AN ANALYSIS OF BENEFITS OF INVENTORY AND SERVICE POOLING AND INFORMATION SHARING IN SPARE PARTS MANAGEMENT SYSTEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BENHÜR SATIR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF PHILOSOPHY OF DOCTORATE
IN
INDUSTRIAL ENGINEERING

JUNE 2010

Approval of the thesis:

AN ANALYSIS OF BENEFITS OF INVENTORY AND SERVICE POOLING AND INFORMATION SHARING IN SPARE PARTS MANAGEMENT SYSTEMS

submitted by **BENHÜR SATIR** in partial fulfillment of the requirements for the degree of
Philosophy of Doctorate in Industrial Engineering Department, Middle East Technical University by,

Prof. Dr. Canan Özgen Dean, Graduate School of **Natural and Applied Sciences**	______________
Prof. Dr. Nur Evin Özdemirel Head of Department, **Industrial Engineering**	______________
Asst. Prof. Dr. Seçil Savaşaneril Tüfekçi Supervisor, **Industrial Engineering Department**	______________
Assoc. Prof. Dr. Yaşar Yasemin Serin Co-supervisor, **Industrial Engineering Department**	______________

Examining Committee Members:

Prof. Dr. Murat Köksalan Industrial Engineering Department, METU	______________
Asst. Prof. Dr. Seçil Savaşaneril Tüfekçi Industrial Engineering Department, METU	______________
Prof. Dr. Fetih Yıldırım Department of Industrial Engineering, Çankaya University	______________
Prof. Dr. Ülkü Gürler Department of Industrial Engineering, Bilkent University	______________
Asst. Prof. Dr. Sedef Meral Industrial Engineering Department, METU	______________

Date: ______________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: BENHÜR SATIR

Signature :

ABSTRACT

AN ANALYSIS OF BENEFITS OF INVENTORY AND SERVICE POOLING AND INFORMATION SHARING IN SPARE PARTS MANAGEMENT SYSTEMS

Satır, Benhür
Ph.D, Department of Industrial Engineering
Supervisor : Asst. Prof. Dr. Seçil Savaşaneril Tüfekçi
Co-Supervisor : Assoc. Prof. Dr. Yaşar Yasemin Serin

June 2010, 184 pages

Inventory management and production control problem of a dealer operating in a decentralized spare parts network is analyzed in this dissertation. Spare parts network is assumed to be formed of two dealers and the problem of a dealer is considered under the assumption that the other dealer has a known policy. These dealers collaborate through inventory and service pooling. Furthermore, the dealers collaborate through sharing information on the net inventory status.

Upon demand arrival, a dealer may request a part from the other dealer, in which case a payment is made. Under this competitive and collaborative environment, the optimal operating policy of an individual dealer is characterized under full information. Through computational analysis, the conditions under which the dealer under consideration is most profitable are identified. Finally, by comparing different pooling strategies and several information availability levels, the benefit of information sharing is quantified.

Main findings are as follows. Using a centralized model with a central authority, comparison of optimal operating policies shows that there are opposite behaviors between centralized and decentralized system structures for rationing and transshipment levels. In decentralized model, for a few instances, non-monotonic behavior of optimal operating policy is observed. It is shown that dealer under consideration prefers extreme values of payment amount. Observation indicate that this preference depends on the base-stock level of the other dealer. Diminishing returns on profit over customer arrival rate of dealer under consideration is observed, whereas for customer arrival rate of the other dealer effect on profit depends on payment amount. Comparison of pooling strategies indicate that operating without pooling outperforms a poorly designed pooling strategy. Fixed and variable payment amounts are observed to effect the behavior of rationing and transshipment levels of the optimal operating policy of the dealer under consideration oppositely. For high customer arrival rate and low payment or vice-versa, information value is found to be low, whereas it is high for variable payment. The parameter effects to the value of information are analyzed with respect to usage of information in the optimal operating policy of the dealer under consideration. Information on the other dealer is found to be more valuable compared to information on the dealer under consideration itself.

Keywords: Inventory and service pooling, optimal policy characterization, pooling strategies, information availability, information sharing levels

ÖZ

YEDEK PARÇA YÖNETİM SİSTEMLERİNDE ENVANTER VE SERVİS HAVUZLAMANIN VE BİLGİ PAYLAŞIMININ FAYDALARININ BİR ÇÖZÜMLEMESİ

Satır, Benhür
Doktora, Endüstri Mühendisliği Bölümü
Tez Yöneticisi : Yard. Doç. Dr. Seçil Savaşaneril Tüfekçi
Ortak Tez Yöneticisi : Doç. Dr. Yaşar Yasemin Serin

Haziran 2010, 184 sayfa

Bu tezde, merkezî olmayan bir yedek paça ağı içerisinde çalışan bir yedek parça sağlayıcısının envanter yönetimi ve üretim kontrol problemi çözümlemesi yapılmıştır. Yedek paça ağının iki dağıtıcıdan oluştuğu varsayılmıştır ve bir dağıtıcının problemi, diğer dağıtıcının bilinen bir politika altında çalıştığı varsayımı altında düşünülmüştür. Bu dağıtıcılar, havuzlamayı envanter ve servis için yapmaktadırlar. Buna ilâveten, dağıtıcılar net envanter seviyeleri hakkında bilgi paylaşımı ile işbirliği yapmaktadırlar.

Müşteri gelmesi üzerine, bir dağıtıcı diğer dağıtıcıdan parça isteyebilir, ki bu durumda bir ödeme yapılır. Bu rekabetçi ve işbirlikçi yapıda, tam bilgi altında bir dağıtıcının eniyi çalışma politikasının nitelendirilmesi yapılmıştır. Sayısal çözümleme ile, göz önüne alınan dağıtıcı için en kârlı şartlar tanımlanmıştır. Son olarak, değişik havuzlama stratejileri ve çeşitli bilgi mevcûdiyeti seviyeleri kıyaslanarak, bilgi paylaşımının faydası nicelenmiştir.

Ana hatları ile sonuçlar şu şekildedir. Merkezi bir otoritenin sahip olduğu merkezi bir model kullanılarak yapılan kıyas sonucunda, merkezi olan ve olmayan yapılar arasında, tayınlama ve taşıma seviyeleri için zıt davranışlar gözlemlenmiştir. Az sayıda örnek için eniyi çalışma politikasında monoton olmayan davranışlar görülmüştür. Odaklanılan dağıtıcının, gelir paylaşım miktarının uç değerlerini tercih ettiği gösterilmiştir. Gözlemler, bu tercihin diğer dağıtıcının hedef stok seviyesine bağlı olduğunu işaret etmektedir. Odaklanılan dağıtıcıya gelen müşterilerin geliş oranının kâra etkisinin artan ve ardından azalan olduğu, diğer dağıtıcının oranının etkisinin ise gelir paylaşım miktarına bağlı olduğu görülmüştür. Değişik havuzlama stratejilerinin kıyasları, havuzlama yapmadan çalışmanın, kötü tasarlanmış bir havuzlama politikasından daha iyi olduğunu göstermiştir. Sabit ve değişken ödeme miktarlarının, odaklanılan dağıtıcının eniyi çalışma politikasındaki tayınlama ve taşıma seviyeleri için zıt bir davranışa yol açtığı gözlemlenmiştir. Bilgi değerinin, yüksek müşteri geliş oranı ve düşük ödeme altında veya tersi durumda düşük olduğu; değişken ödeme miktarı altında ise yüksek olduğu görülmüştür. Parametrelerin bilgi değerine etkileri, odaklanılan dağıtıcının eniyi çalışma politikasında bilginin nasıl kullanıldığı ile çözümlenilmiştir. Diğer dağıtıcı hakkındaki bilginin, odaklanılan dağıtıcının kendi bilgisinden daha değerli olduğu bulunmuştur.

Anahtar Kelimeler: Envanter ve servis havuzlama, eniyi politika nitelendirilmesi, havuzlama stratejileri, bilgi mevcudiyeti, bilgi paylaşımı seviyeleri

to my wife Ferda and our daughter Ayşe Râna,
and
in memory of Kemâl Çağrı

ACKNOWLEDGMENTS

The major part of my thanks goes to my supervisor Asst. Prof. Dr. Seçil Savaşaneril Tüfekçi and my cosupervisor Assoc. Prof. Dr. Yaşar Yasemin Serin. Their subtle ideas and approaches to the problems that I encountered during the study have always guided me in the correct manner. Especially, I would like to express that although I am the first doctorate student of my supervisor, she managed the study very professionally and also she has contributed significantly to the study with her hard-working personality.

I owe special thanks to the Thesis Monitoring Committee Members, Prof. Dr. Murat Köksalan and Prof. Dr. Ülkü Gürler, for their valuable contributions.

I am thankful to my chairman Prof. Dr. Levent Kandiller, for his tolerance and support to my graduate study. He always shed light on my academic career and my life, as well.

I present my esteem to Prof. Dr. Fetih Yıldırım for guidance and assistance to the decisions about not only my academic studies only but also my life.

I thank to Asst. Prof. Dr. Sedef Meral for her valuable contributions as a jury member of my thesis defense.

I deem myself as fortunate having parents who have never given up their special support throughout my life without examining my rightness. I present my deepest thanks to my mother and father, Mükerrem and İsmail Satır. Special thanks go to my brother M.D. İbrahim Özgür Satır, for his advices about my life, especially for my occupation.

Love of my wife, Ferda, have given me strength and confidence ever since we met. She is a very special person for me. Very special thanks go to her.

The gift of God to our family, our daughter Ayşe Râna, and my love to her motivated me all the time.

I would like to mention that this study is a part of research project supported by TÜBİTAK with the project number 108M004. I would like to thank to TÜBİTAK for the support.

Thanks to everyone I could not mention here.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Manufacturers of automobiles, expensive and complex equipments always have the problem of maintaining high customer service levels while keeping inventory holding costs and loss of goodwill costs at acceptable levels. These equipments may become unusable due to even a single component failure. Therefore, high service level is characterized by part availability. Providing quality after sales service results in increased customer loyalty and repeated purchases, which is the major source of revenue. In after sales business the revenue obtained accounts for a low percentage of the total revenues, while the profit constitutes a much higher proportion of the total profits. For example in automotive industry, GM earned $150 billion from product sales in 2001, and though revenues from after-sales were $9 billion, profits were higher compared to those obtained from product sales (Cohen et al. [2006]). Although there is significant revenue-bringing potential in after sales services, it is not easy to convert this revenue into profits because of high inventory holding and loss of goodwill costs incurred due to the very unpredictable nature of the demand. Since part availability is crucial for providing a timely service, the companies of after-sales services face the difficult task of being responsive while managing the inventory effectively and keeping the inventory holding and loss of goodwill costs low.

In general, the part providers in the spare parts network of an after sales service system operate as follows. When an equipment malfunctions, the customer asks for service from an after-sales service provider in the spare parts network. Until the equipment is fixed (until a spare part is ready at the service provider and the broken part is replaced), a downtime cost accrues to the customer. There are certain parts that are infrequently needed and are relatively costly to hold inventory for. Downtime depends on how quickly the spare part is supplied for the equipment. If the spare part inventory is close to the service provider, downtime is shorter. Original Equipment Manufacturers (OEMs) have large number of distribution centers in order to locate the spare part stocks close to after-sales service providers, but this may result in excessive inventory holding costs in the overall spare parts network. A requirement of a part puts a pressure on service provider, as well as on manufacturer until the spare part is supplied, and this pressure is the highest when the spare part does not exist in the spare parts network and it should be manufactured by the OEM (in this case, the part should pass through the spare parts network to reach to the service provider). In spare parts systems, this problem can be alleviated by sharing/pooling of spare parts inventory among the after-sales service providers of the spare parts network. The lowest-level after-sales service providers, namely **dealers** or **retailers** or **service centers** share their inventory through **lateral transshipment**. With the improvements in transshipment methods in terms of speed and cost, and expanded use of information systems in spare parts networks, the cost of inventory and information sharing has been reduced and inventory pooling is becoming more attractive for after-sales service providers in spare parts networks.

In practice, many applications of inventory pooling can be observed. An example is the Saturn Corporation which re-constructed its service parts supply chain (Cohen et al. [2000]). The key component in this process is pooling of the component inventories. Dealers that are in close proximity with each other were formed into groups such that if one of the dealers is out of stock, upon a demand arrival a part would be transshipped from another dealer in the group. In case of a lateral transshipment, a full reimbursement is made to the sending dealer. If the item does not exist in the group, dealer requests the part from Saturn. Pooling the inventory resulted in significant savings in the inventory holding costs while improving the service levels. After adopting this pooling group strategy in the supply chain, Saturn ranked first in US in terms of part availability. Narus and Anderson [1996] state that, by using right inventory pooling strategies, costs may decrease up to 20% and at the same time, lost business due to stockout and lack of response may decrease up to 75%. Volvo GM is the first example of Narus and Anderson [1996], where Volvo GM is reducing its inventory by 15% by reducing number of warehouses, and has a lower level of business loss at the same time.

In the aircraft industry, in both military and non-military industries, inventory pooling may improve the service level while decreasing the high inventory holding costs. Jung et al. [2003] indicate that in spite of the fact that US Military had 46.4 billion dollars of repairable spare parts inventory in September 2002, 10% of military aircrafts are grounded due to lack of spare part. Authors work with real data from US Air Force and through numerical experiments, the authors conclude that with the same (and even with slightly better) fill rates, lateral transshipment decreases the total cost significantly, which may reach to 80% in heavy-traffic setting. Cooperative pooling in air transport industry has its roots in 1960s, companies such as KSSU and Atlas formed consortiums for maintenance of aircrafts. Although those attempts have a long history, there is still room for improvement in spare parts systems in air transport industry. Kilpi and Vepsalainen [2004] show that through pooling, spare parts inventories can be lowered over 30% with a minor sacrifice in service levels.

There are local applications of inventory pooling and information sharing in Turkey. TOFAŞ is one of the largest automobile manufacturers in Turkey with 152 authorized services (service centers) and around 150 spare parts providers.[1] TOFAŞ built an information system, namely DAS (Dealer Automation System), for its authorized services to locate spare part inventory in the whole chain. Whenever a customer arrives for service to a dealer, the dealer accepts the customer, regardless of how tight his schedule is. After customer arrival, the demand for the required part(s) is met from dealer's own inventory, if possible. In case of a stockout, the dealer may request the part from TOFAŞ or may perform a search over DAS to see the stock level of the parts in authorized dealers. DAS lists all the dealers having the part in stock as well as the inventory at TOFAŞ. DAS enables a dealer to know whether the part exists or not at other dealers. Dealers that have the part might be in the same city or in different cities. The dealer usually prefers to request a part from a dealer that is closer than TOFAŞ. If request is accepted, a price is determined through negotiation. In the extreme case, the owner of the part may put a price including all profit, and the requesting dealer may accept the offer just for customer satisfaction. The transshipment cost is paid

[1] The information on TOFAŞ's system is gathered by an interview with one of the authorized service managers in Ankara, namely Kavaklıdere Service.

by the requesting dealer. If the request is rejected, the dealer orders the part from TOFAŞ and it arrives in a day. But, in this case the transportation cost is paid by the dealer. Although this option may have a higher net profit, requesting from a closer dealer might be more preferable since customer satisfaction has prime importance. Last choice is requesting from a dealer located in another city, but as TOFAŞ does not have any incentives and subsidies, and all the costs are incurred to the requesting dealer, and lateral transshipment may become a real loss in that case.

Another example is about VOLVO, which is one of the largest construction equipment manufacturers in the world.[2] In Turkey, there are twelve aftersales service providers, where five of them belongs to Akça Machinery. All these five service centers use two different software, namely MMI (Manufacturer Managed Inventory) and SS (Solar System). MMI is developed by VOLVO and SS is tailored by a domestic software developer company. SS allows any of these five service centers to see the inventory levels of other service centers fully, in terms of amount and the time already spent in stock. If a required item is not in the inventory but is located at another Akça service center, the time the part spent waiting in stock is checked. If it has been in inventory for more than a year, then it is requested from the service center and as a company policy, the item is immediately shipped. If the age of the part is less than a year, the request is evaluated by the service center. Generally, if there are two or more items in the inventory, the request is accepted, but if there is a single item or no item in the inventory, then the request is rejected. In order to place a lateral transshipment request to a service center other than Akça Machinery, telephone, fax or e-mail is used. Other service centers make a request in the same manner. The service centers that do not belong to Akça Machinery cannot see the inventory level at other service centers. The advantage of lateral transshipment is twofold: the unit cost of the item is the same as a normally ordered item while the request is treated as an expedited order. In case of emergency orders from the central warehouse of VOLVO, the standard unit cost is applied, where there are discounts (usually 10%) in case of regular ordering. On the other hand, emergency orders are expedited and they arrive in one day instead of a week, which is the standard for normal orders.

[2] The information on VOLVO's system is gathered by an interview with Ankara authorized aftersales service coordinator.

1.1 MOTIVATION OF THE STUDY

The motivation of this study comes from this increasing attractiveness in inventory pooling and increasing number of practical applications in different types of spare parts networks. The strategies of the dealers and the effects of **sharing system** (or **pooling system**) on their decisions play a key factor in customer satisfaction, and the success of the manufacturer in the long-run.

