elements include inventory/equipment, personnel, facilities, transportation assets, and real estate.
5. Logistics Response Time: This is the period from the time the logistics demand signal is sent till the satisfaction of the demand. "Logistics Demand" refers to systems, components, or resources, including labor, required for weapon system logistics support (DoD, 2006).

Reliability performance is considered as important as availability by the customer under PBC. As the duration of contract may be rather long, reliability growth planning under performance contract plays an important role to prevent system wear-out. Smith (2006) demonstrates reliability planning process flow under PBC in Figure (1.2). He also highlights reliability improvement through design phase, critical component restoration and conditional maintenance (proactive change of marginal components before failure) as three primary activities to improve reliability under PBC.

Fig 1.2 PBC Reliability Planning Process Flow (Smith 2006)

1.3.2 Metric

The term “metric” defines performance measurement, including the method of calculation, the person who would carry out the calculation, and the source from where the data would be

collected (Neely et al. 1995). Metric is at the heart of PBC, where the performance is linked to the metrics specified in the contract. Therefore, the metric should be formulated properly to support the desired performance. Using an inappropriate metric can make the PBC implementation ineffective and increase the program costs.

Maskell (1991) summarizes the metric principals as: 1) being directly linked to defined performance; 2) simple and easy to use; 3) providing fast feedback; and 4) being able to stimulate continuous improvement. Table (1.2) illustrates the applicable metrics for applying the specified system performance, which have been indicated by Naval Air Systems Command Instruction (NAVAIRINST) 4081.2A (2004b).

1.3.3 Incentives

Incentives are to incentivize the supplier to satisfy the customer performance requirements. Incentive can be in monetary, non-monetary, positive or negative. As per Office of the Secretary of Defense(2001), “Regardless of the final composition and structure of the contract, the goal is to encourage and motivate the contractor’s optimal performance for logistical support through reduced cycle time, reduced inventory, improved reliability/obsolescence management, and reduced total ownership costs”.

1.4 Research Objectives and Motivation

The PBC strategy is often named “performance based logistics” (PBL) in defense sector, or called as “power by the hour” (PBH) in commercial segment. PBL is becoming a dominant logistics support strategy, especially in capital-intensive industries where the systems and subsystems require high availability and are expensive to maintain. Considering the high cost and long lifetime of the subsystems in such industries, the failed subsystem is repaired rather

than discarded. Therefore, a repairable parts inventory approach is required to support such systems.

Table 1.2 Performance / Metrics

Performance Outcomes	Metrics/Considerations
Operational Availability	Readiness Mission Capable Rates Sortie Generation Rate Turn- Around-Times Surge Requirements Reduced Down Time
Operational Reliability	Sortie/Mission Completions Time On Wing Mean Time Between Failures (MTBF) MTBF Improvement No Fault Found/Reduction Elimination
Cost Per Unit Usage	Cost Per Flight Hour Annual FFP Cost (prorated by units) Obsolescence Management Attrition Replacement Sustaining Engineering/ECP Costs Total Ownership Cost (TOC)
Logistics Footprint	Maintenance Planning Reliability improvement Reduced Cannibalizations Support Equipment Training Publications Inventory Needs Staffing Levels
Logistics Response Time	Parts Availability First Pass Effectiveness Maintainability P,H,S&T Mean Logistics Down Time Supply Chain Management

Many models have been developed in repairable parts inventory domain over the last 40+ years. METRIC, MOD-METRIC and VARI-METRIC models (Muckstadt, 1973) are the fundamental models in this area. These models started out with some restrictive assumptions that were later relaxed by some researchers. While most of the earlier repairable parts

inventory models try to minimize the total cost subject to satisfying a specified service level into a centralized approach, PBC considers the system in a decentralized framework and tries to satisfy both the customer and the supplier's requirements, simultaneously.

Recently some studies (Kang et al. 2006; Kim et al. 2007; Kim et al. 2011; Nowicki et al. 2008; Oner et al. 2010) have focused on repairable parts inventory models under performance contract, but the models still have restrictive assumptions which reduce the supplier's flexibility and may give indication of unrealistic performance to the customers. Firstly, they ignore the dependency of subsystems failure rate to the number of operational units in the closed-loop system. Secondly, they do not consider the impact the supplier effort may have on component reliability and efficiency of the repair facility, while the PBC should empower the provider with the authorization and responsibility to control those elements required to successfully support the program.

The barriers and enablers for PBC successful implementation have been studied by many researchers (Cunic 2003; Devries 2004; Doerr et al. 2004; Berkowitz et al. 2005). Devries (2004) research, as a dominant study in this area, indicates that performance, metrics and incentives are the most frequent enablers of PBC, and that enhancing them can lead to its successful implementation. While the performance of a system under PBC can be defined in many ways, Sherbrooke's (1968) findings introduce the backorder as the most reasonable performance measure as compared to fill rate, ready rate, and operational rate.

The availability of a repairable part inventory system can be improved by managing three parameters: (1) base stock at the warehouse, (2) component reliability, and (3) efficiency of the repair facility. Thus, the following question becomes important in the successful implementation of a PB contract: How do inventory management, component

reliability and repair facility efficiency influence the availability of systems under PBC? The first research objective is defined to answer this question.

Research Objective 1: Study the repairable part inventory system with performance contract under real-life settings and relax restrictive assumptions, such as fixed failure rate, and study the effect of component failure rate, server repair rate, number of servers and the base stock level on system availability.

The supplier may be unwilling to reveal his efforts level and achievements during the contract period. Also, the wide of variety of contracts available and the complexity of these contracts is major hindrance in deciding the contract type and content terms. These issues make it difficult to develop a universal service contract guideline.

The difficulty in sharing information is another problem in establishing a proper coordination between customer and supplier. Information sharing is difficult due to the fears of information leakage and conflicts of interest between customer and supplier. As a result, it is hard to build an information sharing platform for information sharing and delivery coordination.

Establishing a right mechanism to allocate the incentives of coordination among the supplier and customer, is another challenge. It is difficult to find a balance point and even harder to justify it, especially when information sharing is partial and asymmetric.

PBC realizes expected performance, and payment is linked to successful delivery of outcome. PBC also uses motivation mechanism to ensure that required performance is received. To achieve PBC's ultimate goal, an appropriate incentive should to be in place to motivate the supplier. In previous studies (Kim et al. 2007; Kim et al. 2010), the average

number of backorders, the average downtime and the cumulative downtime have been used as three efficient incentives under performance contracts. Actually, designing a contract based on these three performance measures may cause different responses from the supplier depending on the system features. Also, anticipating the supplier's response to each incentive mechanism is very important in performance contract where the customer cannot directly specify the supplier's investment decision.

Mirzahosseini and Piplani (2011) show that to achieve a desired performance level in a repairable parts inventory system, the supplier has to improve the component reliability and the repair time, rather than invest in building up a stock of spares. The most effective incentive is defined based on the sensitivity of parameters to the performance of the system. This means that the best incentive has to have the most sensitivity to the failure rate and repair rate as it can motivate the supplier to improve those parameters.

The failure rate in a repairable parts inventory system is a function of the number of operational systems and the service rate. The number of operational systems, on the other hand, is directly dependent on the total number of systems which are supported by the supplier. Therefore, the number of systems is one of the main parameters that affect the contract terms and the supplier's response under a performance contract. Against this backdrop, the following question becomes important in the successful implementation of a performance contract: How does the number of supported systems influence the performance incentives and the supplier's decision under a performance contract? Thus, the second research objective is defined as follows:

Research Objective 2: Evaluate the role of the total number of systems on effectiveness of the penalty and cost sharing rates as negative and positive incentives and the supplier's response under a PBC.

One way to coordinate between parties is to enforce an incentive contract according to the principle-agent model. In principle-agent model, the agent makes an effort, which incurs cost for the agent, to meet the principle's desired performance. The principle as the other party in the contract will compensate the agent based on achieved performance level. The moral hazard problem and conflict of interests are two major issues under a principle-agent model that can be solved by designing an appropriate incentive mechanism. The moral hazard problem happens because of information asymmetry where the principle cannot observe the agent's efforts level during the contract time. The second issue comes when both parties (principle and agent) look for his own profit maximization.

The classical agency model considers the tradeoff between incentive and insurance, wherein the principal and agent sign a contract and the principal compensates the agent for receiving the outcome as specified in the contract. Gibbons (2005) reviews the incentives between firms in the subject of "agency theory" (i.e. the economic theory of incentives). His research arrives at three key insights. First, "you get what you pay for" problems with objective performance measure are as important as the tradeoff between incentives and insurance in the agency theory. Second, the incentive contracts are not the only way to implement incentives. For example, the market competition can lead to creation of incentives without the need for contracts. Finally, a powerful technique in the incentive contracts is integration decision which can be provided by applying asset ownership.

PBC has emerged as a new service strategy in capital-intensive industries which aim to compensate for performance outcomes. In performance contracts, availability is widely defined as one of the key performance criteria that can be affected by various factors such as reliability and response time. To achieve a desired availability level, the supplier has to improve the component reliability and the repair time.

Kim et al. (2007) consider a linear compensation model for application in performance contracts. They assume that both the principal and the agent have expected mean-variance utility functions. Chen and Miller (2009) work on principal-agent models where the agent chooses actions over time and is reimbursed based on the final value of output. They use numerical simulation to find the difference between the linear contracts outcome performance and the piecewise-linear-threshold contract's relative performance where the agent is risk-averse. The literature of capturing risk under performance contracts reveals a lack of understanding of how incentive contracts should be designed based on the supplier's risk attitude. The third research objective defined in the present research focuses on this area.

Research Objective 3: Examine the impact of the supplier's utility function on the reimbursement policy under PBC.

1.5 Organization of the Thesis

This thesis is organized into six chapters. Chapter 2 discusses the relevant literature in repairable parts inventory management and PBC. First the methods in previous researches for relaxing the assumptions in repairable parts inventory systems under PBC are studied. Then, the literature is reviewed with the objective of pinpointing integral issues of PBC and

understand the most common performance, metric and incentives which are used in PBC, and finding their gaps.

In Chapter3 the structure of the inventory system is presented and a single-location spare-parts inventory system is modeled under a PBC. First, the closed-loop inventory system is discussed, and some assumptions relaxed by applying the Markovian approach. Then, the system performance and the metric for calculating the performance are defined. Next, an example based on real data from defense sector is used to illustrate the findings. Finally, a parametric analysis is conducted to find the factors which have significant effect on the system performance.

Chapter 4 focuses on the effect of operating fleet size on incentive design under PBC. First, the contract between customer and supplier and the objective function and constraints are described. Then, the contract model is optimized to specify the reliability level and its interaction with service time decisions under varying contract terms. At the end, a numerical analysis is illustrated using real data from semiconductor industry to describe how the minimum acceptable incentive should be adjusted according to the number of systems changes.

In Chapter 5 the incentive contract is designed based on the supplier's risk attitude. In the first section, different forms of revenue functions are defined in the presence of various types of utility functions which the supplier may use to capture the risk. Then, an illustrative experiment using healthcare systems data is conducted to find the optimal reimbursement policy with respect to the supplier's risk attitude.

Chapter 6 provides a summary of findings and suggestions for future work.

Chapter 2

Literature Review

Many researchers have studied the PBC program implementations to find their barriers and enablers, appropriate frameworks and structure for their successful implementation (Cunic 2003; Devries 2004; Doerr et al. 2004; Berkowitz et al. 2005). PBC is a preferred approach in both private and public sectors, especially in capital-intensive industries where the systems and subsystems require high availability and are expensive to maintain. In such industries, the failed parts need to be repaired and cannot be scrapped because of their high cost and long lifetime. Therefore, a repairable parts inventory approach is required to support such systems where their repair process is closed-loop and the number of total subsystems is constant.

Recently some of the researchers have applied repairable parts inventory approach to PBC and optimized the elements of the contract (Kim et al. 2007; Kim et al. 2011; Nowicki et al. 2008; Oner et al. 2010). To simplify the analysis they make some restrictive assumptions, such as fixed failure rate, fixed repair rate and ample repair capacity. Relaxing these assumptions, thus, offers an opportunity to develop a more robust model and influence the contract elements, especially metric formulation and compensation model. Also, these assumptions are in conflict with the PBC principle, which states that “the PBC should empower the provider with the authorization and responsibility to control those elements required to successfully support the program” (NAVAIRINST 4081.2A 2004b).

In this chapter first a step is taken towards relaxing the above-mentioned assumptions in repairable parts inventory systems under PBC. Secondly, the literature is reviewed with the objective of pinpointing integral issues of PBC and understanding the most common

performance, metrics and incentives which are used in PBC implementations, and finding the research gaps this research could fill.

2.1 Repairable Parts Inventory Management

Over the last 40+ years researchers have developed many models to analyze repairable parts inventory systems. METRIC, the first practical model in this area is a mathematical model of a base-depot supply system with compound Poisson demand, which specifies base and depot stock levels to optimize system performance (Sherbrooke 1968). In the case of a failure at base, the unit is repaired at the base, or is sent to a depot in a specific fraction and the failed unit is replaced by a ready to use one. When the failed unit is sent to the depot, a replenishment order is immediately triggered from the base to the depot; lateral resupply between bases is not allowed in this model. METRIC's objective is to minimize the expected number of backorders, calculated by taking the expected value of the number of backordered days during a fixed period of time. Sherbrooke (1968) uses one-for-one replenishment inventory policy and models the system based on the mean repair time rather than its distribution.

2.1.1. METRIC extension models

Muckstadt (1973) extends the METRIC model, calling it MOD-METRIC, to control a multi-item, multi-echelon, multi-indenture inventory system for repairable items. The system consists of several bases and one central depot to support these bases. An item may be demanded at any base and the base replaces the item from the central depot. To improve maintainability, the subsystems (e.g. engine) are considered modular in design containing identical modules. When a subsystem fails, it is replaced by a ready to use one and the failed module (containing the defectives parts) is repaired either at the base or at the central depot.

In the latter case, a demand is issued to the depot to replace the module. Applying this procedure decreases the repair time and increases the serviceable time of the subsystem. The model is used to compute the base stock levels of different echelons by considering this logistical relationship between line replacement unit (e.g. engine) and shop replaceable units (e.g. engine's modules). MOD-METRIC model has been implemented by Air Force to support the F-15 weapon system.

Graves (1985) considers a two-echelon inventory system for a repairable item which includes a repair depot and several operating sites. The failure process is in accordance with a compound Poisson which only depends on the required number of working items, not the actual number of working items. The steady-state distribution of net inventory levels obtained is very close to the exact distribution and dominates the METRIC approximation. Sherbrooke (1986) later develops a VARI-METRIC model which focused on the expected backorders. He changes the METRIC model from first-order to second-order by incorporating variances of repair items. Simulation shows that the VARI-METRIC provides a more accurate estimate of the expected backorders. These four models are the fundamental models of repairable spare parts inventory systems. Many researchers have later enhanced these models by relaxing one or more of their assumptions.

2.1.2. Relaxing METRIC assumptions

An infinite repair capacity is one of restrictive assumptions researchers have tried to relax. Gross et al. (1985) study the benefits which can be gained by considering the interaction among reliability, queuing and inventory in a multi-echelon repairable-item inventory system using closed-network queue. They consider a limited number of simultaneous repair channels and a revised failure process, which is dependent on the number of operating units, to determine the optimal combination of the number of spares and repair channels to deliver the

service-level performance. Their parametric analysis shows the influence of failure rate and repair rate on steady-state availability.

Gupta and Albright (1992) model a two-echelon multi-indentured repairable item inventory system with ' N ' bases where each base has a maximum number of online channels which are used by the machines (Figure 2.1). When the number of operating machines in a base is less than the number of channels, some channels are considered down. Each machine is composed of several modules and a specific number of modules are kept at the base to support the machines. In case of failure, the defective module is replaced by a spare one and the failed module is fixed either at the base or the depot, depending on the failure type. There are limited repairmen in the system and a failed module is repaired by its own specific repairman at the base or the depot. The authors use Markovian approach to calculate the steady-state operating characteristics of the system.

Fig 2.1 Depot repair (Gupta and Albright 1992)

Diaz and Fu (1997) consider multi-echelon models of repairable parts under a one-for-one replenishment policy. They relax the assumption of ample repair capacity in the inventory model of repairable items and introduce an approximation to deal with limited repair capacity. Their proposed models can be divided into three kinds of queuing models: (1) A model leads to exact results for repair time distribution in case of limited repair facilities; (2) Second model for supporting systems with general repair distribution, in which they relax the exponential service time distribution assumption by applying a two-moment queuing model approximation; and (3) Modeling the system that receives different classes of parts with different repair distribution. They also conduct numerical experiments to show their model's superiority against traditional models when faced with high repair facility utilization. These models have been used by Caracas Metro subway system.

Sleptcheko et al. (2003) analyze the trade-off between inventory level and the number of servers at repair facilities in multi-echelon, multi-indenture service part supply systems. Lau and Song (2004) study a repairable item inventory system under limited repair capacity and non-stationary Poisson demands to minimize the cost of the system.

2.1.3. Lateral transshipments

Some of the research in the domain of repairable parts inventory systems also takes the lateral resupply into account. Lee (1987) develops a continuous-review multi-echelon model for repairable parts when emergency lateral transshipments between bases are allowed. The system has several bases to support the operative items and one central depot to serve the bases in case of backorders. There may be instances where a base is nearer to a neighboring base than the central depot. When a backorder takes place at a base it may be more efficient to receive the item from a neighboring base for reducing lead time and improved service performance. He calls this concept as emergency lateral transshipments. In a follow-up paper,

Axsater (1990) extends the use of emergency lateral transshipments approach in a two-echelon inventory system with repairable items. He focuses on modeling the demand at the base correctly; the demand rate at a base depends on the inventory. In this case only a base with positive on hand inventory receives the demand of emergency transshipments from other bases. This technique leads to an improvement in performance when the proportion of emergency transshipments is high.

The backorder time is an important criterion in repairable parts inventory models and can be influenced by many factors; one of those is the replenishment lead-time, which is the time taken to replenish a part in the warehouse. Dhakar et al. (1994) focus on the stocking levels of high cost, low demand and critical repairable spares from the point of view of cost and production scheduling. The objective of the research is to find the optimal stocking level by applying three different lead time processes (normal, emergency and expedited). The typical lead time process is composed of several states, such as waiting for transportation, in-transit to repair shop, waiting in the repair queue, being repaired and returning to warehouse. The expediting function models the lead time reduction based on the item's current position in the lead time process. They develop a realistic model by considering the failure rate as a function of the number of machines in operation and determine the optimal stocking levels. Their simulation results validate the key role of such a function in the analysis of the optimal number of spares.

Wang et al. (2000) consider a two-echelon repairable inventory system consisting of a central depot and multiple stocking centers. Their research is related to the depot replenishment lead times (DRLTs) which are assumed independent and identically distributed (IID) in previous models. This common assumption in the literature is justified by assuming ample capacity at the depot repair facility. To highlight the problem with this assumption, they define DRLT as the sum of the defective lead time, which is the lead time

from placing an order (by central depot) until receiving the defective part (at the repair center), plus the repair lead time. The DRLTs can be IID when both the defective lead time and the repair lead time are independent and identically distributed. IID repair lead time can be justified by ample capacity at the depot repair facility, but this assumption does not hold for the defective lead time which is dependent on the distance between the center and the depot. Their results show that this assumption can lead to an unrealistic system performance.

Wong et al. (2005) develop an analytical model to determine the spare stocking levels in a single-item, multi-hub, multi-company, repairable inventory system under stock pooling among the hubs and companies. Their model considers delayed lateral transshipments which have been ignored in the previous models. This model can be applicable to an airline's inventory system where the airline has a central store at its hub airport and several smaller stores at other airports it serves. In such a system, inventory pooling can improve the company's logistics performance where the hub store can use the other stores' inventory in case of a backorder. The model minimizes the total system cost consisting of inventory holding, downtime and transshipment costs.

De Smidt-Destombes et al. (2007) consider an installed base of k -out-of- N systems in which each system is composed of identical, repairable components; a k -out-of- N system consists of N identical components in which at least k components are needed for a functional system. They focus on the maintenance interval, the spare part inventory level and repair capacity as the factors that influence the system availability. They use the data of Active Phased Array Radar and Active Towed Array Sonar systems to analyze the relation between the above control variables and systems availability.

Tao and Wen (2009) model a closed loop multi-echelon repairable inventory system to evaluate the effect of repair capacity on expected backorders, and optimize spares location. Their results illustrate their model's superiority over METRIC.