Although sharing resources (spare parts and repair service) with other dealers in the service network seems to be an attractive strategy to improve customer satisfaction and to increase profitability, in practice the firms usually adopt naive resource sharing strategies. Those naive strategies include, meeting the transshipment requests all the time or meeting them only if stock level is very high, placing a request only in case of a stockout, sharing information on stock levels only partially, and so on. These practices bring in several questions, such as whether resource sharing is always profitable than not sharing, what the right resource sharing strategy should be, and what the benefit of a fully transparent information sharing strategy across the spare parts network is. This study is motivated by these research questions. In this thesis, an analysis of benefits of inventory and service pooling is made. Besides resource pooling, the benefit of information sharing is also quantified to take a comprehensive approach to the resource sharing problem. In this study, it is assumed that information sharing could take place in one of the three levels: no-information sharing, partial-information sharing and full-information sharing. For no-information sharing, the dealers do not know the stock levels in the system except their own status. Full-information sharing indicates that the dealers have all the information about the stock levels in the system. Partial-information sharing is an intermediary information sharing level between no-information sharing and full-information sharing, where a dealer has incomplete information on the stock levels in the system, including own status. For instance, a case where a dealer knows whether he has positive stock or there is a stockout situation, and that dealer has all the remaining information fully about the stock levels in the system is an example for partial-information sharing. Another example is a dealer knows whether another dealer may accept his lateral transshipment request without knowing the other dealer's exact inventory levels and the dealer has all the remaining stock level information of the system fully. TOFAŞ case is also an example for partial-information sharing, where a dealer has only the information that the spare part exists or not at another dealer.

It is possible for an after-sales service provider to have lack of information on amount of own inventory levels or waiting customers. It is very common in practice that a company has a mismatch of inventory database with real levels. In a recent experience with a large plastics manufacturer having three distinct warehouses, among 50 randomly selected stock keeping units (SKUs), mismatch problems are detected with 10 of them. Inventory database gave a stockout for an SKU, although it was in stock. For three SKUs, the situation was just the reverse. Although database gave positive stock levels, these SKUs were not in stock. For the remaining six SKUs, quantity mismatches compared to the inventory database values were observed.[3] It may also be the case that an after-sales service provider is having lack of information on the number of waiting customers in the queue.

In order to clarify the structure of pooling system used in this study, a summary is as follows based on the discussions made previously. Both inventory and service pooling is assumed to take place, i.e. after-sales service providers may exchange items and redirect customers, which is equivalent to sharing capacity in the system. Information unavailability on own status (net inventory level) or status of another after-sales service provider is considered for the information pooling in this study.

1.2 OUTLINE OF THE STUDY

In this section, an outline of the study without the technical details is given to simplify readers' following-up of the thesis. In addition, more detailed outline is given about the main technical chapters.

Organization of the study is as follows. In Chapter 2, previous literature on inventory pooling in spare and service parts, inventory rationing, competition in spare parts

[3] The company name cannot be provided due to confidentiality policy of the company. The company is a successful Turkish company having innovation capabilities and performing research and development projects with international institutions. The management level was not expecting such a result as the company had invested in and has been using an Enterprise Resource Planning (ERP) System.

management systems, collaboration among competitors in spare parts, information sharing in operations management, and methodology for partially observable Markov decision processes is presented and this study is positioned in the literature.

In Chapter 3, under the assumption of a decentralized spare parts network consisting of two dealers, the benefits of resource pooling under full information sharing is analyzed. In the analysis, four main issues are considered, viz. monotonicity properties of the optimal policy, parameter effects on performance measures, effect of competition on resource pooling and comparison of pooling strategies. Monotonicity properties of the optimal policy refer to the monotonicity of the optimal control variables defining the policies of after-sales service providers with respect to the system state. In the decentralized system setting, independent after-sales service providers of the spare parts network maximize their own profits. The effects of system parameters, such as customer arrival rate, holding cost value, etc. on performance measures (profit, service level, flow rates and revenue components) are analyzed. In order to reveal the competition effect on resource pooling, a centralized system setting is also considered, where a central authority controls the spare parts network and the performance of the whole network is considered. For comparison of pooling strategies, four strategies considered, namely dynamic-pooling, static-pooling, full-pooling and no-pooling.

In Chapter 4, information availability effects on pooling benefits are analyzed for the two-dealer decentralized system. System is redefined under partial-information availability. Different information availabilities on both own status and other dealer's status are considered, which are no-information on other dealer's net inventory level, partial-information on own inventory status and partial-information on own customer status. The dealer's profit under full-information is compared with no-information and partial-information cases.

Conclusions and future research directions are given in Chapter 5.

A more detailed outline about the main technical chapters is as follows. Basically, this dissertation is about analysis on collaboration effects on spare parts management systems and is composed of two main parts, which are collaboration by means of inventory and service pooling (Chapter 3) and collaboration by means of information sharing (Chapter 4).

The chapter on analysis of inventory and service pooling benefits is organized as follows. Firstly, a decentralized setting is considered which is composed of two dealers. In Section 3.1, this spare parts system is defined and the modelling of the system using a Markov Decision Process approach is described. To simplify the analysis, the Continuous-Time Markov Decision Process is converted to the equivalent Discrete-Time Markov Decision Process under the discounted profit criterion using uniformization approach. In short, the optimal operating policy for one of the dealers (D1) given the other dealer (D2) is considered in this system. After this, the optimal policy is characterized for D1 under the decentralized setting. Conditions defining optimal operating policy of D1 are stated and it is proved that these conditions are these conditions form the necessary and sufficient conditions for the optimal monotone operating policy of D1. For the payment levels, it is proved that extreme values are more profitable for D1 compared to other intermediate values. Similar steps are followed for a centralized setting in Section 3.3. The optimal policy is characterized for the central authority and it is proved that a set of given conditions form the necessary and sufficient conditions for the optimal monotone operating policy. In Subsection 3.3.2, the behaviors of optimal policies under decentralized and centralized settings are compared. In Section 3.4, firstly the mathematical programming model used in the solutions is given, performance measures of the system are described and truncation of the state-space (which includes infinitely-many states) is explained. Then, the pooling strategies used under decentralized setting are defined (optimal-pooling, full-pooling, no-pooling and static-pooling) and numerical setting of parameters for decentralized and centralized settings are given. This section finalizes the chapter with the analysis of parameter effects and comparison of pooling strategies on the results of the numerical study.

The chapter on analysis of information sharing benefits is Chapter 4 and is organized as follows. In Section 4.1, observation process and internal process definitions are made, which are used to model the incomplete information situations. Two main information incompleteness models are described, viz. incomplete information of a dealer (D1) on own net inventory level and incomplete information on the other dealer's (D2) net inventory level. For D2's model, several information availability levels are described for partial-information cases, besides the extreme cases of no-information (a special case of partial-information) and full-information. The mathematical programming model used for solution of the partial-information models is given. For D1's model, two cases for information incompleteness on inventory status and customer status are explained. In Section 4.2, policies are compared for full-information and partial-information models to make the differences clear. In Section 4.3, firstly the solution strategy developed is explained, which is composed of non-linear solvers for the mathematical programming model and a neighborhood search algorithm. Then, numerical setting of parameters are given and two general observations on variable payment effect are stated. This section finalizes the chapter with the analysis of parameter effects on the information availability benefit for D2's model and comparison of inventory status and customer status cases for D1's model based on the results of the numerical study.

CHAPTER 2

LITERATURE REVIEW

In this chapter, the studies in the literature that are related with this study are summarized. The subjects of the papers, research questions that are investigated, their differences with the previous works, models built and main results are explained. In Sections 2.1, 2.2, 2.3 and 2.4, literature related to benefits of pooling is reviewed. Literature is considered under main headings as inventory pooling in spare and service parts, inventory rationing, competition in spare parts management systems, and collaboration among competitors in spare parts. Related literature of benefits of information availability is reviewed in Sections 2.5 and 2.6. Literature on information sharing in operations management is introduced first, and literature on methodology for partially observable Markov decision processes are given.

In the followings, the review of related literature is presented in six sections; 2.1, 2.2, 2.3 and 2.4 discuss previous work on inventory and/or service pooling, and 2.5 and 2.6 discuss information sharing topics.

2.1 INVENTORY POOLING IN SPARE AND SERVICE PARTS

Before starting to inventory and service pooling related literature, the two building-blocks should be mentioned, which are Sherbrooke [1968] and Muckstadt [1973].

One of the early works on service parts management is the METRIC model suggested by Sherbrooke [1968] where minimization of back-order levels in a two-echelon recoverable parts system is aimed under a budget constraint. METRIC had found many practical applications although it is not an exact model.

Muckstadt [1973] suggests a model called MODMETRIC to extend Sherbrooke [1968]'s work by including multi-indenture part structure. In both Sherbrooke [1968] and Muckstadt [1973], no lateral transshipment are considered and therefore inventory pooling is not made, and parts are assumed to be repairable.

Studies on multi-echelon setting (without lateral transshipment) still continue. One of the new research directions is to minimize inventory related costs. Wong et al. [2005a] is one of such works. They propose four heuristics and show their efficiencies using lower bounds that are generated by decomposition and column generation methods.

Lee [1987] extends Muckstadt [1973] and Sherbrooke [1968] by including lateral transshipment among members of pooling groups with the assumption of identical group members, which actually puts a restriction on pooling. Emergency lateral transshipment in continuous review systems with one-for-one replenishment and Poisson demand is considered. In case of an emergency transshipment, the source base in the group is selected randomly, which is a non-optimal decision rule. The model considered provides an approximation to expected level of backorders and number of emergency lateral transshipment.

Axsäter [1990] considers a similar setting to Lee [1987] including the random-selection decision rules for transshipment, and assumes that group members are not necessarily identical. The fraction of time demand is backordered, met from stock or met through emergency transshipment are evaluated and compared to values obtained from a simulation study. Results are closer to simulation results when proportion of emergency lateral transshipments is large. The author improves on the approximation technique used in Lee [1987] and shows this by obtaining better results under the same setting of Lee [1987]. Both Lee [1987]'s and Axsäter [1990]'s methods are approximate.

Jung et al. [2003] extend the infinite repair capacity and depot only repair assumptions of Lee [1987] and Axsäter [1990]. In their paper, both depot and bases have finite repair capacities and can repair spare parts, but only bases hold spare parts inventory. For minor failures, a failed part may be repaired at the base, but for more critical failures, it should be sent to depot. An algorithm is proposed for determining the optimal spare inventory level, where the objective is minimization of the total expected cost of the system.

The study by Dada [1992] is similar to Lee [1987], but differs in two ways. These are the use of exact model to provide a benchmark for approximate model, and use of error bounds by an aggregate model for system's approximate performance. Aggregation assumes combining all centers into an aggregate center and aggregate solution is disaggregated to approximate the performance of the exact system.

Alfredson and Verrijdt [1999] relaxes the two assumptions of Dada [1992] about local warehouses, viz. identical replenishment lead times from central warehouse and single unit stocking at all local warehouses. Deterministic, exponential and lognormal lead times are used in the study. They also use aggregation and disaggregation in solution method. Their main findings are two-folds: emergency supply strategy always pays off and lead time distribution does not change results significantly, which indicates exponential distribution can safely be used in their study. By using exponentially distributed lead times, they are able to obtain exact analytical results, like many other researchers.

Herer et al. [2002] propose using lateral transshipment instead of product postponement[1] and propose a cost-effective and customer-responsive solution. The solution has lean and agile properties and called as leagile solution. The authors show that transshipments can decrease the overall inventory levels and improve service, and further improvements can be achieved in non-identical retailers environment by using subgroups.

Herer and Rashit [1999] consider a two-location inventory system with lateral transshipments in a single period. The authors assume nonnegligible fixed and joint replenishment costs and analyze their impact on the optimal replenishment policy. The main finding of the study is that the optimal replenishment policy is not of an min-max form unlike the classical newsboy problem.

Herer and Tzur [2001] assume dynamic deterministic demand and consider a two-location inventory system with lateral transshipments over finite horizon. They as-

[1] Product postponement, or product differentiation postponement, is an attempt for delaying some production activities until the customer demand is revealed in order to have fast customer-response and lower system costs by redesigning the product or production processes. Product design can be made for using common parts in many different end-products. By using common parts, number of stock-keeping units are reduced and only few component inventories are held, and products are stocked as semi-finished forms. Final assembly of the product, which is the last stage of production, can even be carried out at the distribution center. Therefore, cycle time is reduced and fast customer-response is achieved, with the expense of redesigning the product.

sume nonnegligible fixed and joint replenishment and transshipment costs. They propose a polynomial time algorithm to find the optimal replenishment and transshipment policy.

Herer et al. [2006] consider a multi-retailer setting and find the transshipment and replenishment quantities that minimize the expected long-run average cost over an infinite horizon. The authors show that order-up-to policy is optimal and calculate order-up-to levels using infinitesimal perturbation analysis method.

Kranenburg and van Houtum [2009] model a real-life inventory control problem as a network consisting of two types of local warehouses, viz. main and regular. In this multi-item and multi-location system, lateral transshipment is allowed from main local warehouses only, which lead to partial-pooling. Even if the number of main warehouses is small, the authors show that the system benefits are close to full-pooling setting. The authors use a heuristic algorithm which uses an accurate and fast approximate evaluation method for the optimization of base stock levels. The algorithm is implemented by ASML, the company owning the real-life problem, and substantial improvement in terms of waiting times and cost are realized. Authors show that even partial pooling results in significant savings.

van Wijk et al. [2009] characterize the structure of the optimal lateral transshipment policy in a two-stock point system providing spare parts with repairing. They define sufficient conditions under which rationing (they call as hold-back) or complete-pooling policy is optimal.

Research in Wong et al. [2006] is motivated by an air carrier company that wants cooperation in spare parts. They use a centralized model and use the company data to model a two identical carrier system in order to estimate potential benefits of pooling. Under waiting time constraints and under the assumption of complete pooling, the authors determine optimal base-stock levels of a two-location system. A local search for the multi-item problem that uses initial solution obtained from Lagrangian relaxation and a sub-gradient method is used. The authors find that cost savings are significant by pooling and increase with shorter transshipment lead-times and tighter service constraints. Wong et al. [2006] also provides an overview lateral transshipment literature, where literature is categorized in terms of the number of echelons (single, two or multi), the number of items (single or multi), review (periodic or continuous), inventory policy, and the type of analysis done(exact or approximate and evaluation or optimization).

Wong et al. [2005a] analyze pooling of repairable spare parts in a single echelon setting where lateral transshipments may be delayed, have non-zero times and occur by proximity rule. They propose an approximation method for service level calculation and show its accuracy and efficiency by numerical experiments. The authors also show that delayed lateral transshipments can reduce the expected number of backorders significantly and instantaneous lateral transshipment can lead to non-optimal stocking decisions.

2.2 INVENTORY RATIONING

In this study, a single type of customer is assumed to arrive to the system. A dealer faces with this customer demand, as well as other dealers' lateral transshipment requests, which is another type of demand for the dealer. Consequently, in this dissertation, a dealer faces with multiple customer classes. This situation brings the inventory rationing concept and therefore the related literature is reviewed.

Topkis [1966] is one of the early works on stock rationing, which assumes only two different demand classes. Kaplan [1969] follows this work by changing lost sales assumption with backlogging. Both studies assume that there are two different types of demands. Kaplan [1969] assumes periodic review and models the rationing level as a function of time to replenishment. Optimal rationing level, which he calls as scheduled reserve level, decreases by decreasing time to replenishment. He not only makes a computational experiment but also performs an analysis based on the real data from US Army Materiel Command.

Topkis [1968] extends the assumption of Kaplan [1969] to multi-demand case and considers both lost sales and backlogging cases. He found that calculation of optimal rationing levels require much more effort for backlogging case. Although the system is assumed to be single period where procurement is only once made at the beginning, the analysis is made by dividing the period into intervals. The results are extended to a multi-period model. He concludes the study with analysis on some myopic policies and gives conditions under which the myopic policies are optimal for multi-period model.

Ha [1997a] considers stock rationing problem of a manufacturer of single item in make-to-stock system with several demand classes and lost sales. Optimal production policies (i.e. production manager decides on *continue production* or *stop production*) and rationing policies (i.e. sales manager decides on *serve customer* or *reject customer*) of a single manufacturer with a single product and multiple customer types are investigated. Results show that rationing (threshold) levels for customer types are increasing with decreasing lost sales costs (for a customer type, if inventory is under rationing level, reject customer). There exists a base stock level and base stock and rationing policies are stationary. Cost reduction by using such a policy instead of First-Come-First-Serve (FCFS) system is increasing with increasing lost sales costs ratio.

Ha [1997b] is similar to Ha [1997a], but two types of items are produced for stock. In his study, he identifies switching curve for production, which indicates optimal production scheme, i.e. produce product type one, two or do not produce at all. He also compares three priority rules with optimal production scheme defined by the switching curve.

de Véricourt et al. [2000] partially characterize optimal hedging point policy for rationing the production capacity between two products. Partial characterization of the structure of the optimal hedging policy in a certain region of the state space is a straight line defining the monotone switching curve that rations production of the products, which generalize the results of Ha [1997b].

Nahmias and Demmy [1981] consider single and multi period periodic review, and also continuous review structures. They concentrate on expected backorder rates instead of optimal policies derived in the previous studies. In their numerical experiments, for given inventory control policies, they calculate backorder rates for high and low priority demands.

Rationing level in lateral transshipment related studies may be zero or positive, i.e. an actor (after-sales service provider, dealer, participant, etc.) may share his all inventory with another actor at the same echelon level or may spare some items for expected future own-customers. When this level is zero, rationing is not actually made between different demand classes. Pooling policy of such studies are called as complete or full pooling. On the other hand, in the studies where positive rationing levels are used, rationing is not made necessarily optimal. Indeed, in most of the studies, rationing is not optimal. Such pooling policies are called as partial pooling or hold-back inventory policies. For example, in Çömez et al. [2007] and Xu et al. [2003], rationing level is called as hold-back inventory level. The studies that compare complete and partial pooling mostly end up with the conclusion that full pooling is better compared to partial pooling, for example Tagaras and Cohen [1992] and Wong et al. [2007], where the latter makes this conclusion for for centralized setting. In this dissertation, it is shown that full pooling may perform worse than rationing.

Axsäter [2003] proposes a decision rule for lateral transshipment for parallel local warehouses in a single-echelon system that face with compound Poisson demand. While forming the decision rule, no further transshipments after some specific time is assumed. Expected savings for a lateral transshipment is calculated under this assumption and maximum saving by transshipment is used as the decision rule (maximizer may be transshipment of all demand, part of it or even none), which is then used repeatedly as a heuristic. Local warehouses are replenishing from an outside supplier and reorder policies are found by simulation, except no-pooling case. With a numerical study, decision rule is compared with a simple alternative heuristic and no-pooling. In the numerical study, several scenarios are considered, viz. two and three local warehouses, lower and fixed transshipment costs, and using no-pooling ordering policy for warehouses. Results support that the decision rule performs very well and it is especially valuable in more complex decision situations compared to simple alternative heuristic.

In Kukreja et al. [2001], stocking decisions for low-usage and expensive items in a multilocation inventory system is modelled and a real life application at a large utility company is made. The benefit of transshipment among inventory locations (for real life application, these are power plants), i.e. inventory pooling, compared to a system where inventory locations have individual inventory strategies (without considering inventory transshipment) is questioned. Inventory pooling policies considered in the study are simple policies and are not required to be necessarily optimal. The study is similar to Axsäter [1990], but the selection of plant of a stock out plant for lateral transshipment is made according to transshipment costs with all other plants. In this way, Kukreja et al. [2001] alters the random decision rule of Axsäter [1990], which is one of the non-optimal decision rules, with making the lateral transshipment from the lowest-cost source and finds cost-effective stocking levels. Another similarity is usage of the time fractions (demand met by on-hand stock, by lateral transshipment or backordered). The numerical study on part of the real life data is made and cost savings around 70% are observed.

Grahovac and Chakravarty [2001] is another study similar to Axsäter [1990]. The subject is the benefits of sharing and lateral transshipment of low-demand and expensive items in a vertically integrated (or autonomous) retailers via one central depot. It considers a system with low-demand items with transshipment to show the benefit of inventory pooling for both centralized and decentralized settings. In this study transshipment policies are not necessarily optimal. Similar to Kukreja et al. [2001], simple and not necessarily optimal inventory pooling policies are considered. There are two important differences that differentiates the study from Axsäter [1990], viz. individual locations of the lowest-echelon can face with different levels of demand and emergency orders are allowed not only for stockout but also for arbitrarily chosen levels of net stock. The assumptions of the model are more representative of a commercial supply chain, where the model in Axsäter [1990] is more descriptive of military systems. Main findings are as follows: In a centralized supply chain, in case of lateral transshipment, retailer stocking levels are at least equal to and distribution center (DC) stocking level is at most equal to those in without lateral transshipment. In few cases, overall inventory level is larger in case of lateral transshipment. Free riding by the distributor without lateral transshipment is more severe. For larger relative shares of customer waiting and emergency transshipment costs of retailers, they tend to be more motivated for sharing and transshipment of inventory, and DC is just the opposite. Overall costs are reduced by around 20% by lateral transshipment, but this savings is not accompanied by reduction in overall inventory level.

2.3 COMPETITION IN SPARE PARTS MANAGEMENT SYSTEMS

In this dissertation, a decentralized system is considered with two dealers. Due to the decentralization, competition occurs, which leads to a review on competition in spare parts management systems.

Two competing dealers that maximize their individual profits and collaborate through lateral transshipments is assumed in Çömez [2009]. A single sales season is studied which is divided into periods. A Game Theoretic setting is used to model the problem. A Generalized Nash Bargaining formulation is used for optimal transshipment policies, which are shown to be dynamic and a Non-Cooperative Cournot Game is used to determine optimal order quantity decisions, which are shown to form a pure strategy equilibrium provided that order quantities can be non-integer. An important result is that retailers always benefit from optimal transshipment policies compared to complete or no pooling strategies.

Karsten et al. [2009] define a set of companies separately stocking spare parts of the same item, where the companies are independent and maximizing their own profits. They analyze effects of pooling in such a system in a cooperative cost game context. They show that the core of the game is non-empty, i.e. a stable cost allocation exists for collaboration. For non-identical demand rates, base stock levels and downtime costs, core is non-empty for most of the cases, but there are exceptions for some combinations of those conditions.

Rudi et al. [2001] considers a decentralized single-period news-vendor model where the lateral transshipment price is negotiated. The model represents the real-life problem of independently operating Bosch automotive parts distributors of Norway. Unique Nash equilibrium is shown to exist for the decentralized model and the behavior of optimal policy with respect to transshipment price is characterized. Comparing optimal centralized and decentralized solutions, authors show that transshipment prices can be adjusted to achieve the centralized solution.

Zhao et al. [2005] is one of the studies that consider a two dealer decentralized model, which is about a decentralized dealer network in which each independent dealer is given the flexibility to share his inventory. Dealers are assumed to make replenishment from a manufacturer in constant lead times. The strategies of dealers under full or fixed portion sharing and rationing conditions, the affect of dealers' decisions on each other and end customers, and manufacturer's impact on these strategies via incentives and subsidies are the research questions. A two-dimensional strategy (base-stock and threshold-rationing levels) is used for the dealers. Steady-state probabilities of dealers for inventory levels are used to calculate expected cost functions. For three different situations, the strategies of dealers are investigated, which are full sharing (threshold-rationing level is zero and base-stock level is the strategy), fixed sharing (for a given threshold-rationing level, base-stock level is the strategy) and inventory rationing (for a given base-stock level, threshold-rationing level is the strategy) games. Results are somehow surprising. Cost function cannot be shown to be supermodular. The authors are able to show monotonicity for inventory sharing game, but not for inventory rationing game. This does not mean that there is no equilibrium solution for the games. Nash equilibria are checked using an extensive numerical study (including more that a thousand instances), which shows that for most of the cases equilibrium solutions exist and no equilibrium is observed in very little cases. Main findings are as follows: Dealers respond to higher incentives by decreasing their threshold-rationing levels rather than increasing their base-stock levels, manufacturer subsidies increase backorders (meaning worsens customer service level, where it is just the reverse for incentives), inventory sharing in decentralized system for very expensive items increases backorders (for other items, backorders decreased, as it is always the case for centralized system).

Optimal inventory transshipment policies in a decentralized dealer network is subject of Zhao et al. [2006], who were also able to characterize the optimal operating policy. In the paper, decisions of both the sending and accepting a request are included. The study is similar to Zhao et al. [2005] and the research questions are the same, except requesting decision (i.e. sending a request to another dealer for lateral transshipment) and demand filling decision (i.e. accepting a lateral transshipment request made by another dealer) of the dealers. Mainly, two networks are analyzed: with a large number of dealers and with two dealers. Dynamic programming is used for the former for optimal policy determination and a game theoretic approach is used for the latter for finding equilibrium conditions. When there are large number of dealers, the effects of the actions of one dealer on others is negligible, so other dealers' actions are considered as exogenous constants. Inventory process of a dealer is modelled as a continuous time Markov chain, and uniformization is used to transform the process into the equivalent discrete time Markov chain. Based on steady state probabilities, a long-run average cost function is used and a search algorithm is run to find optimal decisions. For two dealers case, requests from each other is modelled endogenously. Existence of equilibrium is shown given some levels of decisions. An approximate expected cost function is used and an iterative method is used to find optimal levels. Extensive numerical study is used for two dealers case and four different policy types are investigated, viz. only base-stock, base-stock and rationing, base-stock, rationing and requesting policy under decentralized network and base-stock, rationing and requesting policy under centralized network. The key findings are as follows. With increasing transshipment cost, dealers stock more and share less, where high-volume dealers respond more to an increase. An increased transshipment incentive makes dealers to decrease both stocking and rationing levels in the two dealer network, whereas in a large network stocking levels are increased (rationing levels are decreased as in the former). The dealers are very sensitive even the incentive is very small, which is meaningful in practice. Including a requesting threshold in a base-stock and rationing policy in a decentralized network makes costs, stocks and backorders less in equilibrium. In Table 2.1, the two parts of this study (pooling benefits in the presence of full information availability, and benefits of information availability in the presence of pooling) are included and information about them is given.

The demand side and dealer relations of Zhao et al. [2005] and Zhao et al. [2006] are similar with the ones in this study (two demand classes are considered, one for own customer and one for the other dealer, and lateral transshipment demands are endogenously modelled). Each dealer uses a base-stock and threshold-rationing policy for his inventory-stocking and inventory-sharing decisions in Zhao et al. [2005] and requesting threshold level in addition to them in Zhao et al. [2006], as in this study. In the first part of Zhao et al. [2006], a dealer knows the fulfilment possibility for a lateral transshipment request based on his experiences. In this study, a dealer is assumed to know the other dealer's situation in terms of inventory and/or customer. The main difference in supply side is manufacturer's exponential lead time, which can process one unit at a time. Zhao et al. [2005] assumes manufacturer's lead time to be constant. In Zhao et al. [2006], each dealer has a production facility associated with and production times are exponential, which is also assumed in this study. Actually, this is an important difference both in terms of practical meaning and modelling complexity.

Zhao et al. [2008] aim to characterize the optimal operating policies in a centralized dealer network. A network with two dealers which are linked to make-to-stock production facilities is assumed. A centralized system is considered and an expected discounted system cost is minimized. Lateral transshipments are allowed both before and after the demand realization in this study, which differs from the previous literature. State is represented by two dimensions, viz. the inventory levels of each dealers. Structure of the optimal policy is analyzed which includes three control variables, namely order-up-to level, production transshipment and demand-filling transshipment. The authors show that optimal base-stock level and rationing levels dynamically change depending on the inventory status at the dealers, i.e. with the increasing inventory level of one dealer, other dealer's order-up-to level decreases, while other control variables increase. Two heuristic methods are compared with optimal policy, since they claim that the optimal policy is not easy to implement in practice. Optimal values of control variables are found (i.e. the optimal policy is defined) by search algorithm, but also a newsvendor heuristic is proposed for finding the values of control variables.

Leng and Parlar [2005] make a review on game theoretic applications in supply chain management. For a detailed classification of the related studies, the paper can be referred to.

2.4 COLLABORATION AMONG COMPETITORS IN SPARE PARTS

There exist a body of work on production and inventory management of spare parts in the literature, but most of the studies consider spare parts from a single organization's perspective. Moncrief et al. [2005] provide practical solution techniques for spare parts management problems, such as their stock management (re-order point calculation, economic order quantity analysis,etc.), getting rid of excess spare parts inventory (identifying excess inventory items, avoiding unnecessary purchases, etc.), and setting and monitoring inventory goals (defining related measure and methods of measurement). As spare parts are mostly expensive and and used infrequently, the success of methods considering a single organization are limited. Collaboration for management of spare parts among similar organizations, and even competitors, brings significant management advantages to the companies. In a short chapter, Moncrief et al. [2005] provide collaboration examples with distributors, original equipment suppliers, consignors and others (enterprize software providers, parts finding firms, stocking consultants). The authors quantify the benefits using real-life cases.

Cachon [2002] analyze inventory management in a competitive supply chain. Using a game-theoretic approach, even in a simple single-supplier and single-retailer environment, the agents are shown to have incentives to deviate from the Nash equilibria. In supply chains that re composed of independent agents, a centralized solution (that optimize overall performance) cannot be achieved due to competition. Some coordination techniques (various forms of transfer payments and imposing service constraints) are shown to manipulate the competitive behavior and improve the overall solution.

Keskinocak and Savaşaneril [2008] model a single-period collaboration efforts of two competitive manufacturers, which make joint-procurement. They analyze profitability of the buyers (i.e. manufacturers) and supplier. They are able to characterize conditions of profitable collaboration for a buyer, where procurement is assumed both capacitated and uncapacitated. Buyers are also modelled as the same size and different sizes, where in the latter case, bigger buyer is shown to have more impact on the collaboration process.

Çömez et al. [2007] consider a two dealer system operating under periodic review policy where the system is controlled by a central authority and dealers ration the stock between own-customer demand and other dealer's demand. A cycle is composed of periods and multiple lateral transshipments are allowed within a cycle. Positive transshipment and replenishment lead times are assumed. They characterize optimal policy and show that rationing level decreases by increasing number of periods to the next replenishment. They also propose a heuristic and show its efficiency by comparing with optimal solutions.

Wong et al. [2007] study a spare parts inventory system with lateral transshipments, where parts can be repaired and delayed lateral transshipments are possible. Using a game-theoretic approach, the authors show that there are cost allocation policies for decentralized setting which are acceptable for all participants. The authors also give an example about how false information causes the companies to become worse-off to show the importance of building mutual trust between the cooperating companies.

In spare parts management literature, the main characteristics of the studies are presented below.

Objective Formulations are with cost/profit driven objectives or service-level driven objectives. In some formulations, cost/profit driven objectives with service-level constraints or service-level driven objectives given budget restrictions are used.

Review Periodic review or continuous review are assumed.

Echelon Spare parts management system under consideration is either a single-echelon or multi-echelon system.

Item The analyses are made under single-item or multi-item setting. There exist studies that consider multi-component or multi-indenture items.

Period According to the number of periods considered, single-period or multi-period models exist.

Customer In the studies, single-customer class or multi-customer classes are assumed. Depending on the setting of the problem, a service center may face a single customer class or multiple customer classes. If there exist customers that bring different cost/revenue to the service center, then the study considers a multi-customer class setting. This may be the case if there exist several service centers and customers from other centers are regarded as lower-priority classes, even a single type of customer arrives to the system.

Demand Endogenous or exogenous stochastic demand models exist. If demand is subject to changes due to interaction of decisions among after-sales service providers, such models are assumed as endogenous models. In exogenous demand models, the demand that a location faces is independent of the decisions made in the system.

Transshipment Transshipment is solely vertical or vertical with lateral transshipment. In some studies, transshipment is irrelevant, which can be considered as no transshipment.

Authority Centralized or decentralized systems are considered. In some decentralized systems, affect of a central authority on after-sales service providers via interventions or subsidies are modelled. Both systems are considered for comparison purposes in some studies.

Optimality There exist studies that either seeks for the optimal policy and/or propose heuristic policies or simply evaluates the system performance under a given policy. Searching, iterative and exact procedures are used for optimal policy algorithms. Simulation models are regarded as heuristic policy seeking algorithms. Studies that evaluate or approximate performance measures for given settings (policies) are performance evaluation type.

Information Partial, full and approximate information availabilities are considered. The studies where after-sales service providers know each others net inventory levels are regarded as full information. Partial information is assumed as information incompleteness. Using previous experiences of after-sales service providers for estimations is regarded as approximate information availability.

Table 2.1 gives a taxonomy of the studies included in the literature review up to this point based on the the main characteristics of the studies proposed above. Kennedy et al. [2002] is a review paper about spare parts inventories literature, emphasizing the unique properties of spare parts (which are different than work-in-process or finished-goods inventories, such as repair or replace decision, dependent part failures, unplanned repair requirements, etc.) and classifying the recent literature on the topic. Paterson et al. [2009] makes a literature review for inventory models with lateral transshipments. Interested readers may refer to those reviews.