According to the literature, application of the Markovian approach and consideration of the system as a closed-network queue can be a useful way to relax the restrictive assumptions of earlier researchers and obtain more accurate results for repairable parts inventory systems. Besides relaxing the assumptions, the system needs to be modeled in a way that facilitates the computation of desired performance measures. This issue becomes more important under PBC where both parties (customer and the supplier) need some accurate measures of the system performance to set their own controllable parameters. While most of the earlier models in the literature aim to find optimal (minimum cost) base stock taking a centralized approach, PBC tries to consider the system in a decentralized framework and satisfy both the customer and supplier's requirements, simultaneously. Recently some studies have been done to close this gap (Kim et al. 2007; Kim et al. 2011; Nowicki et al. 2008; Oner et al. 2010), but the models still contain restrictive assumptions which decrease the supplier's flexibility and may indicate unrealistic performance to the customers.

2.2 Performance-based Contract

After-sales services are, often, covered in a contract between the two parties (customer and the supplier) and sometimes contractual problems do come up during the contract period. These issues become even more important in PBC where the customer needs a minimum level of performance and the supplier is only rewarded or penalized in accordance with the delivered outcome. As a first step in PBC, the customer selects an appropriate performance measure based on the system properties and the ultimate objective. Next the customer defines an effective metric to measure the specified performance during the contract period. Finally the customer offers a compensation model which determines the payment to the supplier during the lifetime of the contract. The supplier, as the other interested party in the contract, will then set the parameters of the system (such as inventory level, reliability of components

and repair rate) to deliver the performance required. In this section the literature on the most important issues in PBC is reviewed.

2.2.1 Performance and the metric

The performance of an inventory system with repairable parts can be defined in different ways. Modeling the performance in such systems has been studied extensively in the literature (McCall 1965; Pierskalla & Voelker 1976; Sherif & Smith 1981; Valdez-Flores & Feldman 1989; Cho & Parlar 1991; Dekker 1996; Wang 2002; Cassady et al. 2004). Sherbrooke (1968) compares several system performance measures, such as fill rate (defined as the fraction of demands that are immediately filled by supply upon receipt of the requisition), ready rate (defined as the fraction of items that are not in backordered), operational rate (defined as the probability that k or less aircraft have a supply shortage at base, at a random point in time) and backorder (defined as unsatisfied demand at base level). He concludes that backorders are the most reasonable performance measure.

The literature offers some guidelines on the concept of PBC and its fundamental issues by looking at real cases of PBC implementation. Devries (2004) conducts a study to find the most frequent barriers and enablers to effective PBC implementation. Three following questions are posed in this research: a) What are the barriers and how do they influence PBC implementation; b) What are the enablers and how do they influence PBC implementation; and c) What strategy/actions would lead to more successful PBC implementation? The data is collected from program officers, who implement PBC, using a survey. The data analysis supports a direct relationship between the number of enablers, the enhancement of enablers, the influence of enablers (after enhancement) and the success of PBC implementation.

Defining a baseline for the system performance has been studied by many researchers (Billinton et al. 2002, Smith 2006) as it is the basis for the design of an appropriate reward and penalty mechanism. Billinton et al. (2002) study performance-based regulation regime development in electricity industry where the historic data is collected to specify service reliability standards. They use data from Ontario Energy Board to define the reliability standards and evaluate the impact of the reliability baseline on incentives determined by the regulator.

Doerr et al. (2004) present the following propositions about the impact of measurement on performance of PBC to address the unanswered questions related to objectives and measurement: 1) When operational risk is high or difficult to measure, PBC strategies should seek less commercial sector involvement; 2) When commercial sector vendors assume less (measurable) operational risk under a PBC, the duration of that contract should be short; 3) PBC strategies involving less than comprehensive logistical support of a weapon system (e.g. for a component) should nonetheless have integrated weapon system models in support of their business case study; 4) In the management of ongoing PBC, metrics should primarily concern themselves with valued outcome, and should be related to weapon system cost, readiness, and agility. Process measures should only be applied when key operating decisions depend on the status of the process itself; and 5) Operational risk in key performance measures should be tracked for ongoing PBL contracts. When essential to mission support, a reduction in variability should be supported with appropriate incentives.

Berkowitz et al. (2005) conduct a study on implementing PBC as a mechanism for integrating the acquisition and sustainment of systems in the Department of Defense and suggest the following six guidelines for successfully implementing PBC: 1) Assign responsibilities clearly throughout the firm; 2) Design metrics to motivate the right behavior; 3) Manage failures to limit disincentives for risk-taking; 4) Develop a supportive

organizational context for tools; 5) Manage relationship with stakeholders; and 6) Benchmark to promote continuous improvement.

Dang et al. (2009) evaluate PBC as a way to integrate acquisition and sustainment of systems, and leverage commercial best practices to reduce costs, improve performance, and ensure operational readiness. They introduce public law, culture, funding, training and education, risk, and contractors on the battlefield as the barriers to implementing PBC. They find that the key to successful implementation of PBC is a workforce of Program Managers (PMs) who are knowledgeable and trained to develop straightforward, measurable, and achievable metrics focused on top-level objectives. Their research also highlights the tracking of performance data to ensure PBC is achieving the expected results of reduced costs and improved performance and reliability.

Ng et al. (2009) look at outcome-based contracts as an excellent evidence of organizations shifting from goods-dominated logic to service-dominated logic. They analyze two contracts between defense contractors and the UK Ministry of Defense (MoD) to support (maintenance, repair and overhaul) the fighter jets and the missile systems for their operations life. Interviews are used to collect data from stakeholders to find the value of co-creation within outcome-based contracts. In co-creation strategy firms and customers share, combine and renew each other's resources and capabilities to create value through new forms of interaction, service and learning mechanisms. They find that in outcome-based contracts both the supplier firm and the customer should join together to reach an effective value co-creation model rather than individually efficient models. In other words, combination of two local optimums can increase overall system costs, leading to a more expensive contract.

As availability is a preferred performance measure under PBC, identifying the elements which influence the availability and the efforts that can improve the availability of the system becomes important. Kang et al. (2005) find that many PBC concentrate on the

contractual agreement for improving component level performance, while what really matters is the total system operational availability. Therefore, they develop a methodology to calculate the system operational availability based on component-level reliability and maintainability data. They believe that the PBC manager should have a sense of the distribution of operational availability around the target levels. They focus on the performance and define readiness risk as the probability that the availability remains below a certain level. Then, they develop a discrete-event simulation model, using Arena simulation software, and conduct a weapon-system-level risk analysis by evaluating the interaction between reliability, time-to-repair and operational availability. They find that the time-between-failure, the spare inventory and time-to-repair are three key metrics which have higher impact on the availability of the system.

In a follow-up paper, Kang et al. (2006) extend the simulation model to consider both availability (average percentage of assets that are available for operations) and readiness risk (probability that a vendor will fail to deliver a desired threshold of operational availability) as two measures of performance. This approach introduces the most effective alternative for increasing the availability and reducing the readiness risk of the system, simultaneously. Using simulation they model the outcome performance and its cost for a combination of four logistics elements: inventory service level, failure rate, repair rate and logistics delay. They then use a design-of-experiments approach to determine the combination of logistics elements which can facilitate achieving the greatest readiness availability. Their results show that the readiness risk is affected by failure rate and logistics delay.

PBC is one key defense procurement strategy which uses pre-defined metrics to evaluate operational availability and total cost of ownership. PBC aims to improve the availability, reliability, maintainability and supportability at a lower cost. Kumar (2007) introduces reliability, maintainability and supportability as three integral components of

system operational effectiveness (SOE). He calculates the operational availability by dividing the mean time between maintenance by the sum of the down time and the mean time between maintenance. He extends a mathematical model, using Global Programming, to optimize the following measures, simultaneously: (1) Total cost of ownership (TCO), (2) Probability of No stock-out for spares (PNS), (3) System mean-time-to-repair (MTTR), and (4) Logistics footprint.

Kim et al. (2007) report that after-sales support services in the U.S economy earn twice as much profit as do sales of original products. They concentrate on applying PBC in capital-intensive industries such as aerospace and defense where the maintenance contracts are not based on a fixed payment for warranties, because of high uncertainty in costs and repair process. They model a single-location spare-parts inventory system under a principal-agent contract in which the customer (principal) contracts with individual suppliers (agents) who maintain the inventory and repair facilities to support N identical systems where each system has n subsystems. A closed-loop repair process is assumed because of subsystems' high cost and long lifetimes (Figure 2.2). They assume a fixed failure rate for each subsystem and a fixed repair lead time at the repair facility. When a subsystem fails, it is replaced by a ready-to-use subsystem from inventory and the failed subsystem is transported to repair facility which has sufficient capacity. If there is no stock of repaired units, the system stays down until it receives the required subsystem from the repair facility. They analyze this model under three most common types of contracts in high-technology industries (fixed price, cost-plus, and PBC) to find the optimal combination of parameters that result in the greatest outcome for the customer. In their compensation model, the customer pays a part of the supplier's cost, which can be reduced by the its efforts, and penalizes it depending on the number of backorders. To prevent backorders, the supplier may resort to keeping higher inventory, which may not be optimal for the customer. To address this issue, Kim et al.

(2011) focus on the tradeoff between spare asset management and investment into product reliability, by considering a variable failure rate. They find that the spare asset ownership plays a key role in achieving a good balance between inventory and reliability levels.

Fig 2.2 Closed-loop cycle for repairable items (Kim et al. 2007)

There are many estimation models which study the cost structure of the weapon systems (Kirkpatrick 2004; Yoner 2001; Kang 2001). Restrictive assumptions such as a constant failure rate and lack of attention to relationship between the initial purchase cost and the maintenance cost in such models can lead to the acquisition of a less expensive weapon system but one with lower performance. For addressing this issue, Sohn et al. (2009) extend a cost of ownership (COO) model to assess the lifetime cost of weapon systems. COO model is applied in investment decision for systems which have high initial purchase price but ultimately result in lower operational costs. They implement this model to the purchase of spare engines for the Korean navy.

Keeping the availability of complex technical systems (such as computer networks, defense, and medical systems) at a certain level is vital since operations may suffer when the

systems fail. In such systems, the failure of systems leads to significant losses for the users; if the duration of the backorders is long, the losses may grow exponentially (Oner et al. 2010). They believe that the availability of the system is specified by its reliability and the speed of its repair facility. The speed of system repair activities is improved by the repair-by-replacement approach in which a failed part is replaced by a ready-to-use one. In practice, spare parts are kept at a nearby warehouse to support the systems in case of failure. They name the process of replacement from warehouse as ordinary process. In case of out-of-stock situation, they consider an emergency process in which the demand is fulfilled from a more distant warehouse, instead of waiting for a part from the repair facility. The emergency process is more costly but can decrease the system down time, which is the real goal of PBC. They formulate the system life cycle costs (LCC) based on the component's reliability and spare part inventory level, which consists of design costs, production costs, and maintenance and downtime costs. Their results show that the value of components, number of systems, down time penalty rate and length of exploitation phase are the key parameters that influence component reliability.

A review of existing literature on performance and metrics for repairable systems under PBC strongly suggests using availability as the key performance measure. The literature also reveals a common approach to measuring the availability of the system by the average number of backorders. Kim (2008) also studies a repairable inventory model under PBC in which the availability of the system is formulated based on such an approach. In this research the failure rate and repair rate have been assumed constant as relaxing these assumptions may alter the results and render them unfit for comparison.

The PBC have been used widely in both commercial and military sectors and quite a few recent articles (Kim et al. 2007, Nowicki et al. 2008, Oner et al. 2010, Mirzahosseini and Piplani, 2011) focus on their design and implementation. In a single echelon repairable

part inventory system, the supplier can improve the system availability by managing the base stock level at the warehouse, component reliability, and efficiency of the repair facility. The interactions between service time or reliability and the base stock decisions have been modeled in recent papers where the supplier controls more than one action (Sleptchenko et al. 2003; Kim et al. 2007), however no work has been found that studies reliability improvement and its interaction with service time decisions under different contract terms.

2.2.2 Incentives scheme

The classical agency model considers the tradeoff between incentive and insurance, wherein the principal and the agent sign a contract and the principal compensates the agent for delivering the outcome specified in the contract. Fearnley et al. (2004) design a PBC between the Ministry of Transport and Communication (MOT) and the Norwegian State Railways (NBS) to fulfilling the social goals of MOT and NBS's profit-maximizing objective. They believe in giving complete freedom to the supplier under PBC as it has superior knowledge of logistics functions. The results highlight the importance of maintaining a balance between penalties and rewards under PBC to minimize the risk of achieving a low service quality. In a similar work, Hensher and Houghton (2004) study performance-based quality contracts for the bus service in Australia to ensure that operators meet the service levels in line with the government's objectives. They aim to maximize the social surplus subject to the following constraints: minimum required service level from community service obligations (CSO), the commercial imperative of the operator and existing cap of budgeted subsidies. Their study reveals the importance of an appropriate definition of system profit under PBCs.

A reward or penalty plan under PBC, according to a consideration of lower and upper limit on system performance, has been developed by several studies, including Billinton et al. (2002). The concept of such an incentive mechanism is to award bonus if the performance

exceeds upper limit and penalize for a performance outcome below the lower bound. Three zones are defined as per outcome performance: the penalty, bonus and dead zone. The three zones are illustrated in Figure (2.3).

Fig 2.3 The concept of Bonus, Dead and Penalty zones (Sols et al. 2013).

As can be seen, there is no reward or penalty if the system performance falls in the dead zone. The amount of penalty and reward under Penalty and Bonus zones are calculated in accordance with negotiated rates, subject to a pre-specified cap.

As delays in industrial projects are costly for owners, establishing a time reduction strategy is becoming important in many projects. One of such strategies is termed “schedule incentives and disincentives” where the contractor receives rewards for early completion and is penalized for each day of delay. Bubshait (2003) studies the influence of using incentive/disincentive contracting on the contractor’s decision to manage and control the project duration. His findings also emphasize the effect of continuous evaluation of performance on productivity of the contractor.

Gibbons (2005) reviews the incentives between firms under the subject of “agency theory” (i.e. the economic theory of incentives). His research arrives at three key insights:

- (1) First, “you get what you pay for” problems with objective performance measure are as important as the tradeoff between incentives and insurance in the agency theory.
- (2) Second, the incentive contracts are not the only way to implement incentives. For example, the market competition can lead to structural incentives without the need for contracts.
- (3) Finally, a powerful technique to use in incentive contracts is the integration decision which can be enforced by creating asset co-ownership.

Kim et al. (2010) show that when the supplier does not have any control on the system failure rate, using average downtime (rather than cumulative downtime) leads to higher performance. They also find that when the supplier has ability to reduce the system failure rate, the cumulative downtime turns out to be superior to average downtime. They focus on systems with infrequent failure; however, there are many examples of systems with a high failure rate. To simplify the analysis, they also assume a constant failure rate and ignore the dependency of the failure rate to the number of operational systems and the service rate.

Mirzahosseini and Piplani (2013) consider a single echelon repairable part inventory system and conduct a parametric analysis to evaluate the impact of the total number of systems on the supplier’s decisions under a performance contract. Their results show that the number of systems affects the supplier’s response when the customer uses the average number of backorders as the negative incentive.

In accordance with agency theory literature, incentive and insurance are two integral parts of a compensation model. Insurance, which decreases the supplier risk, can be used in the reimbursement model as an appropriate fixed payment. The second important part of compensation model, especially in PBCs, is an effective incentive to motivate the supplier into delivering the desired outcome. Incentives can be monetary or non-monetary, positive or

negative. Reimbursement for supplier's costs as a positive incentive has been widely used in the literature (Scherer 1964; McAfee and McMillan 1986, Kim et al. 2007). Utilizing both positive and negative incentives at the same time can satisfy both the parties by reducing their risk under uncertain conditions.

The customer's operating fleet size supported by the supplier in a performance contract may change over the contract's lifetime. Jin and Tian (2012) study the performance-based service operation in semiconductor industry where the fleet size changes as new generation of equipment is released to the market. They examine the trade-offs between component reliability and inventory level. The spare parts inventory model considering stochastic growth of an installed base is developed by Jin and Liao (2009). Their objective is to determine how an increase in the fleet size impacts the service demand.

These studies show that any changes on the supported fleet size may impact the supplier's response over the contract's lifetime. In such a case, the customer needs to update the contract terms to ensure a desired outcome performance. The literature review reveals a gap in how performance contracts should be designed and adjusted as the fleet size changes.

2.2.3 Supplier's risk attitude

Much research has been done to capture the risk associated with long-term implementation of a PBC by designing a reliability growth plan (Wasserman and Lamberson 1992, Smith 2006). Using an appropriate motivational mechanism to mitigate the risk of unstable decisions by the supplier during the contract period is taken into account by Berends (2000). He analyzes eight case studies from the chemical and oil industries to show a strong positive correlation between the performance and incentive mechanism. He also highlights the importance of comprehensive involvement in continuous evaluation of performance by the customer, over the contract life-time.

Prendergast (1999) evaluates the existing research for the provision of incentives, as the essence of economics. He studies evidence related to the classical agency model to study the compensation issues. The classical agency theory assumes that the contract function is linear, while the optimal contract is linear only under very special assumptions (Holmstrom 1979, and Banker and Datar 1989). A linear compensation function (or revenue function) will be of the form $s + b \times y$, where s is the fixed payment and $b \times y$ represents the performance incentive.

Mirrlees (1974) studies various nonlinear contracts under principal-agent models. His results show superiority of the nonlinear contract (in particular, a step-function contract) to the best linear contract. Murphy (2000) studies the application of a piecewise-linear-threshold in 177 annual incentive plans for companies in different industries, such as utility, finance and industrials. Zhou and Swan (2003) evaluate a piecewise-linear-threshold with a discrete jump at the threshold. They focus on the contractual nonlinearity associated with performance thresholds to indicate optimality of the performance threshold payment scheme.

The impact of risk aversion in reimbursement contracts has been studied in literature. Scherer (1964) discusses the financial risk provision at the defense performance contracts; McAfee and McMillan (1986) study government incentive contracts under moral hazard situations. They model an optimal linear contract in the presence of several risk-averse suppliers. Their results indicate that the optimal contract is an incentive contract and not a cost-plus. Holmstorm and Milgrom (1987) consider the problem of providing incentives for a risk-averse agent with an exponential utility function. Their findings show that a linear contract is optimal when the agent's problem is multi-period. So and Tang (2000) examine the impact of reimbursement contracts on the healthcare industry. They consider a linear incentive function to specify the drug manufacturer's response to the insurer's contract terms by capturing the clinic's risk. Hemmer (2004) focuses on the Linear-Exponential-Normal

(LEN) principal-agent model which assumes a linear contract, exponential utility function and normally distributed output. He finds that the LEN model performs inappropriately as the piecewise-linear-threshold contract is superior to the best linear contract.

Kim et al. (2007) consider a linear compensation model for in PBC consisting of three components: (1) a fixed payment, (2) reimbursement for supplier's cost (see Scherer 1964; McAfee and McMillan 1986 for more detail on cost reimbursement contracts), and (3) a backorder-contingent incentive payment. In this case the principal reimburses a percent of agent's cost but penalizes the agent for each backorder at a constant penalty rate. They assume that both the principal and the agent have expected mean-variance utility functions, widely used in operations management literature (Chen and Federgruen 2000; Van Mieghem 2007).

Chen and Miller (2009) consider principal-agent models where the agent chooses actions over time and is reimbursed based on the final value of output. They use numerical simulation to find the difference between the linear contract's outcome performance and the piecewise-linear-threshold contract's relative performance. The latter contract applies a compensation model in which the payment is constant until a pre-specified threshold and increases linearly for performance that exceeds the threshold, until reaching a cap (see Murphy 2000; Zhou and Swan 2003 for more discussion). They test two well-known agent utility functions in economic analysis, called exponential and power (Raju and Srinivasan 1996; Bliss and Panigirtzoglou 2004), under linear and piecewise-linear-threshold contracts. Their results support the application of the linear contract only when the agent's utility function is exponential. The piecewise-linear-threshold leads to a higher performance only when the agent has a power utility function.

Nowicki et al. (2008) use revenue, which is a function of operational availability, as the supplier reward under PBC. They test three types of revenue functions (linear,

exponential and step) in a multi-echelon, multi-item supply chain to show the significant impact the revenue function has on inventory level, availability and total profit. They show that the form of revenue function plays a key role in a performance contract; the failure rate and repair rate are assumed to be constant in this model. They also assume that the agent (supplier) is risk neutral and do not incorporate risk consideration in their research. Thus, the literature on risk under performance contracts reveals a lack of consideration of the supplier's risk attitude on the design of incentive contracts.