Table 2.1: Taxonomy of Related Literature

STUDY	CHARACTERISTICS											
	OBJ	REV	ECHL	ITEM	PERD	CUST	DMND	TRANS	AUTH	OPTI	INFO	LOC
Alfredson and Verrijdt [1999]	cost/prof.	cont.	multi	single	multi	single	exo.	lateral	centr.	eval.	full	multi
Axsäter [2003]	cost/prof.	cont.	single	single	multi	single	endo.	lateral	centr.	heur.	full	multi
Axsäter [1990]	cost/prof.	cont.	multi	single	multi	single	exo.	lateral	centr.	eval.	full	multi
Çömez et al. [2007]	cost/prof.	perd.	single	single	multi	single	endo.	lateral	centr.	optim.	full	multi
Çömez [2009]	cost/prof.	perd.	single	single	multi	multi	endo.	lateral	decentr.	optim.	full	multi
Dada [1992]	service	cont.	multi	single	multi	single	exo.	lateral	centr.	eval.	full	multi
Grahovac and Chakravarty [2001]	cost/prof.	cont.	multi	single	multi	single	exo.	lateral	both	optim.	aprx.	multi
Ha [1997a]	cost/prof.	cont.	single	single	multi	multi	exo.	no	centr.	optim.	full	single
Ha [1997b]	cost/prof.	cont.	single	multi	multi	multi	exo.	no	centr.	optim.	full	single
Herer and Rashit [1999]	cost/prof.	perd.	single	single	single	single	endo.	lateral	centr.	optim.	full	multi
Herer and Tzur [2001]	cost/prof.	perd.	single	single	multi	single	exo.	lateral	centr.	optim.	full	multi
Herer et al. [2006]	cost/prof.	perd.	single	single	multi	single	exo.	lateral	centr.	optim.	full	multi
Herer et al. [2002]	cost/prof.	perd.	single	single	single	single	exo.	lateral	centr.	optim.	full	multi
Jung et al. [2003]	cost/prof.	cont.	multi	single	multi	single	exo.	lateral	centr.	eval.	full	multi
Kaplan [1969]	cost/prof.	perd.	single	single	multi	multi	exo.	no	centr.	optim.	full	single
Karsten et al. [2009]	cost/prof.	cont.	single	single	multi	multi	endo.	lateral	decentr.	optim.	full	multi
Kranenburg and van Houtum [2009]	cost/prof.	cont.	single	multi	multi	multi	endo.	lateral	centr.	eval.	full	multi
Kukreja et al. [2001]	cost/prof.	cont.	single	single	multi	single	endo.	lateral	decentr.	optim.	full	multi

Table 2.1 (continued)

STUDY	CHARACTERISTICS											
	OBJ	REV	ECHL	ITEM	PERD	CUST	DMND	TRANS	AUTH	OPTI	INFO	LOC
Lee [1987]	cost/prof.	cont.	multi	single	multi	single	exo.	lateral	centr.	optim.	full	multi
Muckstadt [1973]	service	cont.	multi	multi	multi	single	exo.	vertical	centr.	optim.	full	multi
Rudi et al. [2001]	cost/prof.	perd.	single	single	single	multi	endo.	lateral	both	optim.	full	multi
Sherbrooke [1968]	service	cont.	multi	multi	multi	single	exo.	vertical	centr.	optim.	full	multi
Topkis [1968]	cost/prof.	perd.	single	single	single	multi	exo.	no	centr.	optim.	full	single
van Wijk et al. [2009]	cost/prof.	cont.	single	single	multi	single	endo.	lateral	centr.	optim.	full	multi
de Véricourt et al. [2000]	cost/prof.	cont.	single	multi	multi	multi	exo.	no	centr.	optim.	full	single
Wong et al. [2005b]	cost/prof.	cont.	multi	multi	multi	single	exo.	vertical	centr.	heur.	full	multi
Wong et al. [2006]	cost/prof.	cont.	single	multi	multi	single	endo.	lateral	centr.	optim.	full	multi
Wong et al. [2007]	cost/prof.	cont.	single	single	multi	single	endo.	lateral	both	optim.	full	multi
Wong et al. [2005a]	service	cont.	single	single	multi	single	exo.	lateral	centr.	heur.	full	multi
Zhao et al. [2006]	cost/prof.	cont.	single	single	multi	multi	both	lateral	decentr.	optim.	aprx.	multi
Zhao et al. [2005]	cost/prof.	cont.	single	single	multi	multi	endo.	lateral	both	optim.	aprx.	multi
Zhao et al. [2008]	cost/prof.	cont.	single	single	multi	multi	exo.	lateral	centr.	optim.	full	multi
PhD Dissertation-Chapter 3	cost/prof.	cont.	single	single	multi	multi	endo.	lateral	both	optim.	full	multi
PhD Dissertation-Chapter 4	cost/prof.	cont.	single	single	multi	multi	endo.	lateral	decentr.	optim.	partl.	multi

(Abbreviations of characteristics' headings are Objective, Review, Echelon, Item, Period, Customer, Demand, Transshipment, Authority, Optimality, Information and Location; in order of appearance. Abbreviations of characteristics are as follows: prof. =profit, cont.=continuous, perd.=periodic, exo.=exogeneous, endo.=endogeneous, centr.=centralized, decentr.=decentralized, optim.=optimization, eval.=evaluation, heur.=heuristic, partl.=partial, aprx.=approximate)

Three of the studies in the literature discussed up to this point are closer to this study than the others, namely Zhao et al. [2006], Grahovac and Chakravarty [2001] and Zhao et al. [2008]. The main differences of this study with those three are as follows. Zhao et al. [2006] and Grahovac and Chakravarty [2001] calculate steady-state distributions and derive performances thereout. In all those three studies, Zhao et al. [2006], Grahovac and Chakravarty [2001] and Zhao et al. [2008], and in this dissertation, cost driven objective is used for a multi-period horizon under continuous review in a multi-location network for a single item, where lateral transshipments are allowed. Throughout this study, a single-echelon system is considered. Grahovac and Chakravarty [2001] is the only one using multi-echelon setting. For the customer types that the supply chain participants confront with, all studies except Grahovac and Chakravarty [2001] assume multi-customer type, including this study. In Zhao et al. [2006], Grahovac and Chakravarty [2001] and Zhao et al. [2008], operating policies of dealers are modelled endogenously, i.e. the interaction between dealers are included. In this study, a decentralized system is basically considered, while centralized system is also considered for comparison purposes in a limited scope. Both centralized and decentralized systems are included in Grahovac and Chakravarty [2001] and comparisons are made. Zhao et al. [2008] use centralized setting, while Zhao et al. [2006] use decentralized setting. All use optimization for obtaining solution. Except this dissertation, Zhao et al. [2006], Grahovac and Chakravarty [2001] and Zhao et al. [2008] assume production systems and transship only items. In this study, customers are assumed to be redirectable and service is also transshipped in addition to production.

Analysis of benefits of information sharing is another main difference of this study. In Zhao et al. [2008], both dealers' inventory levels are included in state representation, therefore full-information is used. Partial-information is used in Grahovac and Chakravarty [2001], which is represented by probabilities. In Zhao et al. [2006], approximate information is used for acceptance of a dealer's transshipment request by another dealer, based on the experience of the dealer. In this study, full-information and partial-information are used and partial-information refers to incompleteness of information, i.e. observability problems.

Above comparisons are made with the most related studies and aimed to position this study. In Section 2.7, the set of closer studies is extended and positioning of the study is elaborated, and contributions of this study with respect to the positioning is made.

2.5 INFORMATION SHARING IN OPERATIONS MANAGEMENT

Information flow over a supply chain is one of the most important factors affecting the efficiency and effectiveness of the chain. Information can flow on vertical and horizontal dimensions, where the former is between suppliers and customers and the latter is between the same level participants. Vertical flow may be upstream or downstream, i.e. in the direction to the supplier or to the customer, respectively.

Bullwhip (or whiplash) effect is an important concept defining the value of information sharing in supply chains. Lee et al. [1997] defines bullwhip effect as the inefficiencies (such as excessive inventories, insufficient capacities, poor customer service, increase lost sales, etc.) in a supply chain due to distortion of information in the vertical dimension. Pileup accidents is a good example to describe the bullwhip effect. Assume a pool of multiple vehicles cruising on highway and consider the three consecutive vehicles in front of the pool. When the vehicle in front of the pool puts on the brakes, the second vehicle also puts on the brakes, but more aggressively due to the lost time passes until the reflex action. It is important how the third vehicle reacts to this event. If the driver of the third vehicle is able to view the first vehicle's stop lights, the driver can have a reflex action similar to the driver of the second car. If it is not possible to view the first vehicle's stop lights (the second vehicle is a truck, for example) or the driver of the third vehicle concentrates on the second vehicle's stop lights, the only available information to the third vehicle's driver is the second vehicles's action. Possibly the third vehicle's driver puts on the brakes harsher than the second vehicle's driver. In this case, the information on the vehicle in front is not shared with the third vehicle. If all the vehicles in the pool have the same situation, pileup accident is inevitable.

Lee et al. [1997] identify four causes of bullwhip effect and propose solutions, such as avoiding multiple demand forecast updates, breaking order batches, stabilizing prices and eliminating excess ordering in shortage situations. These solutions include information systems applications, electronic data interchange (EDI) and computer assisted ordering (CAO). Laudon and Laudon [2006] emphasize the role of information systems to cope with bullwhip effect and explain their usages.

Five information sharing scenarios are considered in single supplier-single retailer supply chain by Zhao and Qiu [2007], viz. centralized information sharing, full information sharing, supplier-dominated information sharing, retailer-dominated information sharing, and not sharing information at all. For centralized setting, Clark and Scarf [1960] is followed. All other information sharing scenarios belong to decentralized setting, and discrete time Markov decision process is used for modelling. With a numerical study, they show that the cost of a centralized inventory system is about 20% - 40% lower than that of a decentralized system where information is not shared. They also find that in a decentralized supply chain, a higher information sharing level does not always lead to a lower system cost due to inventory competition.

Fransoo et al. [2001] and Mitra and Chatterjee [2004] model two-echelon systems. Fransoo et al. [2001] define supply chain as a subset of supply web. Coordination between supply chains brings better utilization of common resources that those supply chains use. Resource planning function, that has the authority over a shared resource, manages the interactions of supply chains. In numerical studies, they use one manufacturer and four-retailer supply web. Two independent supply chains (manufacturer plus combinations of four retailers), one cooperating and the other non-cooperating, are coordinated via resource planning function to use the manufacturer (common resource that holds and supplies inventory) efficiently. Resource planner pools inventory without use of local information of supply chains. They show that coordination brings decrease in total system inventory.

Mitra and Chatterjee [2004] assume a one-warehouse two-retailer system operating under periodic-review and examine the effect of utilizing demand information in this system. In addition to regular shipments from warehouse, both emergency shipments from warehouse and lateral transshipment among retailers is assumed for excess demand. They show that using actual demand information at the warehouse and dynamically setting the order-up-to level of the warehouse, total cost of the system can be reduced.

Herer et al. [2006] is dealing with information sharing in the horizontal dimension.

They define information pooling as the practice of transshipment transparency for visibility of monitoring the movement of material between locations at the same echelon. They indicate such information pooling through transshipments has been less frequent.

Yan and Zhao [2008] is the first paper studying inventory sharing systems with asymmetric information. For two retailers which share inventory but not the demand information, they consider a one-period setting under different levels of information sharing. An interesting result is that sharing demand information does not always benefit retailers. They consider a coordination mechanism and perform incentive analysis under this mechanism to share information.

2.6 METHODOLOGY FOR PARTIALLY OBSERVABLE MARKOV DECISION PROCESSES

In this dissertation, a decentralized system under partial information is modelled as a Partially-Observable Markov Decision Process (POMDP) in order to analyze the benefit of information sharing. In literature, POMDPs are also called as **hidden** or **constrained** Markov decision processes. Fraser [2008] explains that hidden Markov model refers to internal process (or original process) when a particular state of observation process is not sufficient to specify the state of internal process. In other words, for each state of internal process, observation process can be at several states with related probabilities. Fraser [2008] discusses solution algorithms and continuous state problems for hidden Markov models.

Partial information availability leads to partially observable systems. Smallwood and Sondik [1973] is one of the earliest works on the subject. They use a machine-maintenance example and demonstrate that optimal payoff function is piecewise-linear and convex if finite number of control intervals remain. An algorithm for this condition to calculate the payoff function and the optimal control policy is given. Both discounted and undiscounted costs for a finite planning horizon is considered in the study.

Sondik [1978] extends Smallwood and Sondik [1973] to the infinite planning horizon for the case of discounted costs. In both of these studies, an observer does not directly observe the process, but observes an output based on the conditional probabilities corresponding to outputs for the given current state of the process is. Bayesian update is used for deriving steady state probabilities after a transition (with the selected alternative) given the observation. In this model, the steady state probabilities over the continuous state space of form a Markov process.

White [1976] propose a numerical procedure for finding the optimal solution to a partially observed system over finite horizon. Partial observation is defined as noise-corrupted or imperfect observation in the study. The solution method of Smallwood and Sondik [1973] is extended to semi-Markov processes.

Although the method of Sondik [1978] has attracted interest in the literature, the method is not used in this dissertation, because the solution procedure is complicated and for a large number of states, application of the method for finding the optimal policy is not easy. This important disadvantage of the solution procedure leads using another method in this dissertation.

Serin [1989] models both discounted and average cost settings under partial observability. For the solution of the nonlinear system, an algorithm is proposed and several examples are solved via the algorithm. For the mathematical programming model of the partial information model of this dissertation, the mathematical programming model proposed by Serin [1989] is used.

Serin and Kulkarni [2005] consider average cost over infinite horizon and proposes a decision rule for the partially observable system depending only on the observable process. After providing the mathematical programming model, which is nonlinear, they propose a policy iteration algorithm for solution.

2.7 CONTRIBUTIONS OF THIS STUDY

In order to elaborate the positioning of this study within the literature and consequently to clarify the contributions of this study, closer studies to this study among the reviewed literature are selected and comparisons are made. In the comparisons, six criteria are considered for inventory and service pooling (information sharing), viz. analysis and quantification of benefits of pooling (benefits of information), comparison of pooling strategies (comparison of information availabilities), information used in modelling, policy type considered and comparison of authority types, where these criteria are different than the characteristics defined for taxonomy of the related literature, except the information criterion. The comparisons are summarized in Tables 2.2 and 2.3 for benefit of pooling and benefit of information, respectively.

Considering the work in Chapter 3 of this study, one of the main contributions is analysis and quantification of benefits of inventory and service pooling under optimal operating policies. Pooling the service implies an arriving customer could be redirected to another after-sales service provider, even if that after-sales service provider does not have stock. Axsäter [2003] makes analysis and quantification of benefits of inventory pooling, but under optimal heuristic policies. The studies that make analysis and quantification of benefits of inventory pooling under optimal operating policies are as follows. Grahovac and Chakravarty [2001] and Kranenburg and van Houtum [2009] consider static policies, which impedes the use of information in policy determination. Çömez et al. [2007] and Çömez [2009] consider a finite horizon and model the policy dynamic with respect to the remaining periods, i.e. information about the participants in the system is not used dynamically in policy determination. Zhao et al. [2008] makes only the quantification part, but under centralized setting. Decentralization of the system brings complexities in modelling, solution and analysis. In this study, analysis and quantification of benefits of inventory and service pooling under optimal operating policies under full information is made considering both dynamic and static policies in both centralized and decentralized setting, which has not been made in any of the studies in literature, to our knowledge. To make the analysis and quantification, optimal (dynamic), full, static and no pooling strategies are compared. By comparing the optimal policies of dynamic-pooling with no-pooling, the benefits of pooling is shown to be substantial, but with exceptions. As a managerial insight, an inappropriately designed pooling system, such as full-pooling, can be worse than no-pooling system. Under certain parameter values, full-pooling profit is observed to be less than half of the no-pooling profit.

The analysis of benefits of inventory and service pooling includes characterization of

Table 2.2: Comparison of Chapter 3 of the Study with Closely Related Literature

Study	**Benefit of Pooling**	**Comparison of Pooling Strategies**	**Information**	**Policy Considered**	**Comparison of Authority Types**
Axsäter [2003]	analyzed and quantified under heuristic operating policies	heuristic decision rule, simple decision rule and no-pooling are compared	full	static	centralized
Çömez et al. [2007]	analyzed and quantified under optimal operating policies	partial, complete and no pooling strategies are compared	full	dynamic (w.r.t. time)	centralized
Çömez [2009]	analyzed and quantified under optimal operating policies	optimal, complete and no pooling strategies are compared	full	dynamic (w.r.t. time)	decentralized
Grahovac and Chakravarty [2001]	analyzed and quantified under optimal operating policies	pooling and no-pooling strategies are compared under decentralized and centralized systems	approximate	static	centralized and decentralized
Kranenburg and van Houtum [2009]	analyzed and quantified under optimal operating policies	Partial pooling and special cases (no and full) are compared	full	static	centralized
Kukreja et al. [2001]	Benefit of policy determination when taking pooling into account is measured	Only pooling under emergency requirement is considered	full	static	decentralized
Zhao et al. [2006]	not analyzed or quantified	base-stock, base-stock with rationing, base-stock with rationing and transshipment in decentralized system, and base-stock with rationing and transshipment in centralized system	approximate	static	decentralized
Zhao et al. [2005]	not analyzed or quantified	full-sharing, fixed-sharing-level and inventory-rationing strategies compared	approximate	static	centralized and decentralized
Zhao et al. [2008]	quantified under optimal operating policies	optimal dynamic, static and no pooling strategies compared	full	dynamic and static	centralized
PhD Dissertation-Chapter 3	analyzed and quantified under optimal operating policies	optimal dynamic, full, static and no pooling strategies compared	full	dynamic and static	centralized and decentralized

optimal policies, which is another main contribution of this study. In previous works, similar analyses are made for centralized setting only, which is a simpler analysis compared to decentralized setting. A single dealer's problem is focused, where dealers are not necessarily identical to each other. It is observed that decentralized and centralized systems have opposite monotonic behavior for pooling threshold control variables. This interesting behavior identifies the effect of competition in spare parts networks. There are also non-monotonic instances for the pooling system. In those non-monotonic instances, policy might have different behaviors with respect to the status of the other dealers, which is not the case in centralized setting.

Main contribution of this study about the information sharing effects on the system is related with both the works in Chapters 3 and 4, but especially with Chapter 4. By comparing static-pooling and optimal-pooling strategies in Chapter 3, it is observed that if service centers already collaborate through inventory and service pooling, then additional benefit of sharing information and using that information in policy adjustment is almost negligible.

In Chapter 4, benefits of information sharing under several information availability levels are analyzed and quantified including information incompleteness on own status and on the other dealer's status. Parameter effects on those models are also analyzed in this study. Yan and Zhao [2008] consider information availabilities as different demand variance information and consider a static policy. Zhao and Qiu [2007] also consider a static policy and makes analysis and quantification of benefits of information under this assumption. To our knowledge, none of the previous works consider dynamic policy and make the analysis and quantification under dynamic policy.

Table 2.3: Comparison of Chapter 4 of the Study with Closely Related Literature

Study	**Benefit of Information**	**Comparison of Information Availabilities**	**Information**	**Policy Considered**	**Comparison of Authority Types**
Yan and Zhao [2008]	analyzed and quantified under different demand variance information availabilities	Full and approximate demand variance information availabilities compared	full, approximate	static	decentralized
Zhao and Qiu [2007]	analyzed and quantified under different information availabilities	Full and no on other's status information availabilities compared	full, no	static	centralized and decentralized
PhD Dissertation-Chapter 4	analyzed and quantified under different information availabilities	Full, partial and no on other's, and partial on own status information availabilities compared	full, partial, no	dynamic	decentralized

CHAPTER 3

ANALYSIS OF POOLING BENEFITS IN THE PRESENCE OF FULL INFORMATION AVAILABILITY

A decentralized spare parts system consisting of two dealers is considered, where both dealers are rational, i.e. they make the same decisions under the same conditions. The spare parts system is assumed to be stable and the conditions defining the system (economic, structural, etc.) are not changing over time. It is assumed that the dealers operate independently in the system, except that they interact through lateral transshipment of parts and demand. The aim is to find optimal operating policy of one of the dealers under the assumption of full information sharing and for a given operating policy of the other dealer. In this part of the thesis, under the setting described above, first the optimal operating policy of an independent dealer is characterized, and necessary and sufficient conditions for monotonicity are derived. Then, the behavior of the optimal policy under the decentralized system is contrasted with the optimal policy under the centralized system. The effects of system parameters on the profit is analyzed through a computational analysis. Finally, to quantify the benefit of resource pooling for an independent dealer in the presence of full information sharing, several pooling policies are proposed and the performances of the policies are compared with each other.

3.1 DESCRIPTION OF THE DECENTRALIZED SYSTEM

In the decentralized system, the dealer under consideration (Dealer 1, or D1) takes decisions about customer arrival and production. Whenever a customer arrives to

D1, the dealer has two options. In one option, he[1] may give service to the customer using his own resources, i.e. uses his inventory to meet the demand or backorders the demand in case of a stockout. Alternatively, he may ask the other dealer (Dealer 2, or D2) for his resources, i.e. asks for lateral transshipment of a part from the inventory of the other dealer. While taking those decisions, D2's inventory level is known by D1, which indicates full-information availability. Therefore, D1 only requests for lateral transshipment if D2's status allows the realization of the transshipment, otherwise D1 does not place a request. In case of a lateral transshipment request by D2 (upon a customer arrival to that location), D1 may accept the request and send a part or reject the request. Furthermore, at any point in time, D1 can decide whether to continue or stop production in order to control his inventory level.

The assumptions of the model are as follows:

- Customers arrive independently according to Poisson Process with constant rate to dealers.
- Dealers have dedicated production lines with exponential manufacturing times with constant rates, where each line acts as a single server.
- Transportation time is zero for transshipments between the dealers.

The system is assumed to be a single-echelon system. States of the system reflects the net inventory level of D1 (indicated by i) and the net inventory level of D2 (indicated by j), where $i \in I$, $I = \{i : i \in \mathbb{Z}\}$ and $j \in J$, $J = \{j : j = -\infty, \ldots, S_2 - 1, S_2\}$, where S_2 denotes the base-stock level of D2. i and j indicate the queue status for negative values (absolute value is the number of customers in the queue) and on-hand inventory level for positive values. Set of all states is represented by S, where $S = \{s : s = (i, j) \mid i \in I, j \in J\}$.

In the system, D1 makes decisions at the **decision epochs** that are triggered by i) customer arrivals, ii) lateral transshipment requests and iii) production completions. The available actions of D1 when a customer arrives are *Accept* and *ltr*, i.e. $a_1 \in \{Accept, ltr\}$, where the action *Accept* decreases the net inventory level of D1

[1] The parties of the system are referred to using pronouns 'he' and 'she' for dealers and manufacturer, respectively, with no sexist overtones intended.

and the action *ltr* decreases the net inventory level of D2 (if D2 accepts the request) by one. The action taken by D1 when D2 makes a lateral transshipment request is denoted by a_2, where $a_2 \in \{Accept, Reject\}$. If D1 rejects the lateral transshipment request, his net inventory level does not change but the net inventory level of D2 decrease by one, and if he decides to accept the request, his net inventory level decreases by one. Production times are assumed to be distributed with exponential distribution with rates μ_1 and μ_2 for D1 and D2, respectively. For production, D1 either does not continue production ($Reject$) or continues/initiates production ($Accept$). Action regarding the production is denoted by a_3, where $a_3 \in \{Accept, Reject\}$. The production decision is elaborated below after introducing the memoryless property of the system.

Figure 3.1 shows the transition rates of the states originating from state (i, j). In the figure, D1's plausible actions can be followed. Region I, Region II and Region III indicate the results of actions a_1, a_2 and a_3 in terms of transitions. All action sets of D1 are combined together and actions set A is constructed with elements $a = (a_1, a_2, a_3)$, where actions a_1, a_2, a_3 correspond to actions taken upon customer arrival, lateral transshipment request and production completion, respectively. Region IV denotes where D2 accepts production until his status reaches to his base-stock, S_2.

The exponential distribution assumptions make the system memoryless and the Markovian properties are satisfied, which makes a Markov Decision Process (MDP) model applicable. Since the actions are taken at every state transition and time spent in a state is an exponentially distributed random variable, the system can be modelled as a Continuous-Time Markov Decision Process (CTMDP). In the following, transition probabilities and a one-step reward function will be explained.

For production, memoryless property implies that an initiated replenishment order may be revised and adjourned upon an event that causes a decision epoch. This situation is **intermittent production**. In Figures 3.2 (a) and (b), there are two consecutive decision epochs. Initially, a customer arrival occurs and decision is given as $a_1 = Accept$ and the customer served, therefore the inventory level of D1 decreases by one. At this decision epoch, also the decision about production is given and by the production decision $a_3 = Accept$, D1 starts a new production or continues the current or previously adjourned production until the next decision epoch. The next

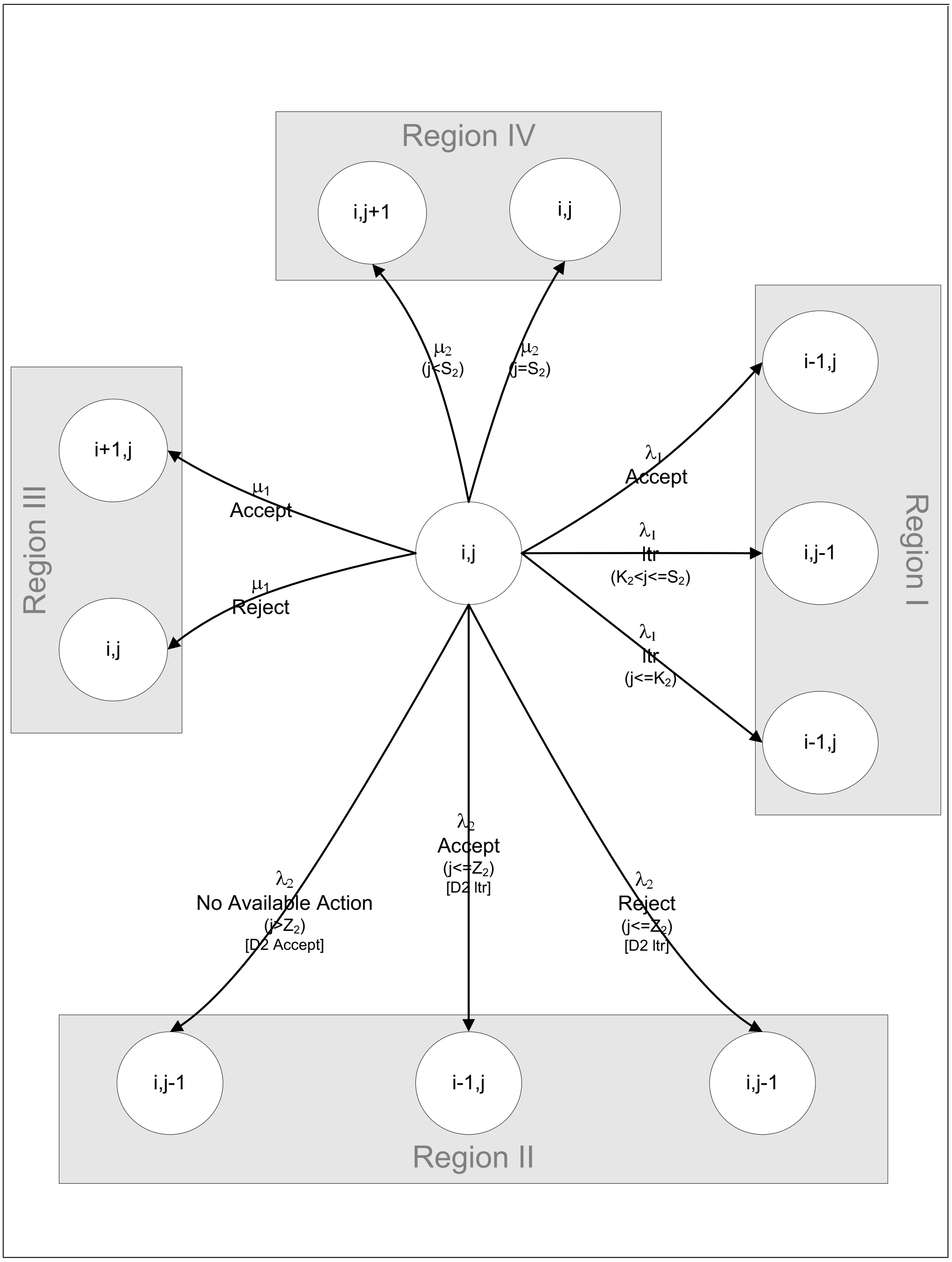


Figure 3.1: Transition Rate Diagram of Decentralized Model

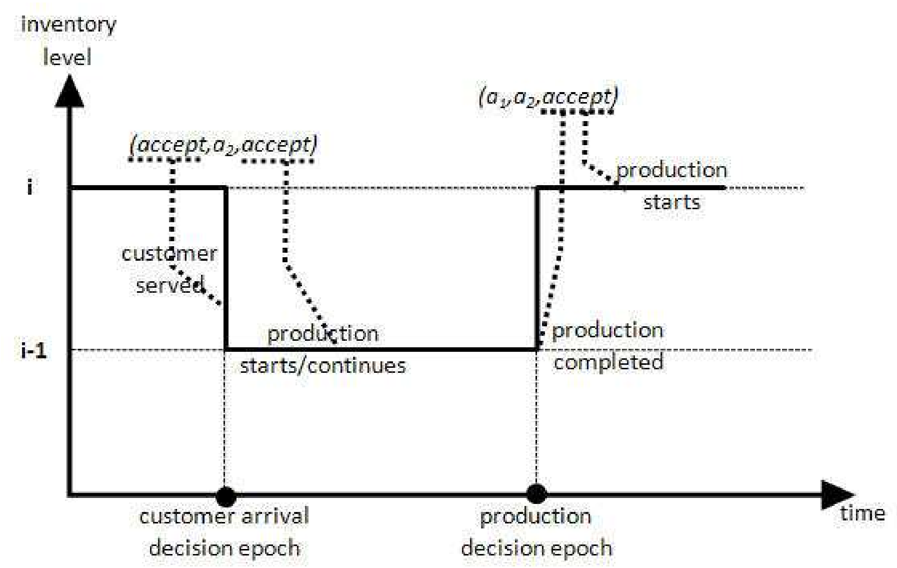
decision epoch is a production decision epoch and D1 makes a different decision about production. In Figure 3.2(a), by the production decision $a_3 = Accept$, D1 accepts the completed product and starts a new production which continues until the next decision epoch. In Figure 3.2(b), D1 adjourns the continuing production by the production decision $a_3 = Reject$.

Production is equivalent to transshipment from an external supplier under the following conditions. Consider a system where there is a manufacturer as the external supplier, which does not hold stock, produces with infinite capacity and exponential manufacturing times with constant rates, the transshipment time from the manufacturer to dealers is negligible, and the dealer has the option of adjourning the production for himself at the manufacturer. This system is equivalent to the system described previously, where dealers produce with their dedicated lines with exponential manufacturing times with constant rates. Another equivalent system is with a distribution center as the external supplier. Distribution center holds infinite stock or has the option of immediate replenishment from a manufacturer, the transshipment time from the distribution center to dealers is an exponentially distributed random variable with constant rates, at most one item can be transshipped to a dealer at a time (transshipment capacity is one), and the dealers has the option of cancelling a transshipment before the item is received.

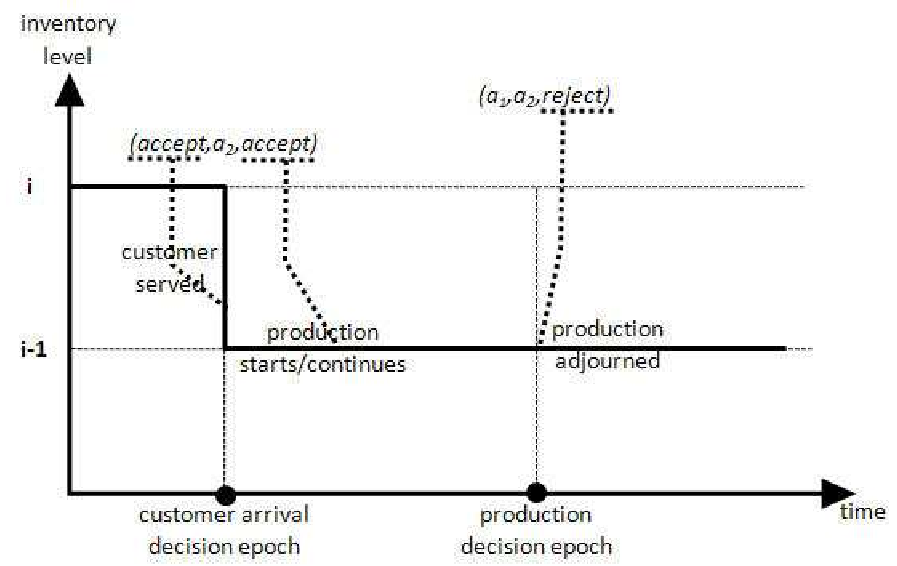
D2's operation strategy is assumed to be a threshold policy characterized by three levels as (S_2, K_2, Z_2), where S_2, K_2, and Z_2 are base-stock, rationing and transshipment levels, respectively. Table 3.1 presents the operating policy of D2.

One may argue that, it is unlikely in practice that D2 has a given policy but D1 optimizes his policy. Such cases may occur while transition of the pooling system (say from centralized setting to decentralized setting) or in hybrid systems. For example, consider the Akça service centers and the other VOLVO service centers. In this hybrid system, the centralized Akça service centers may act like a large dealer with a given policy and one of the other decentralized VOLVO service centers may optimize his policy.