2.3 Summary

PBC is envisioned to lower the asset ownership cost at the guaranteed performance level. While most of the earlier models in the literature aim to find optimal (cost minimizing) base stock subject to satisfying a specified service level using a centralized approach, PBC considers the system in a decentralized framework to satisfy both the customer and the supplier's requirements, simultaneously. Recently some studies have been done to close this gap but their models still have restrictive assumptions which decrease the supplier's flexibility and may give indication of unrealistic performance to the customers. PBL should empower the provider with the authority and responsibility to control those elements required to successfully support the program. Relaxing these assumptions, thus, offers an opportunity to develop a robust model which may influence the contractual elements, especially metric formulation.

Studies on reliability and spare parts logistics have long been carried out in parallel, each striving for high system availability at a lower cost. High system reliability implies a longer MTBF, fewer failures, and a smaller amount of part replacements. Spare parts logistics intends to streamline the material supply so as to mitigate the backorders or ensure

the fill rates, shortening the equipment downtime. Recently, there is a small, but growing literature stream that jointly allocates the reliability and the spare parts to attain the desired service performance.

The customer operating fleet size supported by the supplier in a performance contract may change over the contract's lifetime due to various reasons. Studies show that any changes in the supported fleet size may impact the supplier's response over the contract's lifetime. In such a case, the customer needs to update the contract terms to ensure the desired outcome performance. The literature review reveals a gap in the design and adjustment of performance contracts due to changes in fleet size.

In a PBC, the supplier said to be risk-averse if the value specified in terms of profits is less than the expected profit. Such value is specified by defining a function, called the utility function. The models that evaluate behavior of a risk-averse supplier are known as moral hazard models in the literature. According to the supplier's risk attitude, the utility function can be defined in different forms such as exponential, power or mean-variance. The literature of capturing risk under performance contracts reveals a lack of understanding of how incentive contracts should be designed based on the supplier's risk attitude.

Next few chapters try to fill the above research gaps in PBCs implementation.

Chapter 3

Repairable parts inventory system

PBC is becoming a dominant logistics support strategy in both commercial and defense sectors. PBCs are designed to serve the customer's key performance measures, while the traditional contracts for after-sales services, such as Fixed price (FP) and Cost plus (C+), only provide insurance or incentive. Performance measures can be affected by various factors such as reliability, response time and logistics footprint (e.g. inventory). Controlling these factors and their relationship with metrics and performance measures is vital for the supplier to meet the performance level that satisfies the customer goals. The customer also needs to define effective metrics to monitor the supplier's performance during the contract period and ensure PBC's success.

The after-market support system (under PBC) is traditionally modeled as a repairable parts inventory system. When a component fails it is replaced by a new component keeping in the warehouse. The failed component then transfers to the repair facility for required repair process. In this chapter, the system is modeled as a queuing network by enhancing the classical repairable parts inventory model. The classical model is improved by relaxing some restrictive assumptions, such as fixed failure rate, fixed repair rate and infinite capacity at repair facility. The model focuses on the interaction between inventory management, component reliability and repair facility efficiency. It is then analyzed to determine the factors that have significant impact on the system performance. The influence of component's reliability, service capacity and base stock level on average number of backorders and average waiting time (as two applicable metrics), and some common performance measures, such as fill rate, ready rate and availability are examined. The results

show that the base stock level of the spare parts has negligible effect on the system availability and may lead to a high penalty level by increasing the average down time. Thus, to achieve a desired performance outcome, the supplier has to improve the component reliability and the repair time, rather than invest in building up a stock of spares. In addition, two metrics are formulated that allow the customer to monitor the key parameters affecting the system performance.

3.1 The Inventory System

A closed-loop inventory system consists of a repair facility and a single warehouse, and supports the customer systems under a PBC, similar to the one used in by Kim et al. (2007), Kim et al. (2011), Oner et al. (2010) and Mirzahosseini and Piplani (2011). Each system to be serviced (such as a fighter jet) consists of several major repairable components (e.g. engine, avionics, weapon system). The associated inventory system is illustrated in Figure (3.1).

A component's failure is assumed to follow an independent Poisson distribution with constant rate λ ; the airline industry data supports the validity of this assumption (Wong et al. 2005)(see Appendix D for more complex scenarios). Following this assumption, the component failure distribution in a closed-loop system consisting N systems follows a Poisson process with the variable rate $\lambda(z)$, where z is the number of operational systems. The repair facility is modeled as an $M/M/m$ queue. The repair lead-time of each server in the repair facility, consisting of the repair time and transportation delay, is assumed to be an independent exponential random variable with mean $(\mu)^{-1}$; this assumption is common in the literature (Wong et al. 2005). Alfredsson and Verrijdt (1999) show that the performance of the inventory system is insensitive to the lead-time distribution. The time between replacing components from repair facility to the warehouse follows Exponential distribution with the

variable rate $\mu(y, m)$, which is a function of the number of components in repair facility y and the number of servers m .

A one-for-one base stock S replenishment policy is followed at the warehouse, an appropriate policy for high-cost, low-demand items (Sherbrooke 1968). As per this policy a failed component is replaced by a ready-to-use component (new or refurbished) from the warehouse and the failed one is sent to the repair facility. A backorder takes place when the warehouse is out of stock. The warehouse triggers a replenishment order immediately and the information lead time is assumed to be zero.

Fig 3.1 Closed-loop inventory system

3.1.1 Notation

- N : Total number of systems
- λ : Failure rate of each component
- $1/\mu$: Repair lead-time of each server
- m : Number of servers in repair facility
- z : Number of operational systems
- $\lambda(z)$: Fleet's failure rate

y : Number of components in repair facility (on-order inventory)

S : spare parts inventory at the warehouse

x : The inventory level in the warehouse

π_x : The steady-state probability of inventory level x at the warehouse

$\mu(y, m)$: Replacement rate of components at the warehouse (by the repair facility)

$E(B)$: The average number of backorders

$E(T_{wait})$: The average waiting time spent by a failed part in the repair system

$MTBF$: Mean time between failures

$MTTR$: Mean time to replace

$F(S, \mu, \lambda)$: Fill rate

$R(S, \mu, \lambda)$: Ready rate

$A(S, \mu, \lambda)$: The operational availability

The inventory level in the warehouse is represented by x , which can change from $-N$ to S ($-N \leq x \leq S$), the negative x representing the number of backorders. The relationship between inventory level x , on-order inventory y (the number of components at the repair facility) and the number of operational systems z is formulated in Equation (3.1).

$$z = \begin{cases} N & \text{if } x \geq 0 \\ N + x & \text{otherwise} \end{cases}, \quad y = S - x \quad \text{for } -N \leq x \leq S \quad (3.1)$$

The component failure rate and the component replacement rate in the system can be computed as follows:

$$\lambda(z) = z\lambda \quad \text{for } 0 \leq z \leq N, \quad \mu(y, m) = [\text{Min}(y, m)]\mu \quad \text{for } 0 \leq y \leq S + N \quad (3.2)$$

3.1.2 The Markov model

Application of the Markovian approach and consideration of the system as a closed-network queue is a powerful way to relax the restrictive assumptions of earlier researchers and obtain more accurate results in the repairable parts inventory systems (Gross et al. 1985, Gupta and Albright 1992, Diaz and Fu 1997, and Wong et al. 2005). In this section, the corresponding Markov model is developed to represent the state space x .

There are two events that change the state of the system: (1) A component fails and the supplier replaces the failed component with a ready-to-use one from the warehouse (sending the failed component to the repair facility); (2) A repaired component arrives at the warehouse from the repair facility. The transition diagram of the system is shown in Figure (3.2).

Fig 3.2 The transition diagram of the inventory system.

Let π_x be the steady-state probability of inventory level x at the warehouse. The flow balance equations are laid out in Equation (3.3).

$$\pi_x (z\lambda + [\text{Min}(y, m)]\mu) = (Z_N) \times \pi_{x-1}([\text{Min}(y+1, m)]\mu) + (Z_S) \times \pi_{x+1}([\text{Min}(z+1, N)]\lambda) \quad \text{for } -N \leq x \leq S \quad (3.3)$$

$$Z_N = \begin{cases} 1 & \text{if } -N+1 \leq x \leq S \\ 0 & \text{otherwise} \end{cases} \quad Z_S = \begin{cases} 1 & \text{if } -N \leq x \leq S-1 \\ 0 & \text{otherwise} \end{cases}$$

The left side of Equation (3.3) represents the average rate of transition out of state x . The terms in the bracket indicate the transition due to failed component, which depends on the number of working systems z , and the replenishment rate from repair facility to the warehouse, which is a function of the on-order inventory y and the number of servers at the repair facility m , respectively. The right side of Equation (3.3) represents the average rate of transition into state x . The terms determine the rate at which the replenishment components arrive at the warehouse and the transition to satisfy the demand when a component fails, respectively. Binary variables Z_N and Z_S denote whether the number of backorders is N and the inventory level at the warehouse is S , respectively. The balance Equations (3.3) and the normalizing constraint (3.4) are solved to compute π_x .

$$\sum_{x=-N}^S \pi_x = 1 \quad (3.4)$$

3.2 The Metrics and System Performances

Under PBC, metrics should be related to an element which has a direct link to the defined performance. The average number of backorders has been used in earlier models when the desired system performance is availability. This approach just decreases the number of backorders and does not influence the length of backorders, while “the metric should depend on the length and the number of backorders.” (Sherbrooke, 1968) Therefore, the average number of backorders and average waiting time are developed as two effective metrics under

PBCs. Given the steady-state probabilities, the average number of backorder can be computed as follows:

$$E(B) = \left| \sum_{x=-N}^{-1} x\pi_x \right| \quad (3.5)$$

According to Little's Law, in a queuing process, the average number of components in the system is equal to the arrival rate times the average waiting time. Therefore, the average waiting time spent by a failed part in the repair system, $E(T_{wait})$, can be calculated, by applying Little's Law, as:

$$E(T_{wait}) = \frac{E[y]}{E[\lambda(z)]} = \frac{\sum_{x=-N}^S (S-x)\pi_x}{\sum_{x=0}^S [\lambda(N)]\pi_x + \sum_{x=-N}^{-1} [\lambda(N+x)]\pi_x} \quad (3.6)$$

Where $E[y]$ denotes the average number of components in the repair facility and $E[\lambda(z)]$ the arrival rate of failed units to the repair system.

The performance of inventory systems with repairable parts can be measured in different ways. Performance modeling in such systems has been studied extensively in the literature (McCall 1965; Pierskalla & Voelker 1976; Sherif & Smith 1981; Valdez-Flores & Feldman 1989; Cho & Parlar 1991; Dekker 1996; Wang 2002; Cassady et al. 2004). According to the Markovian approach, the steady state probabilities of the number of units in the warehouse are developed at a random point in time. Therefore, different measures of system performance can be calculated based on probabilities that are related to performance in a single-echelon system. For the different measures, the supplier can improve the system performance by managing three parameters: (1) base stock at the warehouse, (2) component reliability, and (3) efficiency of the repair facility.

The fill rate F is one of common performance measures used in practice. Given a failure rate λ , service rate μ and a base stock level S , the fill rate is defined as a fraction of the demand that can be fulfilled immediately from the warehouse. Thus, the fill rate is the probability of having at least one unit at the warehouse.

$$F(S, \mu, \lambda) = \sum_{x=1}^S \pi_x \quad (3.7)$$

It is intuitively clear that the fill rate is an increasing function of base stock S and service rate μ but a decreasing function of the component's failure rate λ .

The ready rate R is the second commonly used performance measure under PBCs. The ready rate refers to the probability that a component has no backorders at a random point in time meaning that the inventory level is nonnegative. Given the steady-state probabilities, the ready rate can be calculated according to Equation (3.8):

$$R(S, \mu, \lambda) = \sum_{x=0}^S \pi_x \quad (3.8)$$

Similar to the fill rate, as the base stock S and service rate μ increase the ready rate will increase. The ready rate is also a decreasing function of the failure rate λ .

The operational availability is used to measure the performance of the system as the third common measure. Following Kim et al. (2007), and assuming each backordered component results in an un-operational system; the availability of the entire fleet can be defined as:

$$A(S, \mu, \lambda) = 1 - E(B) / N \quad (3.9)$$

Under PBC, the system needs to be modeled in a way that facilitates computation of some effective metrics and allows the customer to monitor the key parameters affecting system availability. This allows both parties to ensure that the delivered performance at the end of the contract meets or exceeds the target. In this research $MTBF$ is defined for

controlling the component reliability and $MTTR_e$ for monitoring the efficiency of repair facility. The $MTBF$ measure based on the current model is more accurate than the previous measures as it considers the dependency of subsystem failure rate to the number of operational units. Also, the proposed model measures the $MTTR_e$, instead of $MTTR$, which considers the repair time and transportation delay, simultaneously. It is easier for the customer to monitor the warehouse replacement rate rather than the repair rate. This issue becomes more important under PBL contracts where the supplier may be more hesitant in offering detailed information, such as repair rate, during the contract's lifetime. These two control measures are modeled by Equation (3.10).

$$MTBF = E[1 / \lambda(z)] \quad 0 \leq z \leq N, \quad MTTR_e = E[1 / \mu(y, m)] \quad 0 \leq y \leq S + N \quad (3.10)$$

The first term of Equation (3.10) represents the Mean Time between Failures and the second one represents the Mean Time to Replace. The $MTBF$ and $MTTR_e$ can be represented as a function of the steady-state probabilities:

$$\begin{aligned} MTBF &= \sum_{x=0}^S [1 / \lambda(N)] \pi_x + \sum_{x=-N+1}^{-1} [1 / \lambda(N+x)] \pi_x \\ MTTR_e &= \sum_{x=-N}^{S-1} [1 / \mu(\text{Min}\{m, (S-x)\})] \pi_x \end{aligned} \quad (3.11)$$

3.3 Numerical Study

An Unmanned Aerial Vehicle (UAV) system consists of four air vehicles (AV's), two ground-control stations (GCSs), modular mission payloads (MMPs), data links, remote data terminals (RDTs) and an automatic landing system. The AV flying hours is estimated at 120 hours per month, per vehicle. The critical components of AV (engine, propeller, avionics

computer) are considered to support 10 UAV systems with 40 air vehicles for one year (Kang et al. 2005). The parameters of the system are listed in Table (3.1).

Table 3.1 Basic parameters of the system

<i>Component</i>	$(\lambda)^{-1} hrs$	$(\mu)^{-1} hrs$	<i>S</i>
<i>Engine</i>	750	123	6
<i>Propeller</i>	500	89	8
<i>Avionics</i>	1000	135	4

As mentioned earlier, the repair lead time of each server $(\mu)^{-1}$ includes the repair time and the transportation delay. For example, when an engine fails, it incurs 4.5 days or 108 hours (2.25 days each way) of transportation delay, along with 15 hours of repair time, on average. Similarly, the repair lead time calculations of propeller and avionics computer are shown in Table (3.1). The mean time between failure of an engine, propeller and avionics computer is expected to be 750, 500 and 1000 hours, respectively. By considering the AV flying hours per year (12×120), the annual failure rate of each engine, propeller and avionics computer is $1.92(1440/750)$, $2.88(1440/500)$ and $1.44(1440/1000)$, respectively. The above example is modeled as a Markovian queue (Equations 3.3-3.8) which is then coded in MATLAB software (Appendix A). The resulting metrics (average number of backorders, average down time) and performance of the system (Fill rate, Ready rate and Availability) are calculated and shown in Table (3.2).

Table 3.2 System performance and the metrics

<i>Component</i>	<i>E(B)</i>	<i>E(D)days</i>	<i>F</i>	<i>R</i>	<i>A</i>	<i>MTBF</i> (hrs.)	<i>MTTRe</i> (hrs.)
<i>Engine</i>	4.26	50	0.27	0.32	0.89	130	118.5
<i>Propeller</i>	6.19	51	0.16	0.2	0.84	92.3	88.1
<i>Avionics</i>	1.64	28.5	0.5	0.6	0.96	159.5	114.9

3.4 Parametric Analysis

A parametric analysis is conducted to study the effect of varying component failure rate, server repair rate, number of servers and the base stock level on the metrics and system performance. The minimum and maximum possible values of parameters of the components are tabulated in Table (3.3).

Table 3.3 Parameter range

<i>Component</i>	$(\lambda)^{-1}hrs$		$(\mu)^{-1}hrs.$		m		S	
	<i>Min</i>	<i>Max</i>	<i>Min</i>	<i>Max</i>	<i>Min</i>	<i>Max</i>	<i>Min</i>	<i>Max</i>
<i>Engine</i>	550	950	105	165	1	3	0	12
<i>Propeller</i>	300	700	70	130	1	3	0	12
<i>Avionics</i>	800	1200	125	185	1	3	0	12

Figure (3.3) illustrates the computational results of the analysis for availability performance (see Appendix A for the other performances and metrics analysis results). Different values of the parameters are used, which range from minimum to maximum values in steps of 100, 15, 1 and 2 for mean time between failure, mean time to repair, number of servers and the base stock level, respectively.

(a) Impact of Repair Rate on Availability

(b) Impact of Base Stock on Availability

(c) Impact of Failure Rate on Availability

(d) Impact of Server Number on Availability

Fig 3.3 Impact of System parameters on Availability

As can be seen in Figure (3.3), decreasing the failure rate, increasing the server repair rate and the number of servers increases the availability of all the components. Figure (3.3)(b) indicates that the base stock level of the components has negligible effect on the availability. Thus, for increasing availability the supplier must invest in improving the

component reliability and the efficiency of the repair system rather than increasing the base stock level.

$MTBF$ can be used for controlling the component reliability and $MTTR_e$ for monitoring the efficiency of repair facility. The $MTBF$ measure based on the developed model is more accurate than the previous definitions, because it considers the dependency of subsystems failure rate to the number of operational units. Also, the proposed model measures $MTTR_e$, instead of $MTTR$, which considers the repair time and transportation delay together. It is easier for the customer to monitor the warehouse replacement rate rather than the repair rate. This issue becomes more important under PBCs where the supplier may be hesitant in offering detailed information such as the facility repair rate during the contract's lifetime.

Thus, for improving the system performance and preventing a high penalty under PBC, the supplier must invest in improving the component reliability and efficiency of the repair system, rather than increasing the base stock level. The next chapter focuses on the reliability improvement and its interaction with the service rate under performance-based contract.

3.5 Summary

In this chapter an enhanced inventory model is proposed for repairable spare parts to illustrate the interaction between key parameters of such systems under the PBC, an increasingly popular strategy in both private and public sectors. Relaxing some unrealistic assumptions in the classical repairable spare parts inventory models, such as fixed failure rate, fixed repair rate and infinite repair facility capacity makes the model more robust. In addition, the model helps with the identification of appropriate levers to the supplier and the

customer, as the two main parties in the PBC, who can manipulate them and guide the system performance towards the target, which is one of the ultimate goals of PBC.

A study of barriers to PBC implementation indicates a lack of effective metrics that allow the customer to monitor the supplier's real performance during the contract's tenure. Two metrics were formulated that facilitate monitoring (and control) of the supplier's actual performance. The effect of key parameters on the system performance (fill rate, ready rate and availability) and the metrics (average number of backorders and average waiting time) also was evaluated that may be used to penalize the supplier. The results show that the base stock level does not have any significant effect on the average number of backorders, in comparison to other parameters, such as repair rate and repair capacity, and may lead to penalties for the supplier. Furthermore, the model recommends focusing on component reliability and repair system efficiency to improve the performance of the system with repairable spare parts.

Chapter 4

Incentive Design in Performance-based Contracts

Business associated with after-sales support represents 8-10% of the gross domestic product (GDP) in the U.S. (total U.S. GDP was \$15.094 trillion in 2011). In public sector, the annual operation and sustainment cost for the U.S. military equipment alone is \$63 billion; the entire military fleet is maintained by 678,000 DoD personnel along with hundreds of private contractors (Smith 2007). In wind-energy industry, the U.S. installed wind-energy capacity increased by 309% between 2006 and 2011, and the global installed capacity has grown 27% each year and reached 237,000 MW in 2011 (GWC 2012).

PBCs have emerged as an innovative after-sale support strategy in capital-intensive industries where compensation is determined by the performance delivered. System availability, widely used as the performance measure, is affected by multiple factors such as reliability, spares, fleet size, and service response time.

The supplier may be unwilling to reveal his efforts level and achievements during the contract period. Also, the wide of variety of contracts available and the complexity of these contracts is major hindrance in deciding the contract type and content terms. These issues make it difficult to develop a universal service contract guideline.

The difficulty in sharing information is another problem in establishing a proper coordination between customer and supplier. Information sharing is difficult due to the fears of information leakage and conflicts of interest between customer and supplier. As a result, it is hard to build an information sharing platform for information sharing and delivery coordination.

Establishing a right mechanism to allocate the incentives of coordination among the supplier and customer, is another challenge. It is difficult to find a balance point and even harder to justify it, especially when information sharing is partial and asymmetric.

Prior studies either focus on the parts availability upon failure, or advocate the reliability growth in design. In this chapter, a single echelon repairable inventory system is considered with focus on reliability improvement and its interaction with decisions affecting service time taking into account the fleet size. It is found that the component reliability in a repairable inventory system is a function of the operating fleet size and the service rate. Then, a principal-agent model is developed to evaluate the impact of the operating fleet size on the incentive mechanism. The results show that the fleet size plays a critical role in determining the penalty and cost sharing rates when the customer adopts the average backorders as the negative incentive.

4.1 Performance-based Contract

The design of a PBC follows a series of decision by the two parties involved. As a first step, the customer offers a compensation model to the supplier which determines the payment to the supplier over the lifetime of the contract. The customer's compensation model incorporates both insurance and incentive, which are integral parts of agency models. For providing insurance, a fixed payment w is included in the model; to incentivize the supplier to improve the system (increase his performance), both positive and negative incentives are considered in the model. As per this policy, the customer pays its share ($0 \leq \alpha \leq 1$) of supplier's total cost TC (Scherer 1964; McAfee and McMillan 1986, Kim et al. 2007) and penalizes the supplier based on the average number of backorders.

$$P(w, \alpha, \beta) = w + \alpha(TC) - \beta E(B) \quad (4.1)$$

where α represents a percentage of supplier's total cost that is compensated for by the customer and parameter β denotes the penalty for each backorder.

4.1.1 Notation

N :	Total number of systems
λ :	Failure rate of each component
$1/\mu$:	Repair lead-time of each server
S :	spare parts inventory at the warehouse
π_x :	The steady-state probability of inventory level x at the warehouse
w :	Fixed payment
α :	Cost sharing rate
β :	Backorder penalty rate
C_S :	Base stock production cost
C_τ^L :	Reliability improvement linear cost
C_τ^E :	Reliability improvement exponential cost
C_μ :	Repair facility improvement cost
$\mu_{\min}$:	Minimum acceptable repair rate
$\tau_{\min}$:	Baseline reliability for the subsystem
$\tau_{\max}$:	Maximum achievable reliability
f :	The feasibility of increasing the subsystem reliability
$E(A)$:	Average operational availability
$A_{\min}$:	Minimum acceptable availability level
$E(B)$:	Average number of backorders

TC : Supplier's total cost

$P(w, \alpha, \beta)$: Suppliers' revenue from the customer

$\Pi^B(\lambda, \mu, S)$: Supplier's profit when it is penalized based on the backorders

$\psi(w, \alpha, \beta)$: Customer's total cost

4.1.2 Customer's problem

The contract is modeled using the principal-agent paradigm in which the principal specifies the terms of the contract to minimize the cost, subject to a defined service level, following which the supplier tries to maximize its profit based on the contract terms. The customer specifies the contract parameters (w, α, β) in which the penalty rate β is defined for each backorder that minimizes its cost subject to existing constraints.

$$\begin{aligned} \text{Min} \quad & \psi(w, \alpha, \beta) = P(w, \alpha, \beta) \\ \text{S.t} \quad & E[A(\lambda^*, \mu^*, S^*)] \geq A_{\min} \quad (AR) \end{aligned} \quad (4.2)$$

$$\Pi^B(\lambda^*, \mu^*, S^*) \geq 0 \quad (IR) \quad (4.3)$$

$$(\lambda^*, \mu^*, S^*) \in \arg \max \Pi^B(\lambda, \mu, S) \quad (IC) \quad (4.4)$$

The first constraint (AR) controls the average availability $E(A)$. As per this constraint, the availability of the system should be maintained higher than the minimum acceptable level $A_{\min}$. (IR) is known as the individual rationality constraint in the principal-agent models which ensures supplier's participation. The incentive compatibility constraint (IC) shows that the supplier always makes decisions to maximize its own profit.

4.1.3 Supplier's problem

Faced with the contract terms (w, α, β) , the supplier then chooses its own set of controllable variables (λ, μ, S) to maximize its expected profit:

$$\Pi^B(\lambda, \mu, S) = P(w, \alpha, \beta) - TC \quad (4.5)$$

The first term in Equation (4.5) denotes the supplier's revenue (from the customer) and the second the supplier's total cost (TC).

4.2 Probability Distribution of Inventory Level

A single echelon repairable part inventory system consisting of a repair facility and a single warehouse is considered, as in Chapter 3, to support the customer's operating fleet. To simplify the analysis, a repair facility with only one server is considered. In such a system, the supplier can improve efficiency of the repair facility by reducing repair time of the server instead of adding a new server. The system operates as per the following assumptions:

- (1) Component failure follows an independent Poisson distribution with constant rate, λ .
- (2) The systems' demand for each component (over the contract's lifetime) occurs according to a Poisson process with the variable rate $\lambda(z)$, where z is the operating fleet size.
- (3) The repair time at the server (in the repair facility), consisting of repair time and transportation delay, is an independent exponential random variable with mean $(\mu)^{-1}$.

- (4) A one-for-one base stock S replenishment policy is used; a failed component is immediately replaced by a ready-to-use component (new or refurbished) from the warehouse and the failed one is sent to the repair facility.

The transition diagram of the system with a constant service rate at the repair facility is shown in Figure(4.1) Markovian approach is applied to such an inventory system to determine the steady-state probability of inventory level x at the warehouse π_x , as described in Chapter 3. The flow balance equations are listed below:

Fig 4.1 The transition diagram of the inventory system with constant service rate

$$\pi_x(z\lambda + \mu) = (Z_N) \times \pi_{x-1} \times \mu + (Z_S) \times \pi_{x+1} ([Min(z+1, N)]\lambda) \quad \text{for } -N \leq x \leq S$$

$$\sum_{x=-N}^S \pi_x = 1 \quad (4.6)$$

By solving the balance equations and the normalizing constraint, the probability distribution of the inventory level can be specified as follows:

$$\pi_x = \frac{\left(\frac{\mu}{\lambda}\right)^x}{\zeta \left[\frac{N\lambda \left(1 - \left(\frac{\mu}{N\lambda}\right)^{S+1}\right)}{N!(N\lambda - \mu)} + \frac{e^{\frac{\mu}{\lambda}} \left(\frac{\lambda}{\mu}\right)^N \Gamma\left(N, \frac{\mu}{\lambda}\right)}{(N-1)!} \right]} \quad \zeta = \begin{cases} (N+x)! & x < 0 \\ N!N^x & x \geq 0 \end{cases} \quad x = -N \dots S \quad (4.7)$$

$$\Gamma\left(N, \frac{\mu}{\lambda}\right) = (N-1)! e^{-\frac{\mu}{\lambda}} \sum_{k=0}^{N-1} \frac{\left(\frac{\mu}{\lambda}\right)^k}{k!}$$

4.3 Incentive Mechanism

Penalizing the supplier based on the expected number of backorders is considered the most reasonable negative incentive mechanism in inventory models (Sherbrooke 1968). The average number of backorders can be calculated by checking the number of backorders at a random point in time. Given the steady-state probabilities, the average number of backorders can be computed by the expected value of x , when x is negative.

$$E(B) = \left| \sum_{x=-N}^{-1} x\pi_x \right| = \frac{\left[\left(\frac{\lambda}{\mu}\right)^N e^{\frac{\mu}{\lambda}} \Gamma\left(N, \frac{\mu}{\lambda}\right) (N\lambda - \mu) + \lambda \right] (N\lambda - \mu) N}{\left[\left(\frac{\lambda}{\mu}\right)^N e^{\frac{\mu}{\lambda}} \Gamma\left(N+1, \frac{\mu}{\lambda}\right) (N\lambda - \mu) - \left(\frac{\mu}{N\lambda}\right)^S \mu + \mu \right] \lambda} \quad (4.8)$$

As can be seen, the average number of backorders is a function of the ratio of the failure rate to service rate; let $\varphi = \lambda / \mu$. Next, the average number of backorders is expressed based on S and φ :

$$E(B) = \frac{\left(\varphi^N e^{\frac{1}{\varphi}} \Gamma\left(N, \frac{1}{\varphi}\right) \left(N - \frac{1}{\varphi} + 1\right) \left(N - \frac{1}{\varphi}\right) N \right)}{\varphi^N e^{\frac{1}{\varphi}} \Gamma\left(N + 1, \frac{1}{\varphi}\right) \left(N - \frac{1}{\varphi}\right) - \left(\frac{1}{\varphi N}\right)^S + \frac{1}{\varphi}} = \frac{\left(\varphi^N e^{\frac{1}{\varphi}} \Gamma\left(N + 1, \frac{1}{\varphi}\right) (N\varphi - 1) + 1 \right) \left(N - \frac{1}{\varphi}\right)}{\varphi^N e^{\frac{1}{\varphi}} \Gamma\left(N + 1, \frac{1}{\varphi}\right) (N\varphi - 1) + 1 - \left(\frac{1}{\varphi N}\right)^S} \quad (4.9)$$

To simplify the above, a parametric analysis is conducted to evaluate the impact of S and φ on the average number of backorders; see Appendix B for more details. It is observed that there are special values of φ (φ_1 and φ_2) for each fleet size where the average number of backorders can be calculated within the following ranges.

$$E(B) = \begin{cases} N - \frac{1}{\varphi} & \varphi_2 \leq \varphi \leq 1 \\ \left(N - \frac{1}{\varphi}\right) \left(\frac{a}{a-b}\right) & \varphi_1 < \varphi < \varphi_2 \\ 0 & 0 \leq \varphi \leq \varphi_1 \end{cases} \quad (4.10)$$

The average operational availability is used to measure the performance of the system; assuming each backordered component results in an un-operational system, the availability of the entire set of systems can be expressed as:

$$E(A) = 1 - E(B) / N \quad (4.11)$$

4.4 Supplier's Costs

The supplier, as the service provider, sets the parameters of the system (such as inventory level, reliability of the components and the repair rate) to deliver the performance level required. The supplier's total cost (TC) consists of subsystem design cost, production cost

and repair facility cost. In this section, two different scenarios for the supplier's cost function are considered. Firstly, a linear cost function is defined for the reliability improvement that results in supplier's profit to be a concave function of the base stock. In fact, a linear cost function for the reliability may not be appropriate for some components as their reliability cost function has an exponential pattern (and not linear). Therefore, in the second scenario two separate cost functions are formulated for the service capacity and component's reliability, based on common functions found in literature.

4.4.1 Base stock production cost

To speed up the maintenance and replacement of failed sub-systems, the supplier sets up a spare parts warehouse in close proximity to the field. For capital-intensive units with sporadic failure, one-for-one base stock policy is widely used (see Sherbrooke 1992). The inventory investment cost, denoted as C_S , is given as:

$$C_S = S \times c_s \quad (4.12)$$

where S and c_s represent base stock level and part unit cost respectively.

4.4.2 Service capacity cost

The supplier can improve the efficiency of the repair facility by increasing the repair rate. The following equation represents the cost associated with increasing the service capacity. It is assumed that this cost is linear in μ with a unit cost c_μ (Allon and Federgruen 2007, Kim et al. 2010).

$$C_\mu = c_\mu (\mu - \mu_{\min}) \quad \mu \geq \mu_{\min} \quad (4.13)$$

where $\mu_{\min}$ is the minimum acceptable repair rate.

4.4.3 Linear reliability cost function

The supplier's TC includes component design cost, production cost and repair facility cost. In the first scenario, it is assumed that all cost functions are linear in S , μ and τ ($1/\lambda$).

$$TC = C_S + C_\mu + C_\tau^L = Sc_S + (\mu - \mu_{\min})c_\mu + (\tau - \tau_{\min})c_\tau \quad (4.14)$$

where c_τ represents the design cost with the baseline reliability. Let c_φ be c_τ/μ . Next, the supplier's total cost can be calculated as follows:

$$TC = Sc_S + (\mu - \mu_{\min})c_\mu - \mu\tau_{\min}c_\varphi + \frac{1}{\varphi}c_\varphi \quad (4.15)$$

Supplier's response

Given the contract terms (w, α, β) , the supplier then sets (φ, S) to maximize his profit.

$$\begin{aligned} \text{Max} \quad & \Pi^B(S, \mu, \lambda) = w - (1 - \alpha)(Sc_S + (\mu - \mu_{\min})c_\mu - \mu\tau_{\min}c_\varphi + \frac{1}{\varphi}c_\varphi) - \beta(N - \frac{1}{\varphi}) \\ \text{S.t} \quad & \varphi_2 \leq \varphi \leq 1 \end{aligned} \quad (4.16)$$

In this range of φ , the supplier's decision for S and φ is as follows:

$$S = 0, \quad \varphi = \begin{cases} \varphi_2 & \beta > (1 - \alpha)c_\varphi \\ 1 & \text{Otherwise} \end{cases} \quad (4.17)$$

When $\varphi_2 \leq \varphi \leq 1$ the base stock does not affect the number of backorders. Therefore, the supplier keeps no base stock and can eliminate the associated cost. Also, the supplier has

motivation to improve the performance only when the penalty rate β is larger than the cost of such improvement $(1-\alpha)c_\varphi$. Offering appropriate penalty β and cost sharing α rates by the customer incentivizes the supplier to choose the lowest φ in this range ($\varphi_2 \leq \varphi \leq I$).

As a next step, the supplier evaluates the benefit of choosing a lower φ ($\varphi < \varphi_2$).

$$\begin{aligned}
 \text{Max} \quad & \Pi^B(S, \mu, \lambda) = w - (1-\alpha)(Sc_s + (\mu - \mu_{\min})c_\mu - \mu\tau_{\min}c_\varphi + \frac{1}{\varphi}c_\varphi) - \beta(N - \frac{1}{\varphi})\left(\frac{a}{a - (\frac{1}{N\varphi})^s}\right) \\
 \text{St} \quad & \varphi_1 < \varphi < \varphi_2
 \end{aligned} \tag{4.18}$$

From the experiments, it is observed that the supplier's benefit is highest around a special value of φ (I/N) regardless of the value of base stock S . To verify this, Table (4.1) shows how φ_1 and φ_2 converge to this value (I/N) as the fleet size is increased.

Table 4.1 φ_1 and φ_2 with increasing number of systems

N	φ_1	I/N	φ_2
50	0.015	0.02	0.023
100	0.007	0.01	0.011
500	0.0018	0.002	0.0022

Now, the optimal base stock level can be found by setting φ ($=I/N$). The supplier's objective function can be simplified as:

$$\text{Max} \quad \Pi^B(S, \mu, \lambda) = w - (1-\alpha)(Sc_s + (\mu - \mu_{\min})c_\mu - \mu\tau_{\min}c_\varphi + Nc_\varphi) - \beta\left(\frac{N^2}{(\frac{1}{N})^N Ne^N \Gamma(N+1, N) + SN}\right) \tag{4.19}$$

Proposition 1. A larger penalty rate β makes a stronger incentive for the supplier to decrease the number of backorders by increasing the base stock. Also, increasing the cost sharing rate

α by the customer motivates the supplier to increase the base stock as the customer reimburses a larger portion of the associated costs.

Proof: The supplier's profit is a concave function of S . Note that:

$$\frac{\partial \Pi^B}{\partial S} = -(1-\alpha)C_s + \frac{N^3 \beta}{\left[\left(\frac{1}{N}\right)^N N e^N \Gamma(N+1, N) + SN\right]^2} \quad \frac{\partial^2 \Pi^B}{\partial S^2} = \frac{-2\beta N^4}{\left[\left(\frac{1}{N}\right)^N N e^N \Gamma(N+1, N) + SN\right]^3} \leq 0$$

Given that $\partial \Pi / \partial S = 0$ at optimality (the supplier's profit is a concave function with respect to S), the optimal base stock is obtained next:

$$S^* = \sqrt{\frac{N\beta}{(1-\alpha)C_s}} - \left(\frac{1}{N}\right)^N e^N \Gamma(N+1, N) \quad (1) \quad \partial S^* / \partial \beta > 0 \quad (2) \quad \partial S^* / \partial \alpha > 0 \quad (4.20)$$

4.4.4 Exponential reliability cost function

The supplier's total cost consists of the cost of component reliability ($\tau = 1/\lambda$) and the cost for making service capacity available. For many components the associated cost of reliability improvement follows an exponential rather than a linear function. Therefore, in this section an exponential reliability cost function is incorporated to evaluate the supplier's decision. In this step, an exponential form is proposed for the cost of component reliability improvement (following Mettas and Kallenberg, 2000):

$$\bar{C}_\tau = e^{f\left(\frac{\tau - \tau_{\min}}{\tau_{\max} - \tau}\right)} \quad \tau_{\min} \leq \tau < \tau_{\max} \quad (4.21)$$

It is assumed that the baseline reliability for each subsystem τ_{min} does not incur any design cost. An upper bound is considered for subsystem reliability τ_{max} that is infinitely expensive for the supplier. Also, the parameter f is interpreted as the feasibility of increasing the subsystem reliability. Larger values of f represent higher difficulty and higher cost of increasing the subsystem reliability. Figure(4.2) illustrates the behavior of the reliability cost function with the initial reliability of 65%, the maximum achievable reliability of 95% and a feasibility of 0.2.

Fig 4.2 The exponential behavior of reliability cost function.

As can be seen, the cost function meets two following important requirements of reliability improvement:

- (1) A low reliability incurs very low cost.
- (2) Cost of meeting a high reliability level is very high, as the marginal cost increases exponentially.

Oner et al. (2010) develop a modified version of the reliability cost function of Mettas and Kallenberg (2000).

$$\bar{C}_\tau = K \left[\exp \left(f \frac{\tau - \tau_{\min}}{\tau_\infty - \tau} \right) - 1 \right] \quad \tau_{\min} \leq \tau \leq \tau_{\max} \quad (4.22)$$

where K is a positive parameter and τ_∞ is a reliability level higher than the maximum achievable reliability. Under this formulation, reaching the maximum achievable reliability is feasible with a finite associated cost. The standard regression analysis indicates that a simple exponential model fits this cost function quite well. Therefore, Equation (4.23) is used in the model as the reliability cost function:

$$C_\tau^E = \nu \times e^{\theta\tau} \quad \tau_{\min} \leq \tau < \tau_{\max} \quad (4.23)$$

Thus, the supplier's total cost (TC) can be expressed as

$$TC = C_s + C_\mu + C_\tau^E = Sc_s + (\mu - \mu_{\min})c_\mu + \nu \times e^{\theta\tau} \quad (4.24)$$

Supplier's response

Faced with the contract terms (w, α, β) , the supplier sets (τ, μ) to maximize his profit:

$$\begin{aligned} \text{Max} \quad \Pi^B(\tau, \mu) &= P(w, \alpha, \beta) - TC \\ &= w - (1 - \alpha)(Sc_s + c_\mu(\mu - \mu_{\min}) + \nu e^{\theta\tau}) - \beta(N - \mu\tau) \end{aligned} \quad (4.25)$$

Proposition 2. The suppliers' incentive to invest in reliability is higher when (1) the customer offers a higher cost sharing incentive α , and (2) the performance incentive β is higher.