To further explain the model, consider the state in Figure 3.1 and the chosen action (a_1, a_2, a_3) which will affect the system in the following event. For example, assume



(a) Production Continued



(b) Production Adjourned

Figure 3.2: Intermittent Production and Memoryless Property

Table 3.1: Policy of D2 for Decentralized Model

	Inventory/Queue Status of D2			
EVENTS	$j \geq S_2$	$S_2 > j > K_2$	$K_2 \geq j > Z_2$	$Z_2 \geq j$
part arrival	reject	accept		
ltr* from D1	accept		reject	
customer arrival	accept			ltr*

(*ltr: lateral transshipment request)

that the next event will be a production completion for D1. In that case, if $a_3 = Accept$, then next state will be $(i+1, j)$, and if $a_3 = Reject$, then next state will be (i, j). This can be seen in Region III of Figure 3.1. Now, assume that the next event will be a customer arrival to D2. In that case, only action a_2 will affect the next state. In case D2 does not ask for lateral transshipment, i.e. $j > Z_2$, no action is allowable for D1. If $j \leq Z_2$, D2 asks for lateral transshipment for a customer arrival, and next state will be $(i, j-1)$ if $a_2 = Reject$, and $(i-1, j)$ if $a_2 = Accept$. This can be seen in Region II of Figure 3.1. Lastly, assume that the next event will be a customer arrival to D1. In that case, action a_1 will affect the next state. If $a_1 = Accept$ is chosen by D1 for a customer arrival, then next state will be $(i-1, j)$. In case of $a_1 = ltr$ is chosen, the next state depends on j, where next state is $(i, j-1)$ if $K_2 < j \leq S_2$ as D2 sends an item to D1 and it is $(i-1, j)$ if $j \leq K_2$ as D2 does not make a lateral transshipment. This can be seen in Region I of Figure 3.1. This implies that D1's action space is defined by D2's operation strategy and D2's status. For example, for $K_2 \geq j$, D2 rejects lateral transshipment requests of D1, and action $a_1 = ltr$ is not available to D1. Table 3.2 presents the available actions of D1.

Let π denote a policy for D1, such that (a_1, a_2, a_3) defines actions for a given state under policy π. Let $v^\pi(i, j)$ be the expected discounted profit of D1 under infinite horizon for D1 starting with initial state (i, j) under policy π. Let $(i, j)_t$ denote the

Table 3.2: Available Actions of D1 for Decentralized Model

	Inventory/Queue Status of D2		
EVENTS	$S_2 \geq j > K_2$	$K_2 \geq j > Z_2$	$Z_2 \geq j$
part arrival	accept, reject		
ltr from D2	no available action		accept, reject
customer arrival	accept, ltr	accept	

state of the system at time t. $v^{\pi}(i, j)$ can be defined as in Equation 3.1 below:

$$v^{\pi}(i, j) = E^{\pi}_{(i,j)}\Big[\int_0^{\infty} e^{-\alpha t}\Big[\{-H[(i, j)_t] - L[(i, j)_t]\}dt + RdN_1(t) + rdN_2(t) + (R-r)dN_3(t)\Big]\Big] \tag{3.1}$$

where $\alpha > 0$ is the discount rate, $N_1(t)$ denotes the number of customer arrivals to D1 by time t, $N_2(t)$ denotes the number of accepted lateral transshipment requests of D2 by D1 and $N_3(t)$ denotes the number of customers of D1 served by lateral transshipments from D2 by time t, $H(i, j)$ and $L(i, j)$ are holding and customer waiting costs which are assumed as linear and defined as follows. For the customers waiting in the queue, waiting cost incurs and c_l denotes waiting cost per customer per unit time. For the unassigned items in inventory at dealer's depot, holding cost incurs and c_h denotes holding cost per item per unit time. For state (i, j), inventory holding cost rate can be defined as $H(i, j) = c_h i^+$ and waiting cost rate can be defined as $L(i, j) = c_l i^-$ for D1, where $i^+ = max\{0, i\}$ and $i^- = max\{0, -i\}$.

The net revenue per customer by using own resources is denoted by R for the single customer type. In case of accepting a lateral transshipment request, the requesting dealer shares part of his own-customer revenue R, i.e. gives r ($r < R$) to the accepting dealer. As a result, the dealer facing with own-customer demand has a net revenue of $R - r$ in case of serving the customer by lateral transshipment. Rejecting the lateral transshipment request has no associated costs, but just the implicit opportunity cost of $R - r$.

Following Lippman [1975], uniformization is used to convert the Continuous-Time Markov Chain into discrete time. For this purpose, the uniform rate β is defined as $\beta = \lambda_1 + \lambda_2 + \mu_1 + \mu_2$. Uniformization leads a uniform transition rate and the infinite horizon continuous time decision process is converted into discrete time decision process. Letting p denote the discrete time periods, expected discounted profit of D1 is rewritten in Equation 3.12 below as:

$$\begin{aligned} v^{\pi}(i,j) &= E^{\pi}_{(i,j)}\Big[\sum_{p=0}^{\infty}\Big(\frac{\beta}{\alpha+\beta}\Big)^{p}\frac{\{-H[(i,j)(p)] - L[(i,j)(p)]\}}{\alpha+\beta} + \sum_{p=1}^{\infty}\Big(\frac{\beta}{\alpha+\beta}\Big)^{p}\Big[R\{N_1(p) \\ &\quad -N_1(p-1)\} + r\{N_2(p) - N_2(p-1)\} + (R-r)\{N_3(p) - N_3(p-1)\}\Big]\Big] \end{aligned} \tag{3.2}$$

Transition probabilities in the discrete time decision process can be defined as follows:

$$p((i',j')|(i,j),(a_1,a_2,a_3)) = \begin{cases} \{\mu_1 I_{[a_3=Accept]}\}/\beta, & \text{if } (i',j') = (i+1,j) \\ \{\mu_2 I_{[j<S_2]}\}/\beta, & \text{if } (i',j') = (i,j+1) \\ \{\lambda_2 I_{[j\le Z_2]} I_{[a_2=Accept]} + \lambda_1 I_{[j\le K_2]} \\ \quad + \lambda_1 I_{[j>K_2]} I_{[a_1=Accept]}\}/\beta, & \text{if } (i',j') = (i-1,j) \\ \{\lambda_2 I_{[j\le Z_2]} I_{[a_2=Reject]} + \lambda_1 I_{[K_2<j\le S_2]} I_{[a_1=ltr]} \\ \quad + \lambda_2 I_{[Z_2<j]}\}/\beta, & \text{if } (i',j') = (i,j-1) \\ 1 - \sum_{(i'',j'')\neq(i,j)} p((i'',j'')|(i,j),(a_1,a_2,a_3)), \\ \quad \text{if } (i',j') = (i,j) \end{cases} \tag{3.3}$$

where $I_{[z]}$ is an indicator function and equals to zero when z is false and equals one when z is true. Objective is to choose a policy π^* which maximizes expected profit of D1. The optimality equation $v^{\pi^*}(i,j)$ can be written as follows:

$$
\begin{aligned}
v^{\pi^*}(i,j) \;=\;& max_{(a_1,a_2,a_3))}\Big\{\frac{[-H(i,j)-L(i,j)]}{\alpha+\beta}+\frac{\beta}{\alpha+\beta}\Big[\frac{\lambda_1}{\beta}(I_{[j\leq K_2]}+I_{[j>K_2]}I_{[a_1=Accept]})R \\
&+\frac{\lambda_1}{\beta}I_{[K_2<j\leq S_2]}I_{[a_1=ltr]}(R-r)+\frac{\lambda_2}{\beta}I_{[j\leq Z_2]}I_{[a_2=Accept]}r \\
&+\sum_{\forall(i',j')}p((i',j')|(i,j),(a_1,a_2,a_3))v^{\pi^*}(i',j')\Big]\Big\} \qquad (3.4)
\end{aligned}
$$

Alternatively, one could express the optimality equation as follows:

$$
\begin{aligned}
v^*(i,j) \;=\;& \frac{[-c_h i^+ - c_l i^-]}{\alpha+\beta}+\Big(\frac{\beta}{\alpha+\beta}\Big)\Big[\frac{\lambda_1}{\beta}\text{Ø}_1 v^*(i,j)+\frac{\lambda_2}{\beta}\text{Ø}_2 v^*(i,j) \\
&+\frac{\mu_1}{\beta}\text{Ø}_3 v^*(i,j)+\frac{\mu_2}{\beta}\text{Ø}_4 v^*(i,j)\Big] \qquad (3.5)
\end{aligned}
$$

where $\text{Ø}_k v^*(i,j)$ for $k = 1,...,4$ are defined for any real-valued function $v(i,j)$ as follows. For any real-valued function $v(i,j)$, let $\text{Ø}_1 v(i,j)$ be an operator (for a customer arrival to D1) expressed as follows:

$$
\text{Ø}_1 v(i,j) = \begin{cases} max\{v(i-1,j)+R, v(i,j-1)+(R-r)\} & \text{if } K_2 < j \leq S_2 \\ v(i-1,j)+R & \text{if } j \leq K_2 \end{cases} \qquad (3.6)
$$

For any real-valued function $v(i,j)$, let $\text{Ø}_2 v(i,j)$ be an operator (for a customer arrival to D2) expressed as follows:

$$
\text{Ø}_2 v(i,j) = \begin{cases} v(i,j-1) & \text{if } j > Z_2 \\ max\{v(i-1,j)+r, v(i,j-1)\} & \text{if } j \leq Z_2 \end{cases} \qquad (3.7)
$$

For any real-valued function $v(i,j)$, let $\text{Ø}_3 v(i,j)$ be an operator (for a production completion for D1) expressed as follows:

$$
\text{Ø}_3 v(i,j) = max\{v(i+1,j), v(i,j)\} \qquad (3.8)
$$

For any real-valued function $v(i,j)$, let $\text{Ø}_4 v(i,j)$ be an operator (for a production completion for D2) expressed as follows:

$$\O_4 v(i,j) = \begin{cases} v(i,j) & \text{if } j = S_2 \\ v(i,j+1) & \text{if } j < S_2 \end{cases} \tag{3.9}$$

For optimal policy π^*, D1's optimal profit function $v^* \equiv v^{\pi^*}$ satisfies the optimality equation (Equation 3.5). By redefining the timescale and assuming $\alpha+\beta = 1$, without loss of generality, optimality equation becomes as follows:

$$v^*(i,j) = -c_h i^+ - c_l i^- + \lambda_1 \O_1 v^*(i,j) + \lambda_2 \O_2 v^*(i,j) + \mu_1 \O_3 v^*(i,j) + \mu_2 \O_4 v^*(i,j) \tag{3.10}$$

Equation 3.10 is the optimality equation that will be used for the decentralized model in the following parts of the study.

3.2 CHARACTERIZATION OF THE OPTIMAL POLICY UNDER THE DECENTRALIZED SYSTEM

In this section, first **optimal policy** is characterized for the problem under consideration and necessary and sufficient conditions are defined for the optimal policy. Then, another sets of conditions for $v(i, j)$, which are sufficient conditions, are explained. Finally, the structure of the optimal policy of the problem is briefly discussed.

Proposition 3.2.1 *Let* ***V*** *be the set of real functions defined on* $\mathbb{Z}^2$*, such that if* $v \in$ ***V****, then*

Q1*:* $v(i,j) - v(i-1,j) \leq 0 \Rightarrow v(i+1,j) - v(i,j) \leq 0, \quad \forall (i,j) \in \mathbb{Z}^2$

Q2*:* $v(i+1,j) - v(i,j) \leq 0 \Rightarrow v(i+1,j+1) - v(i,j+1) \leq 0, \quad \forall (i,j) \in \mathbb{Z}^2$

Q3*:* $v(i,j-1) - v(i-1,j) \leq r \Rightarrow v(i+1,j-1) - v(i,j) \leq r, \quad \forall (i,j) \in \mathbb{Z}^2$

Q4*:* $v(i,j-1) - v(i-1,j) \leq r \Rightarrow v(i,j) - v(i-1,j+1) \leq r, \quad \forall (i,j) \in \mathbb{Z}^2$

If $\exists\, v \in$ ***V*** *that satisfies Equation 3.10, there exists an* ***optimal operating policy*** *for D1 that can be characterized with* $(S_1(j), K_1(j), Z_1(j))$ *as follows.*

Let

$S_1(j) = \min\{i | v(i+1,j) - v(i,j) \leq 0\}$,

$K_1(j) = \max\{i | v(i,j-1) - v(i-1,j) > r\}$ *(defined for* $j \leq Z_2$*),*

$Z_1(j) = \max\{i | v(i,j-1) - v(i-1,j) > r\}$ *(defined for* $j > K_2$*),*

where $S_1(j)$ denotes the base-stock level, $K_1(j)$ denotes the rationing level and $Z_1(j)$ denotes the transshipment level. The ***replenishment policy*** *is a state-dependent base-stock policy with base-stock levels $S_1(j)$, such that for $i < S_1(j)$ produce to replenish the stock, for $i \geq S_1(j)$ stop production and not replenish the stock. The* ***rationing policy*** *is a state-dependent policy with rationing levels $K_1(j)$, such that for $i > K_1(j)$ accept the lateral transshipment requests of D2, for $i \leq K_1(j)$ reject the requests. The* ***transshipment policy*** *is a state-dependent policy with transshipment levels $Z_1(j)$, such that for $i \leq Z_1(j)$ place a lateral transshipment request to D2 (if allowable), for $i > Z_1(j)$ meet the customer demand without transshipment by itself.*
The optimal operating policy has the following property:

$S_1(j),\ K_1(j)$ *and* $Z_1(j)$ *are non-increasing in* j.

Furthermore, if there exists an optimal operating policy characterized as above, then ***Q1-Q4*** *are the necessary and sufficient conditions for the optimal operating policy and the corresponding value function v under the infinite horizon discounted gain criteria satisfies $v \in \mathbf{V}$.*

Proof. See Appendix A.

The policy defined above is monotone in (i,j) in the following sense: (i) The policy is a control-limit type policy in i with three control limits $S_1(j),\ K_1(j)$, and $Z_1(j)$, (ii) Each control limit is a non-increasing function of j.

Operators Ø_1 to Ø_4 (call Ø_n) are defined in Section 3.1 on the set of real-valued functions, v. Consider the following conditions.

C1: $v(i, j) - v(i-1, j) \geq v(i+1, j) - v(i, j), \forall i, j$
C2: $v(i+1, j) - v(i, j) \geq v(i+1, j+1) - v(i, j+1), \forall i, j$
C3: $v(i, j-1) - v(i-1, j) \geq v(i+1, j-1) - v(i, j), \forall i, j$
C4: $v(i, j-1) - v(i-1, j) \geq v(i, j) - v(i-1, j+1), \forall i, j$

Conditions **Q1-Q4** given in Proposition 3.2.1 are necessary and sufficient conditions. The set of conditions **C1-C4** defined above are sufficient conditions for the optimal operating policy defined as in Proposition 3.2.1 to exist, which implies monotonicity of control variables. Table 3.3 summarizes the relations between conditions and control variables.[2] If the sufficient conditions are shown to be satisfied, existence and monotonicity of control variables can be proven. For example, existence of S_1 indicates that there is an $S_1(j)$ for each j such that it is optimal to produce when $i < S_1(j)$ and stop production when $i \geq S_1(j)$. In other words, there is an $S_1(j)$ for each j which maximizes $v(i, j)$ over i. When condition **C1** (or **Q1**) is satisfied, existence of S_1 is implied. Monotonicity of S_1 indicates that, with increasing j, optimal base-stock levels do not increase. In other words, $S_1(j') \leq S_1(j)$ for $j' > j$. When condition **C2** (or **Q2**) is satisfied, monotonicity of S_1 is implied.

Table 3.3: Relation between Control Variables and Conditions for Decentralized System

Condition	Related Control Variable(s)	Indicates
C1, Q1	S_1	existence
C2, Q2	S_1	monotone ↓
C3, Q3	K_1	existence
	Z_1	existence
C4, Q4	K_1	monotone ↓
	Z_1	monotone ↓

Please note that, within the necessary and sufficient conditions, condition **Q1** is quasi-concavity condition and conditions **Q2-Q4** are quasi-sub/supermodularity conditions, and within the sufficient conditions, condition **C1** is concavity condition and conditions **C2-C4** are sub/supermodularity conditions. In order to understand the relation between **Q1-Q4** and **C1-C4**, consider **Q1** and **C1**. Condition **C1** is the sufficient existence condition of $S_1(j)$ which indicates concavity of D1's total discounted expected profit function $v(i, j)$ on i. In Figure 3.3-a, this sufficient condition is sketched. The point that the function $v(i, j)$ maximum is $S_1(j)$, and for the i values less than $S_1(j)$, $v(i, j)$ is smaller than $v(S_1(j), j)$, and vice-versa. Based on the difference of $v(i, j)$ on i, **C1** is obtained. The quasi-concavity condition **Q1** is sketched in Figure 3.3-b, which can be stated as $v(i, j) \geq min\{v(i-1, j), v(i+1, j)\}, \forall i, j$. Please note that con-

[2] For detailed explanation of the relations between conditions and control variables, please see Appendix B.

dition **Q1** holds for $v(i, j)$ seen in Figure 3.3-a, but **C1** does not hold for $v(i, j)$ seen in Figure 3.3-b. Whatever the form of the function seen in Figure 3.3 is, existence and uniqueness of $S_1(j)$ are clear.