Proof: The supplier's cost function is a concave function of τ . Note that:

$$\frac{\partial \Pi^B}{\partial \tau} = -(1-\alpha)\nu\theta e^{\theta\tau} + \beta\mu \quad \frac{\partial^2 \Pi^B}{\partial \tau^2} = -(1-\alpha)\nu\theta^2 e^{\theta\tau} \leq 0$$

$\partial \Pi^B / \partial \tau = 0$ at optimality (the supplier's cost function is concave with respect to τ), leads to

$$\tau^* = \frac{\ln \left[\frac{\beta\mu}{(1-\alpha)\nu\theta} \right]}{\theta} \quad (1) \frac{\partial \tau^*}{\partial \beta} \geq 0 \quad (2) \frac{\partial \tau^*}{\partial \alpha} \geq 0 \quad (4.26)$$

Proposition 3. The supplier is motivated to maintain the availability of the system (by improving the component reliability) higher than the minimum acceptable level ($A_{\min}$), only if the customer's contract terms (α, β) satisfy the following condition.

$$\frac{\beta}{1-\alpha} \geq \frac{\nu\theta}{\mu} e^{\frac{A_{\min} N \theta}{\mu}} \quad (4.27)$$

Proof: Given the contract terms (α, β), the supplier responds by choosing the optimal reliability level (τ^*). The average availability can be calculated as:

$$E(A) = 1 - \frac{(N - \mu\tau^*)}{N} = \frac{\mu}{N\theta} \ln \left[\frac{\beta\mu}{(1-\alpha)\nu\theta} \right] \quad (4.28)$$

In accordance with the availability requirement constraint (AR), the customer should offer contract terms which incentivize the supplier to maintain the average availability higher than the minimum acceptable level ($A_{\min}$). The ratio of penalty and cost sharing rates can be computed as:

$$\frac{\mu}{N\theta} \ln \left[\frac{\beta\mu}{(1-\alpha)\nu\theta} \right] \geq A_{\min} \quad \frac{\beta}{1-\alpha} \geq \frac{\nu\theta}{\mu} e^{\frac{A_{\min}N\theta}{\mu}} \quad (4.29)$$

One key result from Proposition 3 is that the contract terms (β, α) are an increasing function of the number of systems N , as more systems imply a greater equipment failure probability over the contract horizon. As a result, to avoid a large number of backorders, the customer has to provide higher positive α and negative β incentives. Also at the start of the contract the customer has to offer a minimum negative incentive β and positive motivation $(1-\alpha)$ to incentivize the supplier to provide the minimum acceptable performance ($A_{\min}$). This minimum incentive ratio ($I_m = \beta/(1-\alpha)$) can be derived from Equation (4.29) as follows:

$$\frac{\beta}{1-\alpha} \geq \frac{\nu\theta}{\mu} e^{\frac{A_{\min}N\theta}{\mu}} \quad I_m = \frac{\nu\theta}{\mu} e^{\frac{A_{\min}N\theta}{\mu}} \quad \frac{\partial I_m}{\partial N} \geq 0 \quad (4.30)$$

The incentive function (Equation 4.30) offers the following insights into the relationship between the fleet size and the incentive mechanism:

- (1) Minimum acceptable incentive ratio increases exponentially with the fleet size to be supported under the contract.
- (2) The required incentive ratio for a small group of systems is very low.
- (3) The required incentive ratio for a large group of systems is very high.

4.5 Numerical Example

In this section, a numerical analysis using the real data of Automatic Test Equipment (ATE) is carried out (Jin and Tian, 2012). ATE is a capital-intensive machine widely used to test wafers and micro-devices in semiconductor industry. The cost of each machine ranges from \$0.5-3 million depending on the functionality. The machine is designed as modules to

facilitate maintenance and repairs. Upon failure, the machine can be quickly restored to the production mode by replacing the failed module (i.e. component) with a spare one. The failed module is sent back to the repair centre for root-cause analysis. Once fixed it is stored at the local warehouse as refurbished module.

An ATE system typically consists of multiple types of modules each of which is connected in series. PBC is modeled for two types of modules in terms of fleet size and design costs, namely a high volume low cost (HVLC) and a low volume high cost (LVHC) module. In general, more HVLC modules are configured in an ATE system than LVHC modules. The related parameters are used to find the optimal component reliability and the minimum incentive ratio and are listed in Tables (4.2-4.3).

The reliability cost functions of two component types follow an exponential form as per the reliability growth model proposed in Equation (4.22).

Table 4.2. Parameter of HVLC components and contract terms

Parameter	Value	Unit	Parameter	Value	Unit
$\tau_{\min}$	30,000	hours	β	5,000	\$/backorder
$\tau_{\max}$	48,000	hours	N	100	-
τ_{∞}	50,000	hours	f	0.1	-
K	400,000	\$/design	α	0.2	-
μ	4	repair/month	$A_{\min}$	0.99	-

Table 4.3. Parameter of LVHC components and contract terms

Parameter	Value	Unit	Parameter	Value	Unit
$\tau_{\min}$	10,000	hours	β	2,500	\$/backorder
$\tau_{\max}$	20,000	hours	N	30	-
τ_{∞}	22,000	hours	f	0.1	-
K	800,000	\$/design	α	0.2	-
μ	2	repair/month	$A_{\min}$	0.97	-

$$C_{\tau}^{HVL C} = 400000 \left[\exp \left(0.1 \frac{\tau - 30000}{50000 - \tau} \right) - 1 \right] \quad 30000 \leq \tau \leq 48000 \quad (4.31)$$

$$C_{\tau}^{LVHC} = 800000 \left[\exp \left(0.1 \frac{\tau - 10000}{22000 - \tau} \right) - 1 \right] \quad 10000 \leq \tau \leq 20000 \quad (4.32)$$

The “ExponentialFit” (Maple’s Regression package, 2012) is used to fit an exponential function of the form $\nu e^{\theta \tau}$ to data points collected from Equations (4.31) and (4.32). The Regression package fits linear and nonlinear functions by performing least-squares regression that minimizes the sum of the residuals squared. The command runs in Maple 15 to generate 10,000 random data points and estimates parameters ν and θ . The results of regression are illustrated in Figure (4.3). Clearly, a simple exponential function with values (0.3506, 35.31) and (0.00029, 0.00048) for parameters ν and θ , respectively, fits the reliability cost functions shown in Equations (4.31) and (4.32).

Fig 4.3. Fitting an exponential reliability cost function

The coefficient of determination (R^2) is measured (Table 4.4) to show the accuracy of the results. A high value of coefficient of determination (>0.9) indicates that the estimated function fits the actual data points reasonably well.

Table 4.4. Coefficient of determination (R^2)

Component	Regression Sum of Squares (RSS)	Total Sum of Squares (SST)	R^2
HVLC	1960.3	25524.2	0.92
LVHC	1881.7	21008.9	0.91

Then, the optimal reliability of the component can be found as follows:

$$\tau_{HVLC}^* = \frac{\ln\left[\frac{\beta\mu}{(1-\alpha)\nu\theta}\right]}{\theta} = \frac{\ln\left[\frac{5000 \times 0.0055}{0.8 \times 0.3506 \times 0.00029}\right]}{0.00029} = 43900 \text{ (Hours)}$$

$$\tau_{LVHC}^* = \frac{\ln\left[\frac{\beta\mu}{(1-\alpha)\nu\theta}\right]}{\theta} = \frac{\ln\left[\frac{2500 \times 0.0027}{0.8 \times 35.31 \times 0.00048}\right]}{0.00048} = 13000 \text{ (Hours)}$$

A parametric analysis is conducted to evaluate the impact of the customer's contract terms (α, β) on the supplier's optimal reliability decision τ^* . As can be seen from Figure (4.4), increasing both backorder penalty and cost sharing rates incentivize the supplier to invest in reliability improvement. In fact, increasing backorder penalty rate β decreases the supplier's payment by the customer. To avoid the loss, the supplier tries to improve the component's reliability leading to a smaller number of backorders. The customer, on the other hand, may motivate the supplier to invest in reliability improvement by increasing the cost sharing rate α , thereby reimbursing a larger portion of costs associated with the reliability improvement.

(a) Reliability of HVLC

(b) Reliability of LVHC

Fig 4.4 Supplier's optimal reliability decision in response to the contract terms

Fig 4.5. Minimum incentive ratio vs. Fleet size

The fleet size may grow during the contract life-time as new products may be released into the market. Thus, the customer may need to adjust the terms of the contract (based on the fleet size) in order to motivate the supplier. Figure (4.5) illustrates how the minimum incentive ratio should be adjusted as the operating fleet size changes. The minimum incentive ratio increases exponentially as the operating fleet size grows.

4.6 Summary

Even in PBCs, the customer is unable to directly dictate the supplier's investment decisions. Therefore, designing effective incentive mechanisms that draw an appropriate response from the supplier becomes the key to successful implementation of performance-based contracts. An effective incentive is determined based on its sensitivity to the parameters that improve the system performance, such as failure rate and service rate. As failure rate is a function of the number of operational systems and the service capacity, the fleet size is a critical

parameter that affects the supplier's response under performance contract and should be considered in the design of performance incentives.

In this research, the main focus is on the interaction between reliability and decisions affecting service time, under various forms of performance incentives. An analytical model is developed to evaluate the role of fleet size on effectiveness of penalty and cost sharing rates as negative and positive incentives under a performance contract. The results show that the supplier's decision is directly influenced by the fleet size when the customer uses the average number of backorders as a negative incentive. The results also support the thesis that the performance incentive specification should take into account the number of systems to be supported. This finding gives fascinating insight into real-life situations where the fleet size may change over the contract horizon due to release of new products or market saturation. The changes in operating fleet size significantly influence the supplier's decision, and the customer must adjust the contract terms appropriately to avoid facing undesirable performance outcome.

In this chapter, the supplier is assumed to be risk neutral. Chapter 5 presents the design of an effective contract that captures the supplier's risk attitude under PBC.

Chapter 5

Moral Hazard Problem and Risk Sharing under PBC

One way to coordinate different parties is to enforce an incentive contract according to the principle-agent model. In principle-agent model, the agent makes an effort, which incurs cost for the agent, to meet the principle's desired performance. The principle as the other party in the contract compensates the agent based on achieved performance level. The moral hazard problem and conflict of interests are two major issues under a principle-agent model that can be solved by designing an appropriate incentive mechanism. The moral hazard problem occurs because of information asymmetry where the principle cannot observe the agent's effort level during the contract time. The second issue comes when both parties (principle and agent) look to maximize their own profit.

The classical agency model considers the tradeoff between incentive and insurance, wherein the principal compensates the agent for receiving the outcome specified in the contract. The value of a risky decision to the supplier may be different than the expected profit that is realized by the same decision. Some decisions may result in serious losses to the supplier because of the associated risk. The supplier is said to be risk-averse with respect to an option if he prefers a value of profit that it is less than the expected profit for that option. An extensive study has been made on the measures of risk aversion (Szego, 2004). Such measure is specified by defining a function, called the utility function. The exponential, power and expected mean-variance utility functions are widely used in research, economic analysis and practical experience because of their tractability.

Holmstrom (1979) takes into account the moral hazard problem and the supplier's unobservable action in a principle-agent relationship. His model describes the situation where

the supplier's individual action affects the probability distribution of performance outcome. He derives the necessary conditions to improve the performance when the supplier's action is unobservable. Lambert et al (1991) propose a framework that explains how a supplier values a compensation contract. They found that in the presence of a risk-averse supplier tying the compensation model to the supplier's wealth has a superior effect in comparison to only considering the cost of the supplier's effort. Hemmer et al (2000) study the effect of adding a stock option to an optimal compensation contract. Their results show that it is appropriate to consider a convex component in compensation model only if the supplier has a decreasing absolute risk aversion. They also use the first order approach justifying the principle-agent problem developed by Jewitt (1988). Chen and Miller (2009) conduct a numerical simulation to evaluate the performance of a piecewise-linear-threshold when the supplier's action is unobservable and dynamic. Their results support usage of a piecewise-linear-threshold only if the supplier's utility function has the power form.

The literature on performance contracts reveals a lack of understanding of how incentive contracts should be designed taking into account the supplier's risk attitude. Evaluating the impact of the supplier's utility function, which captures his risk attitude, on the reimbursement policy under performance contract is the focus of this research. An analytical model is described, using the basic models of Holmstrom (1979) and Hemmer et al (2000), and used to prove the superiority of considering a convex component in the optimal revenue function under PBC. Then, a numerical analysis is conducted incorporating common utility functions (exponential, power, mean-variance), from practical standpoint, and study their impact on the optimal revenue function that maximizes the outcome performance under PBC.

5.1 Model of Performance-based Contract

The principal (customer) operates identical systems (e.g., airplanes) with a fleet size of N , where each system contains major equipment (e.g., engine, avionic). Random breakdown of equipment occurs according to a Poisson process with a variable rate $\tau(n)$, a function of the number of deployed systems that are in operational condition n and component reliability τ . In capital-intensive industries, the failed equipment need to be repaired and cannot be scrapped because of their high cost and long lifetime. Therefore, after a breakdown the failed equipment needs to be restored to operational condition. To improve performance and decrease repair costs, the customer outsources all service activities to a supplier who is considered to be risk-averse in this model. At the beginning of the contract, the customer offers a compensation model which reimburses the supplier in accordance with the availability performance. The supplier then makes its own decisions to invest in inventory level, service capacity and component reliability, which are obviously affected by his risk attitude. In the following section, the repair and performance measurement processes are explained, the supplier's cost components is discussed and risk aversion model is described.

5.1.1 Notation

- N : Total number of systems
- $1/\mu$: Repair lead-time of each server
- S : spare parts inventory at the warehouse
- C_τ : Reliability improvement cost
- C_μ : Repair facility improvement cost
- $\mu_{\min}$: Minimum acceptable repair rate
- $\tau_{\min}$: Baseline reliability for the subsystem

- τ_{∞} : Maximum achievable reliability
- f : The feasibility of increasing the subsystem reliability
- A : The operational availability
- $A_{\min}$: Minimum acceptable availability level
- $E(B)$: Average number of backorders
- TC : Supplier's total cost
- τ : Component reliability
- T : Contract period
- h : Inventory holding cost per unit per time
- r : Average cost for repairing each failed subsystem
- C_P : Total production cost
- C_H : Total holding cost
- C_R : Repair activities cost
- $E(O)$: Average of on-hand inventories in the warehouse
- $E(R)$: Average number of repairs
- $R(A)$: Revenue function for the operational availability
- $U(\Pi^A)$: Supplier's utility function
- $\Pi^A(\lambda, \mu, S)$: Supplier's profit when the reward is based on availability.

5.2 Repair Process

A closed-loop inventory system consisting of a repair facility and a single warehouse supports the customer systems under a PBC, described in Chapter 3. The repair facility is modeled as an $M/M/1$ queue where the repair lead-time of server in the repair facility, consisting of the repair time and transportation delay, is assumed to be an independent

exponential random variable with mean $(\mu)^{-1}$. The supplier is allowed to expand the service capacity by means of extra investment in, for example, adoption of new technology. A one-for-one base stock $(S-1, S)$ replenishment policy is followed at the warehouse.

5.3 Average and Variance of Performance Outcome

The expected availability of the system is a commonly used performance measure under PBL. “Availability is a measure of the degree to which an item is in operable state and can be committed at the start of mission when the mission is called for at an unknown (random) point in time” (U.S. Department of Defense 2005). The availability of a system can be determined by the number of backorders that are observed at a random point in time once steady state is reached. Similar to Chapter 4, the inventory level’s probability distribution is derived as follows:

$$\pi_x = \frac{(\mu\tau)^x}{\zeta \left[\frac{\frac{N}{\tau} \left(1 - \left(\frac{\mu\tau}{N} \right)^{S+1} \right)}{N! \left(\frac{N}{\tau} - \mu \right)} + \frac{e^{\mu\tau} \left(\frac{1}{\mu\tau} \right)^N \Gamma(N, \mu\tau)}{(N-1)!} \right]} \quad \zeta = \begin{cases} (N+x)! & x < 0 \\ N!N^x & x \geq 0 \end{cases} \quad x = -N \dots S \quad (5.1)$$

Given the inventory level’s probability distribution, average number of backorders B and its variance can be computed.

$$E(B) = \left| \sum_{x=-N}^{-1} x\pi_x \right| \quad Var(B) = \sum_{x=-N}^{-1} \left[|x| - \left| \sum_{x=-N}^{-1} x\pi_x \right| \right]^2 \pi_x \quad (5.2)$$

Let A be a random variable that represents the operational availability. The mean and variance of A can be derived using average number of backorders and its variance as follows:

$$E(A) = E\left(1 - \frac{B}{N}\right) = 1 - \frac{E(B)}{N} \quad \text{Var}(A) = \text{Var}\left(1 - \frac{B}{N}\right) = \frac{\text{Var}(B)}{N^2} \quad (5.3)$$

Thus, the supplier can control the variation in the system performance (i.e. the availability of the system) by managing the following three parameters: (1) base stock at the warehouse, (2) component reliability, and (3) efficiency of the repair facility.

5.4 Risk-averse Supplier and Convex Compensation Model

In this section, a principle-agent relationship is studied to find the optimal compensation model in the presence of a risk-averse supplier. The basic of analysis are the principle-agent models developed by Holmstrom (1979) and Hemmer et al (2000). At the beginning of the contract period, the supplier takes an unobservable action a that can be a combination of all possible efforts (increasing the base stock, improving the components reliability or efficiency of the repair facility) incurring a disutility $TC(a)$. The supplier's efforts will lead to a random performance outcome A (e.g. operational availability) at the end of the contract's life-time which determines the supplier's revenue $R(A)$. The density function of the performance outcome corresponding to the supplier's efforts is $f(A, a)$ over the contract's life-time. Let $U(R, TC)$ represent the supplier's preferences defined as $U(R(A)) - TC(a)$. Now, the customer's problem can be defined as follows:

$$\text{Min} \int R(A) f(A, a) dA \quad (5.4)$$

$$s.t \quad \int [U(R(A)) - TC(a)] f(A, a) dA \geq H \quad (5.5)$$

$$a \in \arg \max \left(\int U(R(A)) f(A, a) dA - TC(a) \right) \quad (5.6)$$

The first constraint (5.5) guarantees the supplier's minimum utility (H) that ensures its participation in the contract; the second constraint (5.6) indicates that the supplier makes its own decisions to maximize its profit. For analytical tractability, Holmstrom (1979) modifies the standard model by replacing the constraint (5.6) with the first order condition (5.7).

$$\int U(R(A))f_a(A, a)dA = TC_a(a) \quad (5.7)$$

where $f_a(A, a)$ and $TC_a(a)$ are the first derivative of $f(A, a)$ and $TC(a)$, respectively with respect to the supplier's action a . He shows that Lagrangian optimization method yields the following characteristic for the modified model.

$$\frac{1}{U'(R(A))} = L_1 + L_2 \frac{f_a(A, a)}{f(A, a)} \quad (5.8)$$

where L_1 and L_2 are Lagrange multipliers for constraints (5.6) and (5.7), respectively.

Jewitt (1988) tries to ensure the optimal solution of the first order approach model can be valid for the original standard model also. He finds that only under certain conditions on the distribution function (listed next) are Holmstrom's (1979) findings for the modified model valid for the original standard model.

- 1) $F(A, a)$ is convex in a for each value of A
- 2) $\int Af(A, a)dA$ is concave in a
- 3) $f_a(A, a)/f(A, a)$ is concave in A for each value of a

Normal, Poisson, χ^2 and Gamma distributions are examples of density functions satisfying the above-mentioned conditions that are similar to operational availability density functions corresponding to the supplier's actions. This research focuses on large class utility functions (the negative exponential, the logarithmic, the quadratic, and the power utility functions) as they are well-known functions used in practice. A general form of large class utility function can be formulated as follows:

$$U(R(A)) = \frac{1-\omega}{\omega} \left[\frac{bR(A)}{1-\omega} + \varepsilon \right]^\omega \quad b, \omega, \varepsilon \geq 0 \quad (5.9)$$

Using Equations (5.8) and (5.9), the optimal revenue model can be derived as

$$R(A) = \frac{1-\omega}{b} \left(\left[bL_1 + bL_2 \frac{f_a(A,a)}{f(A,a)} \right]^{1/\omega} - \varepsilon \right) \quad (5.10)$$

As can be seen from Equation (5.10), there is a convex component $f_a(A,a) / f(A,a)$ in the optimal revenue function that determines the supplier's payment by the customer. An intuitive interpretation is that a convex revenue function ensures that the supplier's risk avoidance does not conflict with the potential benefits of taking an appropriate action to improve the outcome performance.

5.5 Supplier's Costs

The supplier's total cost consists of subsystem design cost, production cost, holding cost and repair activities cost.

5.5.1 Subsystem design cost

It is assumed that baseline reliability for a subsystem τ_{min} does not incur any design cost. Also, an upper bound of subsystem reliability τ_{max} is assumed to be infinitely expensive for the supplier. Based on these assumptions, the following design cost function is used for a subsystem (see Mettas and Kallenberg 2000; Huang et al. 2007 and Oner et al. 2010 for more details):

$$C_{\tau} = K_{\tau} \left[\exp \left(f \frac{\tau - \tau_{min}}{\tau_{\infty} - \tau} \right) - 1 \right] \quad \tau_{min} \leq \tau \leq \tau_{max} \quad (5.11)$$

where K_{τ} is a positive parameter, τ_{∞} is the maximum achievable reliability and f is interpreted as the feasibility of increasing the subsystem reliability; its value can vary between 0 and 1. Larger values of f represent higher difficulty and cost in increasing the subsystem reliability.

5.5.2 Production cost

The production cost of each unit has fixed K_p and variable components. The variable cost depends on the desired reliability level, where producing a subsystem with higher reliability incurs higher cost (Huang et al. 2007 and Oner et al. 2010). The unit production cost function can be formulated as:

$$c_p(\tau) = K_p + \bar{K}_p(\tau - \tau_{min}) \quad \tau_{min} \leq \tau \leq \tau_{max} \quad \bar{K}_p > 0 \quad (5.12)$$

Then, the total production cost (C_p), for the base stock level S , is given by $S \times c_p$.

5.5.3 Holding cost

Given the steady-state probability of inventory level x , the average on-hand inventory in the warehouse, $E(O)$ is simply the expected value of x .

$$E(O) = \sum_{x=1}^S x\pi_x \quad (5.13)$$

Let h be the inventory holding cost per unit per time incurred by the supplier at the warehouse. Then, the average inventory holding cost, C_H , is $h \times E(O)$.

5.5.4 Repair activities cost

When a subsystem fails, the supplier transports it to the repair facility, repairs it and then transports it to the warehouse. Thus, the supplier incurs costs for each failure, which includes transportation and repair activities costs (for using resources). The average number of repairs, $E(R)$, during the lifetime of a contract (T), can be calculated based on $MTTR_e$.

$$E(R) = T / MTTR_e \quad (5.14)$$

If r denotes the average cost for repairing each failed subsystem, total repair activities cost (C_R) is $r \times E(R)$.

5.5.5 Service capacity cost

The supplier can improve the efficiency of the repair facility by increasing the repair rate. It is assumed that service capacity cost is linear in μ with a unit cost c_μ (Allon and Federgruen 2007, Kim et al. 2010) and can be expressed as:

$$C_\mu = c_\mu (\mu - \mu_{\min}) \quad \mu \geq \mu_{\min} \quad (5.15)$$

where μ_{min} can be interpreted as the minimum acceptable repair rate. Now, the total cost (TC) for the supplier can be expressed as:

$$TC = C_{\tau}(\tau) + C_{\mu}(\mu) + C_P(\tau, S) + C_H(\tau, \mu, S) + C_R(\tau, \mu, S) \quad (5.16)$$

5.6 Reimbursement Policy

The customer specifies the compensation policy that determines the payment to the supplier at the end of contract. The compensation model takes into account both insurance and incentive which are integral parts of agency models. In PBL contracts, the supplier's actions over the contract's lifetime determine its reward structure. In this research, a revenue function corresponding to operational availability is used (Nowicki et al. 2008). In this section, different types of reward functions, namely linear, exponential and piecewise-linear-threshold, are used.

The linear compensation function has a form $b + c \times X$, where b represents a fixed payment and the reward function increases proportional to the outcome performance X , with a specified slope c , and is illustrated in Figure (5.1).

$$R_{lin}(A) = \begin{cases} 0 & \text{if } 0 \leq A < A_{min} \\ b + c(A - A_{min}) & \text{if } A_{min} \leq A \leq 1 \end{cases} \quad (5.17)$$

Availability is a preferred performance outcome that is used to motivate the supplier for improving the base stock, service time and reliability level. A minimum acceptable availability level A_{min} is considered to prevent the supplier from delivering an unacceptable performance.

Fig 5.1 Linear compensation function

Another form of compensation function commonly used is exponential, expressed as:

$$R_{\text{exp}}(A) = \begin{cases} 0 & \text{if } 0 \leq A < A_{\min} \\ e^{\alpha + \beta A} & \text{if } A_{\min} \leq A \leq 1 \end{cases} \quad (5.18)$$

Exponential compensation function is illustrated in Figure (5.2), and models a relationship in which the increase in reward for availability improvement is higher closer to the maximum achievable availability level. Such behavior incentivizes the supplier to improve the availability level instead of staying around the minimum acceptable availability.

Fig 5.2. Exponential compensation function

The third reward function used in this research is piecewise-linear-threshold. Here, a_i denotes the fixed payment of the contract providing the supplier insurance and the slope b_i represents incremental reward according to the availability improvement. Similar to the linear and exponential models a quota A_{min} and a cap A_{max} are considered in the piecewise-linear-threshold function.

$$R_{step}(A) = \begin{cases} 0 & \text{if } 0 \leq A < A_{min} \\ a_1 + b_1(A - A_{min}) & \text{if } A_{min} \leq A < A_{penalty} \\ a_2 + b_2(A - A_{penalty}) & \text{if } A_{penalty} \leq A < A_{accept} \\ a_3 + b_3(A - A_{accept}) & \text{if } A_{accept} \leq A \leq A_{max} \end{cases} \quad (5.19)$$

Such reward function can be divided into three zones: penalty, dead and reward zones (Figure 5.3). The idea of the penalty area is to penalize the supplier for achieving an availability level less than the acceptable level. Also, the original definition of the dead zone provides similar award for all availability in the specified range (Brown and Burke 2000). Here, a modified step-wise model from the original form is defined. As can be seen, the

mathematical formula is same for all the zones differentiated by an incremental slope that provides a strong motivation for the supplier to reach an ideal performance.

Fig 5.3. Piecewise-linear-threshold compensation function

5.7 Supplier's risk attitude

Under a PBC, the customer offers the compensation modeled as linear, exponential or step-wise. The supplier then chooses its own set of controllable variables to maximize the expected profit. The supplier is assumed to be risk averse. In this section, the utility theory application from a contract perspective is considered to capture the parties' risk attitude. The utility function may be different for a supplier depending on the risk attitude. In this section, exponential, power and expected mean-variance utility functions are introduced that are widely used in research, economic analysis and practice (Kim et al. 2007, Chen and Miller 2009).

$$U_{\text{exp}} = -e^{-\gamma(\Pi^A)} U_{\text{pow}} = \frac{(\Pi^A)^\delta}{\delta} U_{m-v} = E(\Pi^A) - k \text{var}(\Pi^A) / 2 \quad (5.20)$$

where Π^A represents the supplier's profit based on achieved availability level and parameters γ, δ and k indicate the degree of risk aversion. There are several methods to determine the approximation value of risk aversion degree. For example, imagine an alternative with a fifty-fifty chance of yielding k_0 or lose of $k_0/2$. There is a value of k_0 for which the decision maker would be indifferent between accepting and rejecting the alternative. That value of k_0 approximately indicates the degree of risk aversion for a particular decision maker (Kirkwood 2002).

5.8 Supplier's problem

The customer chooses a compensation policy $R_i(A)$ where the supplier's reimbursement is in accordance with the availability provided A which is a function of base stock S , service capacity μ and component reliability τ . To reach a certain level of availability, the supplier applies effort to improve the system parameters (S, μ, τ) which cause a disutility of $TC(S, \mu, \tau)$ in cost terms. The supplier has a $U_i(\Pi^A)$ utility function, and the profit (Π^A) is calculated as:

$$\Pi^A = R_i(A) - TC(S, \mu, \tau) \quad (5.21)$$

At the beginning of contract period, the supplier sets an inventory level, service capacity and a reliability level for equipment so as to maximize the expected value of his utility function.

$$U_i[R_i(A) - TC(S, \mu, \tau)] \quad (5.22)$$

5.9 Numerical example

In this section, a numerical experiment is conducted using healthcare systems data (Oner et al. 2010) to examine common utility functions, with the objective of finding the optimal revenue function that maximizes the outcome performance under PBC. The cost parameters are tabulated in Table (5.1). Risk parameters of Raju and Srinivasan (1996) are followed which are applicable to a wide range of supplier's utility characteristics.

Table 5.1 Component cost parameters

Parameter	f	μ_{min}	τ_{min}	τ_{max}	K_r	K_p	$\bar{K}_p$	C_μ	h	r
Value	0.5	1	24	360	200,000	1000	10	10,000	20	600

Two scenarios are defined that differ in fixed factors such as the fleet size, contract period, constant risk aversions and revenue models.

Table 5.2 Values of factors

Parameter	N	T	γ	δ	k	R
Case 1	100	60	1/400	2/3	1/50	a
Case 2	60	30	1/333.3	3/4	1/25	b

The specific values of the three revenue functions (linear, exponential and step-wise) are taken from the literature (Nowicki et al. 2008).

$$R_{exp}(A) = \begin{cases} 0 & \text{if } 0 \leq A < 0.6 \\ e^{7.49+4.02A} & \text{if } 0.6 \leq A < 1 \end{cases} \quad R_{lin}(A) = \begin{cases} 0 & \text{if } 0 \leq A < 0.6 \\ 20000 + 200000(A - 0.6) & \text{if } 0.6 \leq A < 1 \end{cases}$$

$$R_{step}(A) = \begin{cases} 0 & \text{if } 0 \leq A < 0.6 \\ 20000 + 100000(A - 0.6) & \text{if } 0.6 \leq A < 0.8 \\ 40000 + 267000(A - 0.8) & \text{if } 0.8 \leq A < 0.95 \\ 80000 + 400000(A - 0.95) & \text{if } 0.95 \leq A \leq 1 \end{cases}$$

Fig 5.4 Linear, exponential and step revenue function for the first scenario

$$R_{lin}(A) = \begin{cases} 0 & \text{if } 0 \leq A < 0.7 \\ 10000 + 241667(A - 0.7) & \text{if } 0.7 \leq A \leq 1 \end{cases} \quad R_{exp}(A) = \begin{cases} 0 & \text{if } 0 \leq A < 0.7 \\ e^{4.29 + 7.031A} & \text{if } 0.7 \leq A \leq 1 \end{cases}$$

$$R_{step}(A) = \begin{cases} 0 & \text{if } 0 \leq A < 0.7 \\ 10000 + 100000(A - 0.7) & \text{if } 0.7 \leq A < 0.8 \\ 20000 + 200000(A - 0.8) & \text{if } 0.8 \leq A < 0.9 \\ 40000 + 425000(A - 0.9) & \text{if } 0.9 \leq A \leq 1 \end{cases}$$

Fig 5.5 Linear, exponential and step revenue function for the second scenario

Tables (5.3) and (5.4) show how each revenue model influences the supplier's mix decisions of base stock, service capacity and component reliability in accordance with the supplier's risk attitude. The numerical results provide evidence that either exponential or step-wise contracts should be considered instead of a linear contract when the supplier's utility function has either exponential or power form. Also, the findings support the superiority of the linear and step-wise contracts to exponential when the supplier's utility function has mean-variance form.

Table 5.3 The supplier's response (case 1)

(S, μ, τ, A) (I^A, ψ)		Revenue Model		
		Linear	Exponential	Step
Utility Function	Exponential	(0, 1.21, 89.69, 0.947) (19868,89400)	(2, 1.28, 93.06, 0.976) (13042,90538)	(3, 1.33, 89.74, 0.98) (13020,92000)
	Power	(0, 1.27, 85.93, 0.949) (19809,89800)	(2, 1.35, 88.37, 0.976) (13079,90538)	(3, 1.37, 87.66, 0.981) (12971,92400)
	Mean-Variance	(2, 1.39, 90.58, 0.982) (16283,97800)	(0, 1.18, 86.16, 0.93) (8512,75252)	(4, 1.41, 92.47, 0.991) (11234,96400)

Table 5.4 The supplier's response (case 2)

(S, μ, τ, A) (I^A, ψ)		Revenue Model		
		Linear	Exponential	Step
Utility Function	Exponential	(2, 1.3, 56.42, 0.969) (37749,75008)	(4, 1.32, 61.01, 0.989) (34127,76394)	(4, 1.36, 60.12, 0.99) (35586,78250)
	Power	(3, 1.23, 58.14, 0.971) (37630,75492)	(4, 1.37, 60.31, 0.99) (34358,76933)	(5, 1.38, 59.02, 0.992) (35365,79100)
	Mean-Variance	(3, 1.42, 61.12, 0.99) (37188,80083)	(0, 1.1, 61.12, 0.935) (19613,52260)	(6, 1.51, 61.07, 0.998) (32877,81650)

The availability performance in repairable part inventory system is directly dependent on the base stock level, service time and component's failure rate. When the supplier's risk attitude is captured by exponential or power function, his efforts to increase the base stock level or improve the service capacity and component reliability will be much lower when a linear contract is offered. A linear contract cannot motivate the supplier to enhance the availability performance as the supplier's risk utility prevents spending more on increasing the base stock level or improving the service capacity and component's reliability.

a) Exponential utility function for supplier

b) Power utility function for supplier

c) Mean-Variance utility function for supplier

Fig 5.6 The impact of the designed revenue function on the availability

As is shown in section (5.4), to control the problem of moral hazard during the life-time of PBC, a convex component should be included in the revenue function when the supplier has a large class utility function. The results in Table (5.3) and (5.4) also support this finding where both exponential and step-wise revenue models lead to a higher availability performance than a linear model in the presence of either exponential or power utility functions. In such a case, a moderate level of risk aversion ensures that the supplier's risk avoidance does not preclude the potential profit.

The results also indicate that the exponential revenue function is inferior to both linear and step-wise revenue models if the supplier's risk attitude follows a mean-variance utility function. Intuitively, a supplier with a mean-variance risk function is sensitive to the variation of its profit, implying that the supplier's risk avoidance increases when the variation of profit increases. Indeed, an exponential reimbursement function has a higher variance in

comparison to a linear and step-wise reimbursement model. Therefore, using an exponential reimbursement function discourages the supplier from increasing the availability (increasing the base stock level, improving the reliability and repair facility) because of increasing risk avoidance. In such a situation, the supplier is not motivated to spend more to achieve higher availability despite higher profits.

5.10 Sensitivity analysis

In this section a sensitivity analysis is performed to evaluate the sensitivity of the optimal availability to the cost parameters. New cases are generated by modifying each cost parameter by $\pm 20\%$ and $\pm 50\%$ and average availability of each modification level $\bar{A}_m$ is calculated. Then the average availability of the main example $\bar{A}$ and the modified instance $\bar{A}_m$ are compared by using the following function.

$$\Delta A = \frac{\bar{A}_m - \bar{A}}{\bar{A}}$$

Table 5.5 Sensitivity of the optimal availability to the inventory holding cost

Utility	Exponential			Power			Mean-Variance			ΔA
Revenue	Linear	Exp	Step	Linear	Exp	Step	Linear	Exp	Step	
50	0.941	0.974	0.974	0.94	0.972	0.973	0.974	0.927	0.988	-0.00562
20	0.947	0.975	0.974	0.949	0.973	0.975	0.975	0.928	0.988	-0.00321
-20	0.956	0.976	0.986	0.952	0.983	0.984	0.986	0.93	0.992	0.003788
-50	0.957	0.98	0.986	0.952	0.983	0.988	0.988	0.936	0.994	0.005969

As an example, the analysis results for inventory holding cost are listed in Table (5.5) where ΔA indicates the average of the percent differences in the optimal availability. Clearly, all deviations in the optimal availability are negligible (less than 0.6%) with respect to inventory

holding cost changes, implying that the results are robust against error in supplier's inventory holding cost estimation. The results of sensitivity analysis with respect to other supplier's cost parameters are presented in Appendix C.

5.11 Summary

In this chapter, the problem of contracting for after-sale service is studied by capturing the moral hazard problem. Moral hazard is created when one of the parties in the contract performs an action that is unobservable to the other. In this study, a risk-averse supplier is assumed whose efforts to enhance the operational availability (such as increasing base stock, improving the component reliability or the repair facility) are not visible to the customer. The customer should design an appropriate contract to control the moral hazard problem because of invisibility of the supplier's response to the contract terms. In this chapter the design of the incentive contract based on the supplier's risk attitude is explored.

Firstly, different forms of revenue function are defined in presence of various types of utility functions which can capture supplier's risk attitude. Then, an illustrative experiment, using healthcare systems data, is conducted to find the optimal reimbursement policy with respect to the supplier's risk attitude. The findings indicate that a linear incentive contract performs poorly when the supplier follows a large class utility function. On the other hand, the customer should avoid proposing an exponential compensation model if the supplier has a mean-variance risk attitude.

Chapter 6

Conclusions and Future Research

Maintenance, repair and overhaul (MRO) business represents 8-10% of the gross domestic product (GDP) in the U.S.; total U.S. GDP was \$15.094 trillion in 2011. These after-sales services span a variety of industries including semiconductor, automobiles, energy, airlines, network computing, healthcare, and defense systems. The annual operation and sustainment cost for the U.S. military equipment alone is \$63 billion, and the maintenance is supported by 678,000 DoD personnel along with hundreds of private contractors (Smith 2007). According to Lucintel (2012), the global wind operation and maintenance market is expected to reach \$10.6 billion in 2016 with annual growth rate of 16.6%. Depending on the complexity of capital equipment, MRO services are subcontracted to the OEM or the 3rd party logistics (3PL) providers. The OEM undertakes the after-sales maintenance service due to the complexity of the technology. Typical examples include aircraft engines, wind turbines, computer servers, and semiconductor equipment.

Two types of contracting mechanisms have been discussed in literature: material-based contract (MBC) and PBC. In MBC, the OEM is compensated for the labor and materials each time the service task is accomplished. Because of the simplicity in implementation, MBC has been widely used in private and public sectors for sustaining equipment operation post installation. The main concern of MBC is that the OEM has less motivation to improve the product reliability. This is because repair and maintenance bring a lucrative revenue stream to the supplier in the after-sales market.

Over the last decade, PBC has emerged as a new service model reshaping the acquisition, operation, and maintenance of capital equipment. A major distinction between

PBC and MBC is that the service provider under the PBC agreement is compensated based on the system performance, not on the actual labor hours and spare parts used. Typical performance measures include system availability, equipment downtime, mean time between failures, and average parts backorders. In the long run, PBC motivates the OEM to improve the reliability during the design, as it is paid off by reduced failures and repair costs post installation. In this chapter, a summary of findings and some suggestions for future work are given.

6.1 Summary

Over the past few years, traditional contracts in after-sales service are giving way to a new approach termed performance based logistics (PBL), in the capital-intensive industries in both commercial and military fields. While the traditional after-sales contracts in capital-intensive industries, such Fixed-price and Cost-plus, are based on buying spare parts, repair action, equipment, transactional goods and services, PBL focus on buying the outcome performance.

Many researchers have studied the PBL programs to find the barriers and enablers, appropriate frameworks and structure for their successful implementation. These studies discover that the performance, metrics and incentives are the most important enablers of PBL, and that enhancing them can lead to its successful implementation. Performance, the key goal of PBL, should be defined in terms of the system objectives, such as operational availability. Metrics are at the heart of PBL, as the performance is linked to the metrics specified in the contract. Incentives (monetary/non-monetary, positive/negative), on the other hand, motivate the suppliers to meet the customer's performance objectives.

While performance based logistics is reshaping the after-sales services in capital-intensive industries, such industries use a repairable parts inventory approach to repair the

failed parts instead of scrapping them, as the systems are expensive and subsystems have long lifetime. Therefore, the repairable parts inventory models literature is reviewed in this research in Chapter 2. The literature review indicates that METRIC, MOD-METRIC and VARI-METRIC models are the fundamental models of the repairable inventory systems which have later been enhanced by relaxing one or more of their restrictive assumptions. Reviewing the repairable part inventory models under PBL clarifies that the PBL approach aims to consider the system in the decentralized framework while satisfying both the customer and the supplier's requirements, simultaneously. Moreover, the literature shows that the PBL models still have restrictive assumptions which decrease supplier's flexibility and may give indication of unrealistic performance to the customers.

Application of the Markovian approach and consideration of the system as a closed-network queue can be a useful way to relax the restrictive assumptions of earlier researchers and obtain more accurate results for the repairable parts inventory systems. In Chapter 3, an inventory model is developed for a repairable parts system, operating under a PBL contract. The inventory system is modeled as an $M/M/m$ queue in which component failures are Poisson distributed and the repair times at the service facility are Exponential. A variable rate is considered for the component failure, which is dependent on the number of operational systems, and the components' replacement rate from repair facility to the warehouse depends on the number of components in repair facility and the number of servers.

An example based on real data from defense sector is used to illustrate the findings. The model is then analyzed to determine the factors that have significant impact on each measure of system performance. The influence of component reliability, service capacity and base stock level are examined on average number of backorders and average down time as two applicable metrics and some common performance measures such as fill rate, ready rate and availability. The results show that the base stock level does not have any significant

effect on the average number of backorders, in comparison to other parameters, such as repair rate and repair capacity and increasing it may result in penalty for the supplier. Furthermore, the model recommends concentrating on the component reliability and repair system efficiency to improve the performance of the system with repairable spare parts.

In PBCs, the customer cannot directly dictate the supplier's investment decisions. Therefore, designing effective incentive mechanisms that draw an appropriate response from the supplier becomes the key to successful implementation of performance-based contracts. An effective incentive is determined as per the sensitivity to the parameters that improve the system performance. According to findings in Chapter 3, to achieve a desired performance level, the supplier has to improve the component reliability and the repair time. As reliability is a function of the number of operational systems, the supported fleet size is a critical parameter that should be considered in the design of performance incentives.