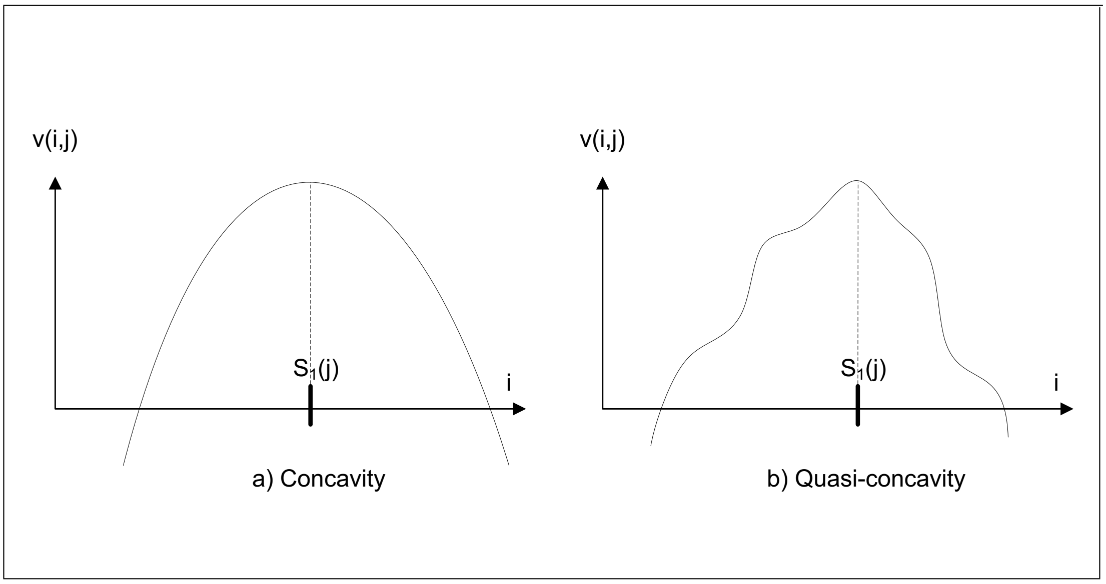


Figure 3.3: Concavity and Quasi-concavity of $S_1(j)$

Monotonicity of $S_1(j)$ simply means that $S_1(j)$ is non-increasing in j. Both forms of $v(i, j)$ in Figure 3.4 satisfy this requirement. In Figure 3.4-a, sufficient submodularity condition **C2** is sketched. The points that $v(i, j)$ and $v(i, j+1)$ maximum are $S_1(j)$ and $S_1(j+1)$, respectively. $S_1(j+1)$ is always smaller than or equal to $S_1(j)$, which is satisfied by **C2**. The quasi-submodularity condition **Q2** is sketched in Figure 3.4-b. Please note that the quasi-concavity condition **Q2** holds for $v(i, j)$ and $v(i, j+1)$ seen in Figure 3.4-a, but **C2** does not hold for $v(i, j)$ and $v(i, j+1)$ seen in Figure 3.4-b. In both forms of the functions seen in Figure 3.4, monotonicity of $S_1(j)$ is clear.

Numerical analysis shows that, under many instances $v(i, j)$ in (3.10) satisfies the conditions in Proposition 3.2.1 and the optimal operating policy of D1 is monotone. However; there are parameter settings under which the conditions **Q1-Q4** are not satisfied and the structure is non-monotone. This observation is formalized as follows.

Observation 3.2.2 *D1 has an optimal control policy that can be characterized as in Proposition 3.2.1 with the control variables* $(S_1(j)^*, K_1(j)^*, Z_1(j)^*)$ *for each j. In many instances* $S_1(j)^*$, $K_1(j)^*$, *and* $Z_1(j)^*$ *are non-increasing in j, however; there are cases where* $S_1(j)^*$, $K_1(j)^*$, *and* $Z_1(j)^*$ *are non-monotone in j.*

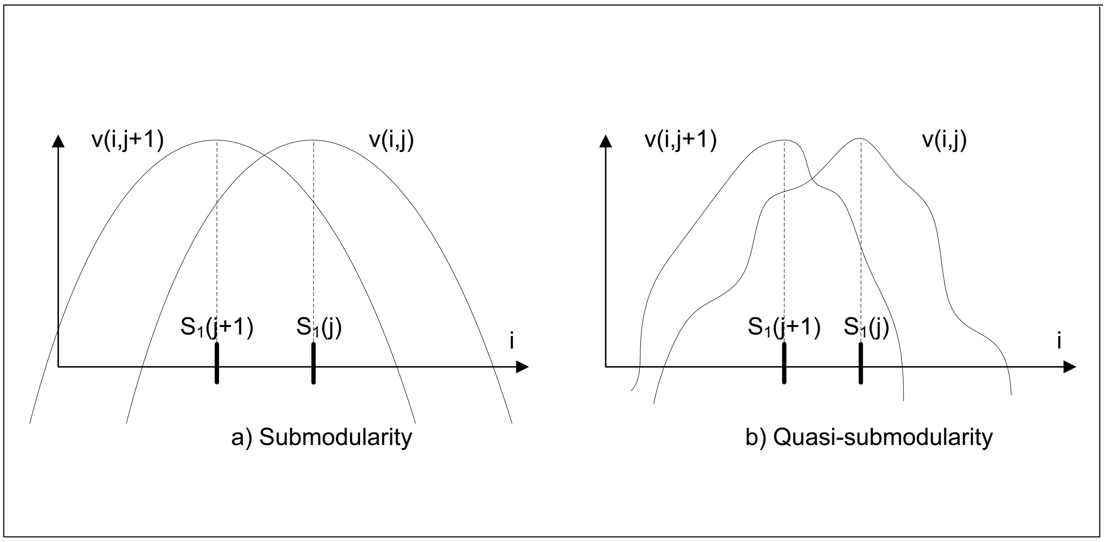


Figure 3.4: Submodularity and Quasi-submodularity of $S_1(j)$

In terms of conditions, non-monotonicity can be stated as follows. The proof of monotonicity fails for decentralized system using the conditions **C1-C4**. Numerical results also support the violation of conditions. Numerical analysis gives the values of $v(i, j)$ and these are evaluated in terms of the above conditions. It is observed that those conditions are not satisfied by numerical results. Figure 3.5 shows the 3-dimensional graph of $v(i, j)$ with respect to i and j.

Numerical evaluations of conditions are made. Figure 3.6 shows the graphs of conditions **C1** to **C4** for a sample solution, where parameter values as follows: $S_2 = 3, K_2 = 0, Z_2 = 0, r = 10, R = 10, c_h = 2, c_l = 2, \lambda_1 = 0.7, \lambda_2 = 0.7, \mu_1 = 1, \mu_2 = 1$. In Figures 3.5 and 3.6, all conditions as rephrased to be nonnegative. For example, **C1** is rephrased as $2v(i, j) - v(i+1, j) - v(i-1, j) \geq 0$. It can be seen from the graphs that all conditions have negative regions, which indicate violation of them. For example, at states $(3, 0)$, $(-5, 2)$, $(-3, -30)$, **C4** is negative.

Furthermore, numerical results show that conditions **Q1-Q4** are not satisfied for some of the instances either.

When optimal policies for the same parameters except the payment level are compared, extreme payment levels have a special property. If payment level were a decision variable, then D1 would prefer the extreme values on payment level. The following result formalize this observation.

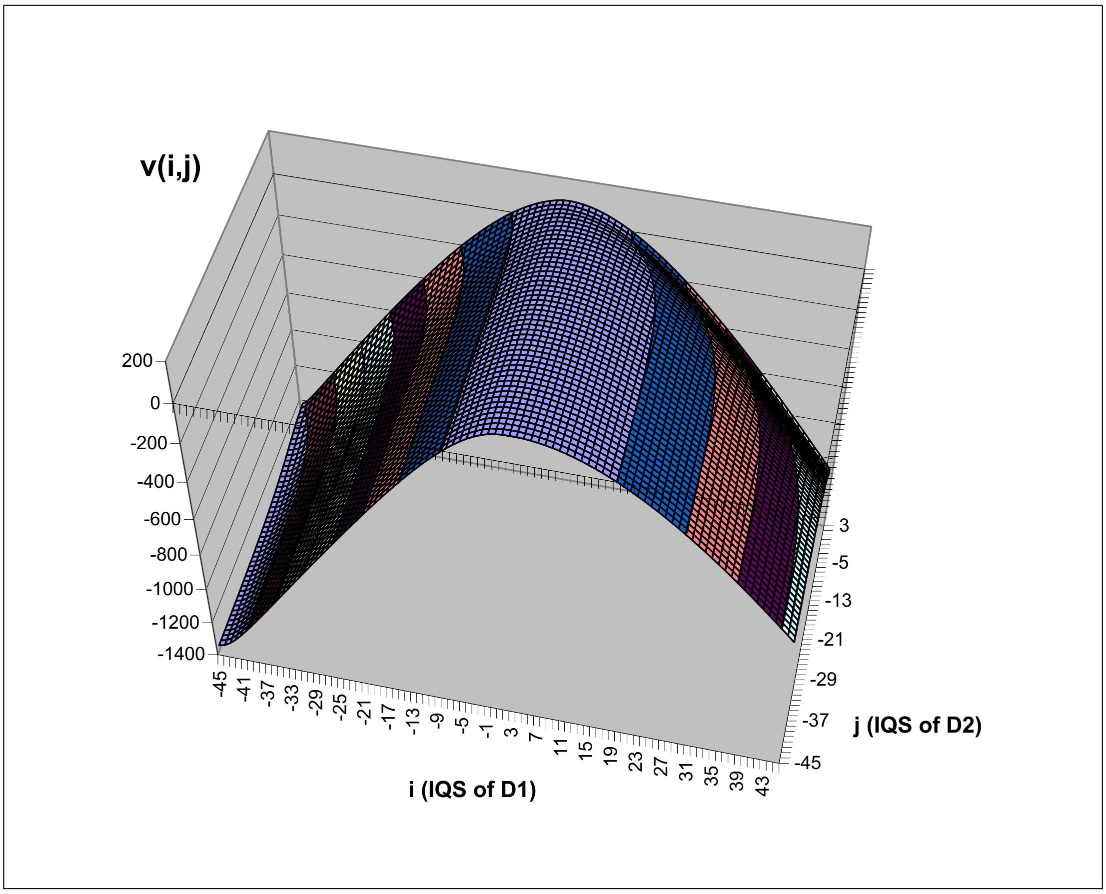


Figure 3.5: $v(i, j)$ with respect to States for a Sample Instance

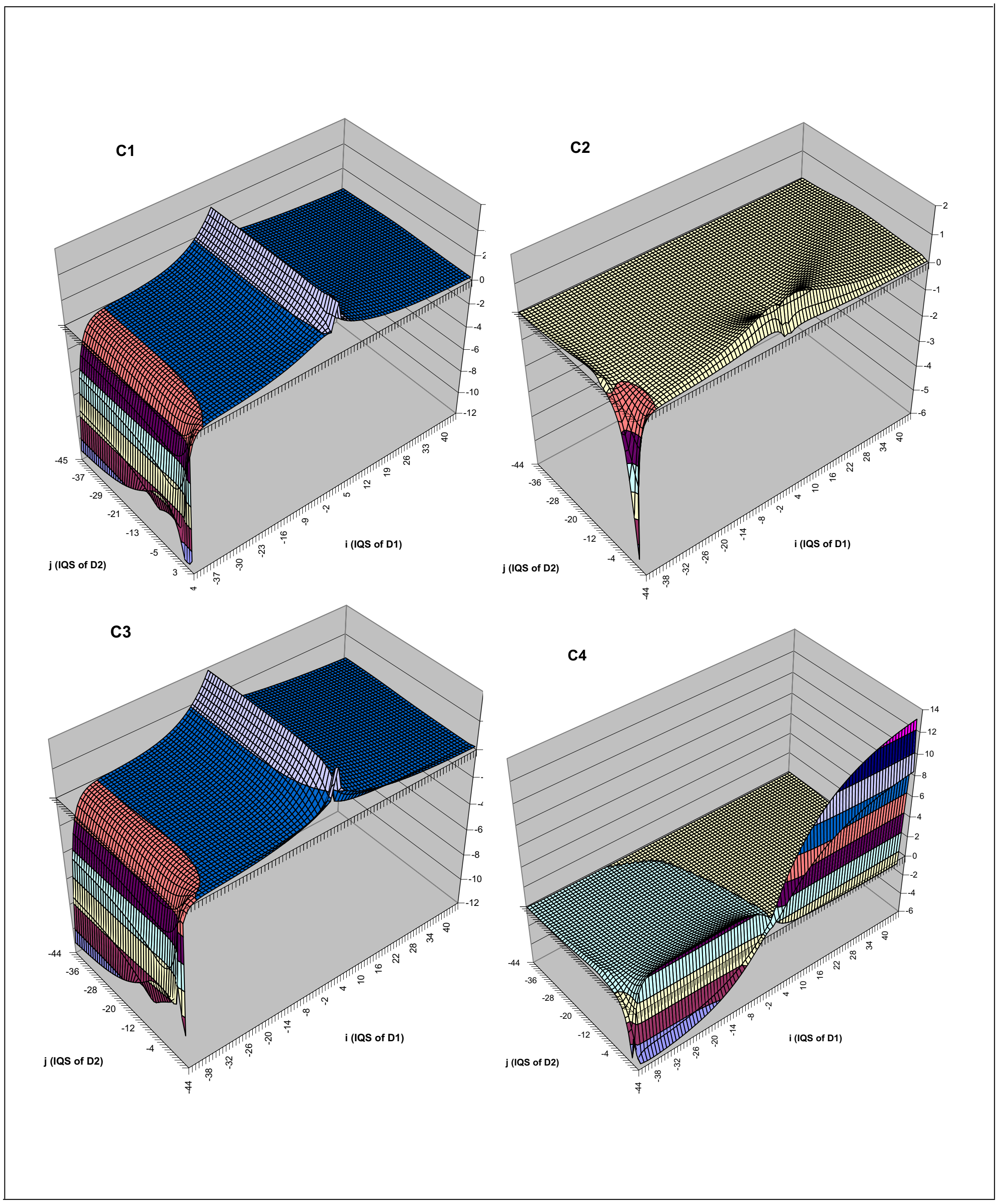


Figure 3.6: Sufficient Conditions with respect to States for a Sample Instance

Proposition 3.2.3 *Under any parameter setting, D1 prefers either the lowest payment level ($r = 0$) or the highest payment level ($r = 10$), but not the intermediate payment levels.*

Proof. See Appendix C.

3.3 COMPARISON WITH THE CENTRALIZED SYSTEM

To better understand the dynamics of the optimal operating policy in a decentralized system, we also study the case where the service centers operate under a central authority (or the manufacturer). In centralized system setting, similar to the decentralized system, a two-dealer spare part service network is considered (Dealer 1 or D1, and Dealer 2 or D2). The difference of this setting to the decentralized system is the existence of a central authority. The central authority controls the dealers and makes decisions in behalf of those dealers.

Whenever a customer arrives to one of the dealers, the central authority has two options. In one option, he may give service to the customer using the same dealer's resources, i.e. uses his inventory for service or backorders the demand in case of stockout. In the other option, he may use the other dealer's resources for customer service, i.e. makes a lateral transshipment of a part from the inventory of the other dealer. The central authority decides whether to continue or stop production at the dealers in order to control their inventory levels. While taking those decisions, the dealers' inventory levels are known by the central authority, which indicates full-information availability.

In this centralized system, system-wide objectives or performance measures are considered, like maximizing expected profit or some measure of customer service, or like minimizing operation costs or some measure of customer dissatisfaction. Clearly these objectives or performance measures are highly dependent on central authority's actions. When a customer service is given, independent of which dealer's resource is used, the unit revenue is earned in the system. Therefore, the total revenue depends on the unit revenue and the arrival rates of customers to the dealers. Similarly, holding and waiting costs are incurred jointly. The system is defined in the next subsection in detail.

States of the system consists of the net inventory levels of D1 and D2 and indicated by (i, j), as in the decentralized system. Figure 3.7 shows the transition rates of the states originating from state (i, j).