Chapter 4 focuses on the interaction between reliability and service time decisions under various performance incentives. An analytical model is developed to evaluate the role of operating fleet size on effectiveness of penalty and cost sharing rates as negative and positive incentives under a performance contract. The results support the thesis that the performance incentive specification should take into account the fleet size to be supported. This finding gives a fascinating insight as the fleet size may change over the contract horizon when new products are released or market reaches saturation. As the operating fleet size changes significantly influence the supplier's decision, the customer must adjust the contract terms to avoid undesirable performance outcome. Finally, a numerical analysis is conducted using real data from an Automatic Test Equipment (ATE) scenario. This example is used to find the optimal base stock and components reliability for both types of cost functions (linear and exponential).

Moral hazard arises when one of the parties in the contract performs an action unobservable to the other. In Chapter 5, a risk-averse supplier is assumed where its efforts for enhancing the operational availability (such as increasing base stock, improving the component reliability or the repair facility) are not visible to the customer. The customer should design an appropriate contract to contain the moral hazard problem because of invisibility of the supplier's response to the contract terms. Firstly, different forms of revenue functions are defined in presence of various types of utility functions which capture the supplier's risk attitude. Then, an illustrative experiment using healthcare systems data is conducted to find the optimal reimbursement policy with respect to the supplier's risk attitude. The findings indicate that the linear incentive contract performs poorly when the supplier's utility follows a large class function. On the other hand, the customer should avoid proposing an exponential compensation model if the supplier has a mean-variance risk attitude.

6.2 Future Work

In this research, some issues arising in inventory models operating under PBL contract have been analyzed. Thereafter, several suggestions have been discussed to enhance the integral enablers of PBL which lead to its successful implementation. However, many unanswered questions remain that warrant investigation. The following are the areas that have potential for future research.

6.2.1 Multiple-Phase and Multi-Level Performance Contracting

A singleechelon repairable inventory system is modeled under performance contracting in this research. Quite a few systems, especially in airline industry, use a multi-echelon repairable inventory system, where the system has many bases to support the operational

items and several depots to serve the bases in case of backorders. In this case, the supplier's decision for locating and allocating the resources among the bases and depots makes the contract more complex. Finding the contract terms when the customer uses a multi-echelon system can be studied as a future research.

Considering one-supplier supply chain structure was a simplifying assumption in this research. In some cases, the customer has to contract with multiple suppliers to support the system with multiple subsystems. This increases the complexity of the model where the customer penalizes the suppliers in accordance with the system downtime not the subsystems downtime. Then, developing a multi-supplier structure and finding a mechanism to incentivize each individual supplier in PBL models can be considered as a future work.

When a new design is introduced to the market, the product reliability is far below the target. The product reliability increases over time through corrective actions or redesign. As a customer who has purchased the early version of the products usually does not expect the system MTBF (or the availability) reach a high level. Rather, the customer expects to see the growth of the MTBF as well as the improvement of the operational availability. Therefore, a more realistic PBC shall be designed as a multi-phase and multi-level service agreement, instead of a stationary performance agreement.

6.2.2 Performance Incentives Trade-off

In PBCs, the customer cannot directly dictate the supplier's investment decisions. Therefore, designing effective incentive mechanisms that draw an appropriate response from the supplier becomes the key to successful implementation of performance-based contracts.

Under the performance contract, the supplier's decision is a function of the ratio of the penalty rate to the cost sharing rate. Also, the individual rationality constraint in the principal-agent models ensures the supplier's participation. Thus, the customer has to find an

appropriate balance between the positive and negative incentives that help to achieve the desired performance and motivate the supplier at the same time. Study of the trade-off between penalty rate and cost sharing rate is another direction for further research.

6.2.3 Performance Control Measures

Under PBL, the system needs to be modeled in a way that facilitates computation of some effective metrics and allows the customer to monitor the key parameters affecting system availability. This allows the parties to ensure that the delivered performance at the end of the contract meets or exceeds the target. In this research *MTBF* is defined for controlling the component reliability and *MTTR_e* for monitoring the efficiency of repair facility. The *MTBF* measure based on current model is more accurate than the previous models, as it considers the dependency of subsystem failure rate to the number of operational units. Also, the proposed model measures the *MTTR_e*, instead of *MTTR*, which considers the repair time and transportation delay, simultaneously. It is easier for the customer to monitor the warehouse replacement rate rather than the repair rate. This issue becomes more important under PBL contracts where the supplier may be more hesitant in offering detailed information such as repair rate during the contract's lifetime. Introducing more control measures to track the performance of a PBL contract over time is another area for future research.

Appendix A

Matlab code for calculating the control measures and the performance is as follow:

```
%-----Input-----
F = input('Failure Rate:');
R = input('Repair Rate:');
m = input('Server No:');
S = input('Base stock:');
N = input('System No:');
%-----Steady State Probability-----
Pro=0;
x=-N;
A= zeroes(N+S+1 , N+S+1);
A(N+S+1,N+S)= min(1,m)*R;
A(N+S+1,N+S+1)= -N*F;
fori=1:N+S+1
A(1,i)=1;
end
fori=2:N+S
x=x+1;
y=S-x;
if x>=0
z=N;
else
z=N+x;
end
A(i,i-1)= min(y+1,m)* R;
A(i,i)=-(z*F + min(y,m)*R);
A(i,i+1)= min(z+1,N)*F;
end
B= zeros(N+S+1,1);
B(1,1)=1;
P = zeros(N+S+1,1);
P = (A^-1)*B;
fori=1:N+S+1
Pro=Pro+P(i,1);
end
Pro;
%-----Metrics-----
p1=0;
j=N+1;
EI=0;
fori=0:S
EI=EI+i*P(j,1);%=====The average inventory in the warehouse=====
p1=p1+P(j,1);
j=j+1;
end
p2=0;
j=1;
EB=0;
fori=-N:-1
```


Appendix

```
EB=EB+i*P(j,1);%=====The average number of backorder=====

p2=p2+P(j,1);
j=j+1;
end
MTBF2=0;
j=2;
fori=-N+1:-1
    MTBF2=MTBF2+(P(j,1)/(F*(N+i)));
    j=j+1;
end
j=2;
EF1=0;
fori=-N+1:-1
    EF1=EF1+ P(j,1)*(F*(N+i));
    j=j+1;
end
j=N+1;
EF2=0;
fori=0:S
    EF2=EF2+F*N*P(j,1);
    j=j+1;
end
mttr=0;
j=1;
fori=-N:S-1
    mttr=mttr+(P(j,1)/(R*min(m,S-i)));
    j=j+1;
end
EX=0;
j=1;
fori=-N:S
    EX=EX+P(j,1)*i;
    j=j+1;
end
EF=EF1+EF2;
EB=-EB;
MTBF1=(1/(N*F))*p1;
EY=S-EX;
WT=(EY/EF)*EB;%=====The average waiting time=====

%-----Control Measures and Performance-----
formatshortg
MTBF=(MTBF1+MTBF2)*8760;
MTTR=mttr*8760;
AV=1-(EB/N)%==The availability of system==
RR=p1%==Ready Rate==
FR=p1-P(N+1,1) %==Fill Rate=
```

Computational results of the analysis for availability performance

The following figures illustrate the computational results of the analysis for availability performance

(a) Impact of Repair Rate on Backorders

(b) Impact of Base Stock on Backorders

(c) Impact of Failure Rate on Backorders

(d) Impact of Server Number on Backorders

Fig A.1 Impact of System parameters on Backorders

(a) Impact of Repair Rate on Waiting Time

(b) Impact of Base Stock on Waiting Time

(c) Impact of Failure Rate on Waiting Time

(d) Impact of Number of Server on Waiting Time

Fig A.2 Impact of System parameters on Waiting Time

(a) Impact of Repair Rate on Fill Rate

(b) Impact of Base Stock on Fill Rate

(c) Impact of Failure Rate on Fill Rate

(d) Impact of Server Number on Fill Rate

Fig A.3 Impact of System parameters on Fill Rate

(a) Impact of Repair Rate on Ready Rate

(b) Impact of Base Stock on Ready Rate

(c) Impact of Failure Rate on Ready Rate

(d) Impact of Server Number on Ready Rate

Fig A.4 Impact of System parameters on Ready Rate

As can be seen in the figures, decreasing the failure rate, increasing the server repair rate and the number of servers improve all metrics (average number of backorders and waiting time) and performance measures of the system (Fill rate, Ready rate). Results also indicate that increasing the base stock level of the components increase both fill rate and ready rate but has negligible effect on average number of backorders. It is clearly observed that the fill rate is always smaller than the ready rate as the probability of zero inventory level is calculated in the ready rate.

As is illustrated in Figure (A.2)(b), adding the base stock increases the average waiting time which can lead to a high penalty for the supplier under PBCs. The main reason is that the average waiting time of a failed component at the repair facility is directly related to the length of queue in the repair system. On the other hand, when the system faces a backorder, the average length of repair system's queue will be equal to the number of backorders plus base stock ($E(y) = S + E(B)$). Increasing the base stock level can effectively reduce on-order components only if it decreases the number of backorders significantly. As can be found in Figure (A.1), the base stock level has negligible effect on average number of backorders. In such a situation, adding base stock only increases the length of queue in repair system and the average waiting time without reducing the number of backorders.

Appendix B

To evaluate the impact of the base stock on the average number of backorders, the sensitivity of the number of backorders to the base stock is defined as:

$$\Delta B(S, \varphi) = B(0, \varphi) - B(S, \varphi) \quad 0 \leq \varphi \leq 1 \quad (\text{B.1})$$

Figure (B.1) illustrates the behavior of ΔB by varying φ for different levels of base stock.

Fig B.1 The sensitivity of the number of backorders to the base stock

As can be seen from Figure (B.1), adding the base stock decreases the number of backorders only around a special value of φ ($\varphi_1 < \varphi < \varphi_2$), especially when the number of systems is high. Also, it can be observed that this value of φ is directly related to the number of systems ($\varphi \approx 1/N$). This suggests that keeping the base stock can be helpful for the supplier only when he chooses a service rate close to the failure rate ($\mu = N\lambda$). If the supplier sets a service rate much lower than the failure rate ($\mu < N\lambda$) the queue at repair facility increases rapidly and the supplier cannot control the number of backorders, even by adding base stock. On the other hand, when the service rate is much higher than the failure rate ($\mu \gg N\lambda$) there is no queue at the repair centre and the supplier does not need to keep any base stock. The average backorders formula can be simplified to:

$$B = \left(N - \frac{1}{\varphi}\right) \times \left(\frac{a}{a-b}\right) \quad a = \varphi^N e^{\frac{1}{\varphi}} \Gamma\left(N+1, \frac{1}{\varphi}\right) (N\varphi-1)+1 \quad b = \left(\frac{1}{N\varphi}\right)^S \quad (\text{B.2})$$

From the numerical experiments, the following observations are made:

(1) For each value of the number of systems (N) there is a range of φ ($\varphi_2 \leq \varphi \leq 1$) in which the value of 'a' is much larger than the value of 'b', regardless of the base stock level. Therefore, the second term in B will be one.

$$\frac{a}{a-b} \approx 1 \quad \text{for} \quad \varphi_2 \leq \varphi \leq 1 \quad (\text{B.3})$$

As an example, Figure (B.2) shows how the value of $(a/a-b)$ converges to one after a certain value of φ (0.023 and 0.011) for 50 and 100 number of systems, respectively.

(a) Number of systems = 50

(b) Number of systems = 100

Fig B.2 The value of $(a/a-b)$ based on φ and S changes.

(2) There is a range of φ ($0 \leq \varphi \leq \varphi_1$), for each value of the number of systems, on which the value of $a/(a-b)$ is almost zero. When the repair centre is operating fast enough it can replace the failed part in a short amount of time, resulting in zero backorders.

$$\frac{a}{a-b} \approx 0 \quad 0 \leq \varphi \leq \varphi_1 \quad (\text{B.4})$$

As can be seen in Figure (B.3), for 50 and 100 systems the value of $(a/a-b)$ is zero when φ is less than 0.015 and 0.007, respectively.

(a) Number of systems = 50

(b) Number of systems = 100

Fig B.3 The value of $(a/a-b)$ based on φ and S changes.

Appendix C

Optimization method

The “GlobalSearch” solver from Matlab global optimization toolbox is used to solve the supplier’s problem. This solver is designed to solve an optimization problem containing both local and global optima. The supplier’s objective is a nonlinear multivariable function with a set of lower and upper bounds on the design variables. Therefore, first the problem is solved using constrained nonlinear solver “fmincon” from the optimization toolbox which is a local solver. “CreateOptimProblem” function from the optimization toolbox is used to define the supplier’s objective and provide a random initial starting point.

```
Problem=CreateOptimProblem('fmincon','x0',[],'objective',Obj,'lb',[],'ub',[],'options',opts)
```

Solver “fmincon” uses one of the four algorithms; active-set, interior-point, sqp, or trust-region-reflective. The trust-region-reflective algorithm is chosen with “optimset”.

```
opts=optimset('Algorithm','trust-region-reflective','GradObj','on');
```

The trust-region-reflective algorithm is a subspace trust-region method and is based on the interior-reflective Newton method. The algorithm iteration involves an approximate solution of a large linear system using the method of preconditioned conjugate gradients (PCG). “fmincon” solver starts at x_0 and attempts to find a local minimum which may not be a global solution, as “fmincon” is designed to find the solution near to the starting point using a gradient based search method. Therefore, the starting point is critical in determining which minimum point, “fmincon” will find. To overcome this, “GlobalSearch” is applied that uses scatter-search mechanism to provide multiple starting points for “fmincon” to run.

Fig C.1 GlobalSearch steps

“GlobalSearch” runs the local solver (“fmincon”) from the starting point (x_0) which has been provided in “CreateOptimProblem”. Then, the solver repeats the following on three steps until convergence is achieved:

1. Formulate the trust-region subproblem. The trust-region subproblem is an approximation of the objective function that reflects its behavior. In the trust-region method such approximation is defined based on the gradient of the objective function and Hessian matrix.
2. Specify the trial step to move to a point with a lower function value. An algorithm is used to determine a trial step by minimizing the subproblem in a neighborhood around the existing point (called trust-region). In the standard trust-region method, the algorithm involves a Newton process and the computation of a full eigensystem which guarantees an accurate solution.
3. Update the current point or adjusting the trust-region. The existing point is updated if the determined step improves the objective function, otherwise, the current point remains unchanged and the region of trust is adjusted and the step calculation repeated.

After reaching to a converged run, “GlobalSearch” records the start point, the end point and the objective function value. Then, “GlobalSearch” generates a set of potential starting points by using the scatter search algorithm. The best starting point is chosen according to the score function evaluation to run “fmincon”. The score function of each point is the summation of the function value and the constraint violations. “GlobalSearch” uses the results of runs from these two starting points (starting point in “CreateOptimProblem” (x_0) and the best starting point among the set of potential starting points (x_b)) to initialize the following two factors:

1. Basins: There is a set of starting points such that their steepest decent paths lead to the same local minimum point. Such set of points is called the basins of attraction.

Fig C.2 Basins of attraction

In “GlobalSearch” the basins of attraction are assumed to be spherical where the radius is the distance between the starting point and the solution. Then, the initial estimation of the basins of attraction is spheres centered at solution points resulting from start points (x_0) and (x_b) .

2. Threshold: It is referred to as the smallest objective function value from the local solver and initially sets on the smaller objective value reached by running “fmincon” on start points (x_0) and (x_b) .

“GlobalSearch” examines the remaining potential starting points generated by the scatter search algorithm and runs “fmincon” on a point that satisfies the following conditions:

- The trial point is not included in any of the existing basins; and
- The score function of the trial point is less than the current threshold value.

The algorithm parameters (threshold and basin radius) are updated, as follow, only if the solver runs for a trial point.

- The threshold changes by the score value of the start point.

Appendix

- A new basin is defined whose radius is the maximum of existing radius and the distance between the start point and solution point.

“GlobalSearch” stops after running out of all potential starting points generated by the scatter search algorithm. In the end, the “GlobalOptimSolution” is created based on the lowest objective function value.

The results of sensitivity analysis

The results of sensitivity analysis with respect to other supplier’s cost parameters are listed in Tables (C.1-C.5).

Table C.1 Sensitivity of the optimal availability to the average repair cost

Utility	Exponential			Power			Mean-Variance			ΔA
Revenue	Linear	Exp	Step	Linear	Exp	Step	Linear	Exp	Step	
50	0.939	0.972	0.979	0.944	0.966	0.975	0.978	0.921	0.988	-0.00574
20	0.94	0.973	0.98	0.946	0.967	0.978	0.978	0.929	0.991	-0.00344
-20	0.952	0.984	0.983	0.949	0.976	0.981	0.986	0.937	0.993	0.003329
-50	0.952	0.984	0.986	0.95	0.979	0.985	0.989	0.937	0.999	0.005624

Table C.2 Sensitivity of the optimal availability to K_r

Utility	Exponential			Power			Mean-Variance			ΔA
Revenue	Linear	Exp	Step	Linear	Exp	Step	Linear	Exp	Step	
50	0.945	0.967	0.977	0.943	0.971	0.973	0.976	0.923	0.984	-0.00608
20	0.945	0.971	0.979	0.945	0.973	0.975	0.982	0.927	0.991	-0.00275
-20	0.947	0.975	0.983	0.954	0.978	0.986	0.985	0.933	0.994	0.00264
-50	0.956	0.977	0.989	0.96	0.978	0.987	0.991	0.939	0.995	0.006887

Table C.3 Sensitivity of the optimal availability to K_p

Utility	Exponential			Power			Mean-Variance			ΔA
Revenue	Linear	Exp	Step	Linear	Exp	Step	Linear	Exp	Step	
50	0.938	0.966	0.975	0.942	0.967	0.976	0.979	0.923	0.984	-0.00712
20	0.939	0.974	0.976	0.943	0.974	0.977	0.98	0.926	0.985	-0.00436
-20	0.951	0.98	0.98	0.956	0.981	0.985	0.986	0.931	0.994	0.003673
-50	0.952	0.981	0.988	0.957	0.983	0.988	0.987	0.931	0.995	0.005739

Table C.4 Sensitivity of the optimal availability to the service capacity cost

Utility	Exponential			Power			Mean-Variance			ΔA
Revenue	Linear	Exp	Step	Linear	Exp	Step	Linear	Exp	Step	
50	0.937	0.97	0.972	0.941	0.968	0.97	0.98	0.921	0.988	-0.00746
20	0.944	0.976	0.977	0.944	0.97	0.972	0.982	0.921	0.988	-0.00436
-20	0.948	0.983	0.989	0.956	0.98	0.984	0.987	0.931	0.997	0.004936
-50	0.948	0.985	0.99	0.954	0.98	0.985	0.992	0.933	0.998	0.006084

Table C.5 Sensitivity of the optimal availability to $\bar{K}_p$

Utility	Exponential			Power			Mean-Variance			ΔA
Revenue	Linear	Exp	Step	Linear	Exp	Step	Linear	Exp	Step	
50	0.937	0.971	0.975	0.94	0.97	0.969	0.972	0.92	0.983	-0.00861
20	0.942	0.979	0.977	0.941	0.971	0.971	0.978	0.929	0.988	-0.00413
-20	0.952	0.981	0.987	0.952	0.979	0.982	0.987	0.937	0.995	0.004591
-50	0.959	0.986	0.989	0.955	0.984	0.983	0.988	0.942	0.996	0.008035

All deviations in the optimal availability are negligible (less than 0.9%) with respect to parameters changes, implying that the results are robust against error in supplier's cost parameters estimation.

Appendix D

Simulation Model

In the analytical model some simplifying assumptions are made regarding the nature of the probability distribution of random variables such as mean time between failures and repair time. In this section, a simulation model is developed using Arena software to validate the analytical model and offer a versatile solution approach for more realistic and possibly more complex scenarios.

Replication 1 Start Time: 500.00 Stop Time: 17,562,386.94 Time Units: Hours

Resource Detail Summary

Usage

	<u>Inst Util</u>	<u>Num Busy</u>	<u>Num Sched</u>	<u>Num Seized</u>	<u>Sched Util</u>
Op System	0.89	35.53	40.00	3,253,027.00	0.89
Repair Server	0.96	0.96	1.00	3,253,022.00	0.96

Fig D.1 Simulation model using Arena software

First, numerical examples in the thesis are used to validate the corresponding analytical model.

Table D.1 Analytical and Simulation results comparison

	Parameters				Analytical		Simulation	
	N	λ	μ	S	$E(B)$	A	$E(B)$	A
1	40	1.92	71.2	6	4.27	0.893	4.49	0.89
2	100	0.2	17	8	15.26	0.847	15.34	0.85
3	30	0.85	24	2	3.91	0.869	4.11	0.86
4	40	1.44	64.8	4	1.65	0.958	2.19	0.95
5	60	0.017	1.23	3	1.59	0.973	1.86	0.97
6	40	2.88	98.4	8	6.2	0.845	6.29	0.84

The repair lead-time of each server in the repair facility is assumed to be an independent exponential random variable with mean $(\mu)^{-1}$. This assumption is common in the literature (Wong et al. 2005) but in this section another realistic distribution (i.e. Gamma) is tested using simulation based method and the results are compared with the first scenario (Table D.2).

Table D.2 Impact of service time distribution on availability

	Parameters				Service Time Distribution	
	N	λ	μ	S	$EXPO(1/\mu)$	$GAMM(1/k\mu, k)$
1	40	1.92	71.2	6	0.89	0.90
2	100	0.2	17	8	0.85	0.85
3	30	0.85	24	2	0.86	0.87
4	40	1.44	64.8	4	0.95	0.95
5	60	0.017	1.23	3	0.97	0.97
6	40	2.88	98.4	8	0.84	0.85

As can be seen, changing the service time distribution will not affect the performance outcome (i.e. operational availability). In fact, the system availability determines based on the average repair time and not the repair time distribution. This also can be justified by referring to Palm's theorem (Sherbrooke 1986).

Appendix

The mean time between failures may follow Weibull distribution as the chance of failure is higher at early and late life of component and relatively flat in between. The simulation model is used to show how the results may change using different distribution of MTBF. If the quantity X is a "time-to-failure", the Weibull distribution gives a distribution for which the failure rate is proportional to a power of time. The shape parameter, k , is that power plus one, and so this parameter can be interpreted directly as follows:

- A value of $k < 1$ indicates that the failure rate decreases over time. This happens if there is significant "infant mortality", or defective items failing early and the failure rate decreasing over time as the defective items are weeded out of the population.
- A value of $k = 1$ indicates that the failure rate is constant over time. This might suggest random external events are causing mortality, or failure.
- A value of $k > 1$ indicates that the failure rate increases with time. This happens if there is an "aging" process, or parts that are more likely to fail as time go on.

Table D.3 Impact of component's failure distribution on availability

	Parameters				Weibull Distribution		
	N	λ	μ	S	$k=0.5$	$k=1$	$k=2$
1	40	1.92	71.2	6	0.97	0.89	0.81
2	100	0.2	17	8	0.99	0.85	0.75
3	30	0.85	24	2	0.96	0.86	0.81
4	40	1.44	64.8	4	0.97	0.95	0.92
5	60	0.017	1.23	3	0.98	0.97	0.94
6	40	2.88	98.4	8	0.97	0.84	0.76

As can be seen, the availability increases if the failure rate decreases over time ($k < 1$) and decrease if the failure rate increases over the time ($k > 1$).

References

- Alfredsson, P. and J. Verrijdt (1999). Modeling emergency supply flexibility in a two-echelon inventory system. *Management Science*, 45(10), 1416-1431.
- Allon, G. and A. Federgruen (2007). Competition in service industries. *Operations Research*, 55(1), 37-55.
- Axsäter, S. (1990). Modelling emergency lateral transshipments in inventory systems. *Management Science*, 36(11), 1329-1338.
- Banker, R. D. and S. M. Datar (1989). Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research*, 27(1), 21-39.
- Berends, T. (2000). Cost plus incentive fee contracting experiences and structuring. *International Journal of Project Management*, 18(3), 165-171.
- Berkowitz, D., Gupta, J. N. D., Simpson, J. T., & McWilliams, J. B (2005). Defining and implementing performance based logistics in government. *Defense Acquisition Review Journal*.37, 255-267.
- Billinton, R., L. Cui and Z. Pan (2002). Quantitative reliability considerations in the determination of performance-based rates and customer service disruption payments. *IEE Proceedings-Generation, Transmission and Distribution*, 149(6), 640-644.
- Bliss, R. R. and N. Panigirtzoglou (2004). Option-implied risk aversion estimates. *The Journal of Finance*, 59(1), 407-446.
- Brown, R. E. and J. J. Burke (2000). Managing the risk of performance based rates. *Power Systems, IEEE Transactions on*, 15(2), 893-898.

References

Bubshait, A. A. (2003). Incentive/disincentive contracts and its effects on industrial projects. *International Journal of Project Management*, 21(1), 63-70.

Cassady, C. R., E. A. Pohl and S. Jin (2004). Managing availability improvement efforts with importance measures and optimization. *IMA Journal of Management Mathematics*, 15(2), 161-174.

Chen, F. and A. Federgruen (2000). Mean-variance analysis of basic inventory models. *Technical manuscript, Columbia University*.

Chen, J. Y. and B. L. Miller (2009). On the Relative Performance of Linear vs. Piecewise-Linear-Threshold Intertemporal Incentives. *Management Science*, 55(10), 1743-1752.

Cho, D. I. and M. Parlar (1991). A survey of maintenance models for multi-unit systems. *European Journal of Operational Research*, 51(1), 1-23.

Cohen, M. A., N. Agrawal and V. Agrawal (2006). Achieving breakthrough service delivery through dynamic asset deployment strategies. *Interfaces*, 36(3), 259-271.

Cunic, B. (2003). *Performance-based contracting*. Hydrocarbon Processing. <http://www.hydrocarbonprocessing.com/IssueArticle/2599519/Archive/Performance-based-contracting.html>

Dang, W., Y. X. Chen and Y. M. Sun (2009). Research on the performance based logistics theory and applying. Processing of Management and Service Science Conference, 1-5, Wuhan.

de Smidt-Destombes, K. S., M. C. van der Heijden and A. van Harten (2007). Availability of k-out-of-N systems under block replacement sharing limited spares and repair capacity. *International journal of production economics*, 107(2), 404-421.

Dekker, R. (1996). Applications of maintenance optimization models: a review and analysis. *Reliability Engineering & System Safety*, 51(3), 229-240.

References

Dennis, M. J. and A. Kambil (2003). Service Managemnet: Building Profits after the Sale. *Supply Chain Managemnet Review*, 7(3),42-48.

Department of Defense (2004).*Defense Acquisition Guidebook (DAG)*.Retrieved April 18, 2005 from <http://dap.dau.mil>.

Devries, H. J. (2004). Performance Based Logistics-Barriers and Enablers to Effective Implementation.*Defense Acquisition Review Journal*. 11, 243-254.

Dhakar, T. S., C. P. Schmidt and D. M. Miller (1994). Base stock level determination for high cost low demand critical repairable spares. *Computers & operations research*, 21(4), 411-420.

Diaz, A. and M. C. Fu (1997). Models for multi-echelon repairable item inventory systems with limited repair capacity. *European Journal of Operational Research*, 97(3), 480-492.

DoD (March 2006). Continuous Process Improvement Transformation Guidebook. *Office of the Deputy under Secretary for Logistics and Material Readiness*.<http://dap.dau.mil>

Doerr, K., Eaton, D.R., & Lewis I.A. (2004). Measurement Issues in Performance Based Logistics. Technical Report, *Naval Postgraduate School*.

Fearnley, N., J.-T. Bekken and B. Norheim (2004). Optimal performance-based subsidies in Norwegian intercity rail transport. *International Journal of Transport Management*, 2(1), 29-38.

Garrett, G. A. (2001). World class contracting: How winning companies build successful partnerships in the e-Business age, *CCH Incorporated*.

Geary, S. (2006). Ready for combat. *DC Velocity*, 4(7), 75-80.

References

Gibbons, R. (2005). Incentives between firms (and within). *Management Science*, 51(1), 2-17.

Global Wind Energy Council (2012) Global Wind Report: Annual Market Update 2011. http://gwec.net/wp-content/uploads/2012/06/Annual_report_2011_lowres.pdf (last accessed on June 6, 2013)

Graves, S. C. (1985). A multi-echelon inventory model for a repairable item with one-for-one replenishment. *Management Science*, 31(10), 1247-1256.

Gross, D., D. R. Miller and R. M. Soland (1985). On some common interests among reliability, inventory, and queuing. *Reliability, IEEE Transactions on*, 34(3), 204-211.

Gupta, A. and S. C. Albright (1992). Steady-state approximations for a multi-echelon multi-indentured repairable-item inventory system. *European Journal of Operational Research*, 62(3), 340-353.

Hemmer, T. (2004). Lessons lost in linearity: A critical assessment of the general usefulness of LEN models in compensation research. *Journal of Management Accounting Research*, 16(1), 149-162.

Hemmer, T., O. Kim and R. E. Verrecchia (2000). Introducing convexity into optimal compensation contracts. *Journal of Accounting and Economics*, 28(3), 307-327.

Hensher, D. A. and E. Houghton (2004). Performance-based quality contracts for the bus sector: delivering social and commercial value for money. *Transportation Research Part B: Methodological*, 38(2), 123-146.

Hölmstrom, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 74-91.

Holmstrom, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, 303-328.

References

- Huang, H.-Z., Z.-J. Liu and D. Murthy (2007). Optimal reliability, warranty and price for new products. *Iie Transactions*, 39(8), 819-827.
- Jewitt, I. (1988). Justifying the first-order approach to principal-agent problems. *Econometrica: Journal of the Econometric Society*, 1177-1190.
- Jin, T. and H. Liao (2009). Spare parts inventory control considering stochastic growth of an installed base. *Computers & Industrial Engineering*, 56(1), 452-460.
- Jin, T. and Y. Tian (2012). Optimizing reliability and service parts logistics for a time-varying installed base. *European Journal of Operational Research*, 218(1), 152-162.
- Kang, S.J. (2001). Study on how to apply the PRICE model in Korean environment. *Korea National Defense University, Technical Report, Seoul, Korea*.
- Kang, K., K. Doerr, M. Boudreau and U. Apte (2005). A decision support model for valuing proposed improvements in component reliability, *2nd Annual Acquisition Research Symposium of the Naval Postgraduate School*.
- Kang, K., K. H. Doerr and S. M. Sanchez (2006). A design of experiments approach to readiness risk analysis. *Proceedings of the 38th conference on Winter simulation, Winter Simulation Conference*.1332 – 1339.
- Kim, S.-H. (2008). Performance-Based Logistics: Incentive contracting in the aftermarket. Ann Arbor, *University of Pennsylvania*.
- Kim, S.-H., M. Cohen and S. Netessine (2011). Reliability or inventory? Analysis of product support contracts in the defense industry.
- Kim, S.-H., M. A. Cohen and S. Netessine (2007). Performance contracting in after-sales service supply chains. *Management Science*, 53(12), 1843-1858.

References

- Kim, S.-H., M. A. Cohen, S. Netessine and S. Veeraraghavan (2010). Contracting for infrequent restoration and recovery of mission-critical systems. *Management Science*, 56(9), 1551-1567.
- Kirkpatrick, D. L. (2004). Trends in the Costs of Weapon Systems and the Consequences. *Defence and Peace Economics*, 15(3), 259-273.
- Kirkwood, C. W. (2002). Decision tree primer. *Department of Supply Chain Management, Arizona State University Tempe, AZ*, 85287-84706.
- Kumar, U. D., D. Nowicki, J. E. RAMÍREZ-MÁRQUEZ and D. VERMA (2007). A goal programming model for optimizing reliability, maintainability and supportability under performance based logistics. *International Journal of Reliability, Quality and Safety Engineering*, 14(03), 251-261.
- Lambert, R. A., D. F. Larcker and R. E. Verrecchia (1991). Portfolio considerations in valuing executive compensation. *Journal of Accounting Research*, 29(1), 129-149.
- Landreth, C. J. and L. L. Corporon (2005). Performance Based Logistics (PBL) for the FA-18/S-3/P-3/C-2 auxiliary power unit (APU) at Honeywell: an applied analysis. *Naval Postgraduate School*, Monterey, CA
- Lau, H. C. and H. Song (2004). Two-echelon repairable item inventory system with limited repair capacity under nonstationary demands. 35th Meeting of the Decision Sciences Institute (DSI). USA: Boston, *Citeseer*.
- Lee, H. L. (1987). A multi-echelon inventory model for repairable items with emergency lateral transshipments. *Management Science*, 33(10), 1302-1316.
- Little, J. D. (1961). A proof for the queuing formula: $L = \lambda W$. *Operations Research*, 9(3), 383-387.

References

Los Angeles Times (January 26, 1999). Firms' Spending on Health Care Up 6.1% in '98, Survey Finds. <http://articles.latimes.com/1999/jan/26/business/fi-1671>

Lucintel Report. Growth opportunities in wind operation and maintenance services market 2011-2016, (last accessed on June 6, 2013), http://www.lucintel.com/reports_details.aspx?RepId=RPT1033

Mangione-Smith, R., M. Schonlau, K. S. Chan, J. Keesey, M. Rosen, T. A. Louis and E. Keeler (2005). Measuring the effectiveness of a collaborative for quality improvement in pediatric asthma care: does implementing the chronic care model improve processes and outcomes of care? *Ambulatory Pediatrics*, 5(2), 75-82.

Maskell, B. H. (1991). *Performance Measurement for World-Class Manufacturing*. Productivity Press.

McAfee, R. P. and J. McMillan (1986). Bidding for contracts: a principal-agent analysis. *The RAND Journal of Economics*, 326-338.

McCall, J. J. (1965). Maintenance policies for stochastically failing equipment: a survey. *Management Science*, 11(5), 493-524.

Mettas, A. (2000). Reliability allocation and optimization for complex systems. Reliability and Maintainability Symposium, 2000. Proceedings. Annual, *IEEE*.

Mirrlees, J. (1974). Notes on welfare economics, information and uncertainty. *Essays on Economic Behavior under Uncertainty*, 243-261.

Mirzahosseini, H. and R. Piplani (2011). Compensation and Incentive Modeling in PBCs for After-market Service. *Proceedings of the 41st International Conference on Computers & Industrial Engineering*, Los Angeles, USA.

Mirzahosseini, H. and R. Piplani (2011). A study of repairable parts inventory system operating under PBC. *European Journal of Operational Research*, 214(2), 256-261.

References

- Mirzahosseini, H. and R. Piplani (2013). Impact of Operational Systems on Supplier's Response Under PBCs. *Procedia Computer Science*, 16, 639-647.
- Muckstadt, J. A. (1973). A model for a multi-item, multi-echelon, multi-indenture inventory system. *Management Science*, 20(4-Part-I), 472-481.
- Murphy, K. J. (2000). Performance standards in incentive contracts. *Journal of Accounting and Economics*, 30(3), 245-278.
- Naval Air Systems Command [NAVAIR]. (2004b, December 01). Policy Guidance for Performance Based Logistics Candidates (NAVAIRINST 4081.2A). Retrieved June 07, 2005 from <https://directives.navair.navy.mil//index.cfm>
- Neely, A., M. Gregory and K. Platts (1995). Performance measurement system design: a literature review and research agenda. *International journal of operations & production management*, 15(4), 80-116.
- Ng, I. C. L., R. Maull and N. Yip (2009). Outcome-based contracts as a driver for systems thinking and service-dominant logic in service science: Evidence from the defence industry. *European Management Journal*, 27(6), 377-387.
- Nowicki, D., U. D. Kumar, H. J. Steudel and D. Verma (2008). Spares provisioning under performance-based logistics contract: Profit-centric approach. *Journal of the Operational Research Society*, 59(3), 342-352.
- Office of the Secretary of Defense, Undersecretary for Acquisition, Technology & Logistics [OSD-ATL].(2001). Incentive Strategies for Defense Acquisitions. Retrieved May 31, 2005, from <http://www.acq.osd.mil/>.
- Öner, K. B., G. P. Kiesmüller and G.-J. van Houtum (2010). Optimization of component reliability in the design phase of capital goods. *European Journal of Operational Research*, 205(3), 615-624.

References

- Phillips, E. H. (2005). A Whole New Approach. *Aviation Week & Space Technology*, 163(17), 52-55.
- Pierskalla, W. P. and J. A. Voelker (1976). A survey of maintenance models: the control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, 23(3), 353-388.
- Prendergast, C. (1999). The provision of incentives in firms. *Journal of economic literature*, 37(1), 7-63.
- Raju, J. S. and V. Srinivasan (1996). Quota-based compensation plans for multiterritory heterogeneous salesforces. *Management Science*, 42(10), 1454-1462.
- Rievley, J. (2001). D., Performance-Based Launch Services Contracts: Better, Faster, Cheaper. Presented to the 38th Space Congress. Air Force Contracting Officer 45 CONS/LGCZL (321) 494-9519.
- Rogers, S. C. (1997). Performance Based Supportability Analysis: A Success Story. *Logistics Spectrum*, 31, 10-14.
- Scherer, F. M. (1964). The theory of contractual incentives for cost reduction. *The Quarterly Journal of Economics*, 78(2), 257-280.
- Scherer, F. M. (1964). The weapons acquisition process: economic incentives, *Division of Research, Graduate School of Business Administration, Harvard University*.
- Sherbrooke, C. C. (1968). METRIC: A multi-echelon technique for recoverable item control. *Operations Research*, 16(1), 122-141.
- Sherbrooke, C. C. (1986). VARI-METRIC: improved approximations for multi-indenture, multi-echelon availability models. *Operations Research*, 34(2), 311-319.
- Sherbrooke, C. C. (1992). Multiechelon inventory systems with lateral supply. *Naval Research Logistics (NRL)*, 39(1), 29-40.

References

Sherif, Y. and M. Smith (1981). Optimal maintenance models for systems subject to failure—a review. *Naval Research Logistics Quarterly*, 28(1), 47-74.

Sleptchenko, A., M. Van Der Heijden and A. Van Harten (2003). Trade-off between inventory and repair capacity in spare part networks. *Journal of the Operational Research Society*, 54(3), 263-272.

Smith, S.J., (2007). Equipping the American forces with a reliable and robust condition based maintenance plus capability. USAF, (last accessed on September 23, 2012), www.acq.osd.mil/log/mpp/cbm+/_LG301T6_FINAL.PDF

Smith, V.D., Searles, D.G., Thompson, B.M., Cranwell, R.M.,(2006). SEM: enterprise modeling of JSF global sustainment, *Proceedings of the Winter Simulation Conference*. 1324-1331.

So, K. C. and C. S. Tang (2000). Modeling the impact of an outcome-oriented reimbursement policy on clinic, patients, and pharmaceutical firms. *Management Science*, 46(7), 875-892.

Sohn, S., Y. Kim and B. Kim (2009). Cost of ownership model for spare engines purchase for the Korean navy acquisition program. *Journal of the Operational Research Society*, 60(12), 1674-1682.

Sols, A. and L. H. Johannesen (2013). The role of transition contracts in performance-based logistics initiatives. *Systems Engineering*, 16 (4), 453-463

Szegö, G. P. (2004). Risk measures for the 21st century, *Wiley*.

Tao, N. and S. Wen (2009). Simulation of a closed loop multi-echelon repairable inventory system. *Management Science and Engineering*, 2009. ICMSE 2009. International Conference on, *IEEE*.

References

- Valdez-Flores, C. and R. M. Feldman (1989). A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics (NRL)*, 36(4), 419-446.
- Van Mieghem, J. A. (2007). Risk mitigation in newsvendor networks: Resource diversification, flexibility, sharing, and hedging. *Management Science*, 53(8), 1269-1288.
- Vitasek, K., J. Cothran, S. Geary and S. Rutner (2006). Performance-based logistics: The changing landscape in support contracting. *Knoxville, TN: The University of Tennessee*.
- Wang, H. (2002). A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3), 469-489.
- Wang, Y., M. A. Cohen and Y. S. Zheng (2000). A two-echelon repairable inventory system with stocking-center-dependent depot replenishment lead times. *Management Science*, 46(11), 1441-1453.
- Wasserman, G. S. and L.R. Lamberson (1992). Spares Provisioning Under Reliability Growth. *Logistics Spectrum Winter*, 26 (4).
- Wong, H., D. Cattrysse and D. Van Oudheusden (2005). Stocking decisions for repairable spare parts pooling in a multi-hub system. *International journal of production economics*, 93, 309-317.
- Yoner, N. (2001). Major weapon systems acquisition and life cycle cost estimation: A case study, *Naval Postgraduate School*, Monterey, CA..
- Zhou, X. and P. L. Swan (2003). Performance Thresholds in Managerial Incentive Contracts. *The Journal of Business*, 76(4), 665-696.