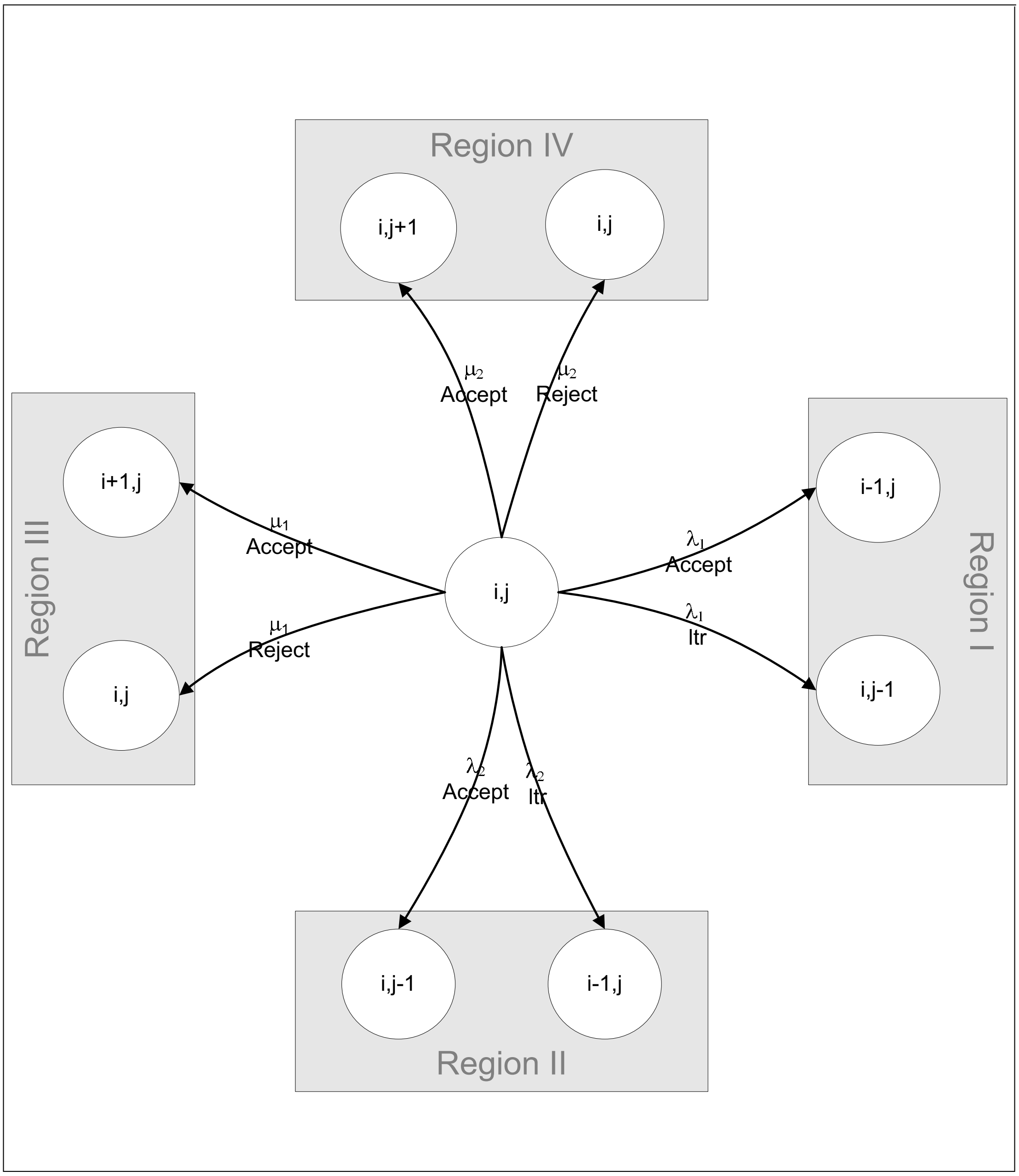


Figure 3.7: Transition Rate Diagram of Centralized Model

In the system, the central authority makes decisions that are triggered by i) customer arrivals and ii) production completions. The available actions, a_1, of the central authority when a customer arrives to D1 are *Accept* and *ltr*. The available actions a_2

when a customer arrives to D2 are similarly *Accept* and *ltr*. Status of the dealer giving service to his customer or sending part to the other dealer decreases by one. Each dealer has dedicated production lines associated with him, with exponentially distributed production times with rates μ_1 and μ_2 for D1 and D2, respectively. The central authority has available actions a_3 and a_4 for production decision at D1 and D2 as *Reject* which does not affect status and *Accept* which increases status by one of the corresponding dealer, respectively. In Figure 3.7, D1's plausible actions can be followed. Region I, Region II, Region III and Region IV indicate the results of a_1, a_2, a_3 and a_4 actions in terms of transitions. All action sets of D1 are combined together and actions set A is constructed with elements $a = (a_1, a_2, a_3, a_4)$, where actions a_3, a_4 correspond to actions taken upon production completion at D1 and D2, and a_1, a_2 correspond to actions taken upon customer arrival to D1 and D2, respectively.

The exponential distribution assumptions make the system memoryless and the Markovian properties are satisfied, which makes a MDP model applicable. In addition, actions are chosen at every state transition by D1 and time spent in a state is an exponentially distributed random variable, therefore the system can be modelled as a CTMDP. Transition probabilities and one-step reward function will be explained in the following parts of this section.

Let π denote a policy for central authority, such that (a_1, a_2, a_3, a_4) defines actions for a given state under policy π. Let $v^\pi(i, j)$ be the expected discounted profit of the system (i.e. sum of D1's and D2's profits) under infinite horizon for D1 starting with initial state (i, j) under policy π. Let $(i, j)_t$ denote the state of the system at time t. $v^\pi(i, j)$ can be defined in Equation 3.11 below:

$$v^\pi(i, j) = E^\pi_{(i,j)}\Big[\int_0^\infty e^{-\alpha t}\Big[\{-H[(i, j)_t] - L[(i, j)_t]\}dt + R\{dN_1(t) + dN_2(t)\}\Big]\Big] \quad (3.11)$$

where $\alpha > 0$ is the discount rate, $N_1(t)$ and $N_2(t)$ denote the number customer arrivals to D1 and D2 by time t respectively, $H(i, j)$ and $L(i, j)$ are holding and waiting costs which are assumed as linear and defined as follows. For the customers waiting in the queue, waiting cost incurs and c_l denotes waiting cost per customer per unit time. For the unassigned items in inventory at dealers' depots, holding cost incurs and c_h

denotes holding cost per item per unit time. For state (i, j), inventory holding cost rate can be defined as $H(i, j) = c_h(i^+ + j^+)$ and waiting cost rate can be defined as $L(i, j) = c_l(i^- + j^-)$ for D1, where $i^+ = max\{0, i\}$, $i^- = max\{0, -i\}$, $j^+ = max\{0, j\}$, $j^- = max\{0, -j\}$. The net revenue generated by per customer to the system is the same for both dealers as there is a single customer type, which is denoted by R. In case of lateral transshipment, the net revenue to the system does not change.

Following Lippman [1975], uniformization is used to convert the Continuous-Time Markov Chain into discrete time. For this purpose, the uniform rate β is defined as $\beta = \lambda_1+\lambda_2+\mu_1+\mu_2$. Uniformization leads a uniform transition rate and the infinite horizon continuous time decision process is converted into discrete time decision process. Letting p denote the discrete time periods and $(i, j)(p)$ denote the state of the system at p, expected discounted profit of the system is rewritten in Equation 3.12 below as:

$$\begin{aligned} v^\pi(i, j) \;=\; & E^\pi_{(i,j)}\Big[\sum_{p=0}^{\infty}\Big(\frac{\beta}{\alpha+\beta}\Big)^p \frac{\{-H[(i, j)(p)] - L[(i, j)(p)]\}}{\alpha+\beta} + \sum_{p=1}^{\infty}\Big(\frac{\beta}{\alpha+\beta}\Big)^p \Big[R\Big(\{N_1(p) \\ & -N_1(p-1)\} + \{N_2(p) - N_2(p-1)\}\Big)\Big]\Big] \end{aligned} \tag{3.12}$$

Transition probabilities in the discrete time decision process can be defined as follows:

$$p((i', j')|(i, j), (a_1, a_2, a_3, a_4)) = \begin{cases} \{\mu_1 I_{[a_3=Accept]}\}/\beta, & \text{if } (i', j') = (i+1, j) \\ \{\mu_2 I_{[a_4=Accept]}\}/\beta, & \text{if } (i', j') = (i, j+1) \\ \{\lambda_1 I_{[a_1=Accept]} + \lambda_2 I_{[a_2=ltr]}\}/\beta, & \text{if } (i', j') = (i-1, j) \\ \{\lambda_1 I_{[a_1=ltr]} + \lambda_2 I_{[a_2=Accept]}\}/\beta, & \text{if } (i', j') = (i, j-1) \\ 1 - \sum_{(i'', j'')\neq(i,j)} p((i'', j'')|(i, j), (a_1, a_2, a_3, a_4)), & \\ \quad \text{if } (i', j') = (i, j) & \end{cases} \tag{3.13}$$

where $I_{[z]}$ is an indicator function and equals to zero when z is false and equals one when z is true. Objective is to choose a policy π^* which maximizes expected profit of the system. The optimality equation $v^{\pi^*}(i, j)$ can be written as follows:

$$
\begin{aligned}
v^{\pi^*}(i,j) \quad = \quad & max_{(a_1,a_2,a_3,a_4)}\Big\{\frac{[-H(i,j)-L(i,j)]}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}\Big[\frac{\lambda_1}{\beta}(I_{[a_1=Accept]} + I_{[a_1=ltr]})R \\
& +\frac{\lambda_2}{\beta}(I_{[a_2=Accept]} + I_{[a_2=ltr]})R + \sum_{\forall(i',j')} p((i',j')|(i,j),(a_1,a_2,a_3,a_4))v^{\pi^*}(i',j')\Big]\Big\}
\end{aligned}
\tag{3.14}
$$

Alternatively, one could express the optimality equation as follows:

$$
\begin{aligned}
v^*(i,j) \quad = \quad & \frac{[-c_h(i^+ + j^+) - c_l(i^- + j^-)]}{\alpha+\beta} + \Big(\frac{\beta}{\alpha+\beta}\Big)\Big[\frac{\lambda_1}{\beta}T_1v^*(i,j) + \frac{\lambda_2}{\beta}T_2v^*(i,j) \\
& +\frac{\mu_1}{\beta}T_3v^*(i,j) + \frac{\mu_2}{\beta}T_4v^*(i,j)\Big]
\end{aligned}
\tag{3.15}
$$

where $T_kv^*(i,j)$ for $k = 1, ..., 4$ are defined for any real-valued function $v(i,j)$ as follows. For any real-valued function $v(i,j)$, let $T_1v(i,j)$ be an operator (for a customer arrival to D1) expressed as follows:

$$
T_1v(i,j) = max\{v(i-1,j) + R, v(i,j-1) + R\} \tag{3.16}
$$

For any real-valued function $v(i,j)$, let $T_2v(i,j)$ be an operator (for a customer arrival to D2) expressed as follows:

$$
T_2v(i,j) = max\{v(i-1,j) + R, v(i,j-1) + R\} \tag{3.17}
$$

For any real-valued function $v(i,j)$, let $T_3v(i,j)$ be an operator (for a production completion for D1) expressed as follows:

$$
T_3v(i,j) = max\{v(i+1,j), v(i,j)\} \tag{3.18}
$$

For any real-valued function $v(i,j)$, let $T_4v(i,j)$ be an operator (for a production completion for D2) expressed as follows:

$$T_4v(i,j) = max\{v(i,j+1), v(i,j)\} \tag{3.19}$$

For optimal policy π^*, system's optimal profit function $v^* \equiv v^{\pi^*}$ satisfies the optimality equation (Equation 3.15). By redefining the timescale and assuming $\alpha + \beta = 1$, without loss of generality, optimality equation becomes as follows:

$$v^*(i,j) = -c_h(i^+ + j^+) - c_l(i^- + j^-) + \lambda_1 T_1 v^*(i,j) + \lambda_2 T_2 v^*(i,j) + \mu_1 T_3 v^*(i,j) + \mu_2 T_4 v^*(i,j) \tag{3.20}$$

Equation 3.20 is the optimality equation that will be used for the centralized model in the following parts of the study.

3.3.1 CHARACTERIZATION OF THE OPTIMAL POLICY UNDER THE CENTRALIZED SYSTEM

In this section, monotonicity properties of the optimal policy of the central authority will be explained. Consider the base-stock level of D1, S_1^*. It is decreasing with the increasing inventory level of D2, i.e. j. Then, it can be said that optimal base-stock level is a function of j, i.e. $S_1^*(j)$, and it is decreasing in j. This example shows the use of monotonicity in this study.

Operators T_1 to T_4 (call T_n) are defined in Subsection 3.3 on the set of real-valued functions, v. Using the following theorem, it is shown that operators T_n preserve the structure of the function v and the optimal policy in the centralized system, called as **centralized optimal policy**, is characterized.

Theorem 3.3.1 *Let* $\mathbf{V}$ *be the set of real functions defined on* $\mathbb{Z}^2$*, such that if* $v \in \mathbf{V}$*, then*

C1': $v(i,j) - v(i-1,j) \geq v(i+1,j) - v(i,j), \forall i,j$

C2': $v(i,j) - v(i,j-1) \geq v(i,j+1) - v(i,j), \forall i,j$

C3': $v(i+1,j) - v(i,j) \geq +v(i+1,j+1) - v(i,j+1), \forall i,j$

C4': $v(i,j-1) - v(i-1,j) \geq v(i+1,j-1) - v(i,j), \forall i,j$

C5': $v(i,j) - v(i-1,j+1) \geq v(i,j-1) - v(i-1,j), \forall i,j$

If $\exists\, v \in \mathbf{V}$ that satisfies Equation 3.20, there exists an ***optimal operating policy*** *for the central authority that can be characterized with* $(S_1(j), K_1(j), Z_1(j), S_2(j), K_2(j), Z_2(j))$ *as follows.*

Let

$S_1(j) = \min\{i | v(i+1, j) - v(i, j) < 0\}$,

$S_2(i) = \min\{j | v(i, j+1) - v(i, j) < 0\}$,

$K_1(j) = \max\{i | v(i, j-1) - v(i-1, j) > 0\}$,

$Z_1(j) = \max\{i | v(i, j-1) - v(i-1, j) > 0\}$,

$K_2(i) = \max\{j | v(i, j-1) - v(i-1, j) < 0\}$,

$Z_2(i) = \max\{j | v(i, j-1) - v(i-1, j) < 0\}$,

where $S_1(j)$ and $S_2(i)$ denote the base-stock levels, $K_1(j)$ and $K_2(i)$ denote the rationing levels and $Z_1(j)$ and $Z_2(i)$ denote the transshipment levels at D1 and D2, respectively. The ***replenishment policy*** *is a state-dependent base-stock policy with base-stock levels $S_1(j)$ and $S_2(i)$, such that for $i < S_1(j)$ produce to replenish the stock and for $i \geq S_1(j)$ stop production and not replenish the stock at D1, for $j < S_2(i)$ produce to replenish the stock and for $j \geq S_2(i)$ stop production and not replenish the stock at D2. The* ***rationing and transshipment policy*** *is a state-dependent policy with rationing levels $K_1(j)$ and $K_2(i)$ and transshipment levels $Z_1(j)$ and $Z_2(i)$, such that for $i > K_1(j)$ and $j \leq Z_2(i)$ transship an item from D1 (or, redirect customer to D1) upon customer arrival to D2, for $j > K_2(i)$ and $i \leq Z_1(j)$ transship an item from D2 (or redirect customer to D2) upon customer arrival to D1.*

The optimal operating policy has the following properties:

$S_1(j)$ and $S_2(i)$ are non-increasing in j and i, respectively.

$K_1(j)$, $Z_2(j)$ and $K_2(i)$, $Z_1(j)$ are non-decreasing in j and i, respectively.

Proof. See Appendix B.

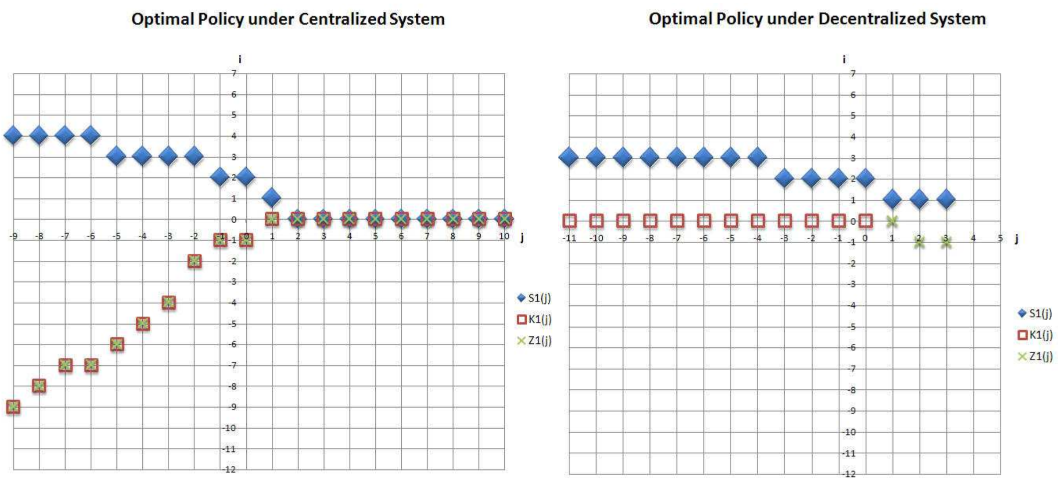
Please note that condition **C5'** is different from condition **C4** of decentralized system with only the direction of inequality.

3.3.2 COMPARISON OF THE OPTIMAL POLICIES UNDER DECENTRALIZED AND CENTRALIZED SYSTEMS

A comparison between decentralized system and centralized system is given here. The behavior of D1 with respect to D2's status changes in both systems.

Theorem 3.3.1 implies that the optimal policy is control-limit type and the control variables are monotone with respect to the state variables. Note that the definition of $K_1(j)$ and $Z_1(j)$ (or, $K_2(i)$ and $Z_2(i)$) are the same, and thus it is possible to define a single control variable that denotes both rationing and transshipment levels. Furthermore, note the dual structure between $K_1(j)$ and $Z_2(i)$ (or, between $K_2(i)$ and $Z_1(j)$); the definition of these control variables imply that for a given (i, j), if $j \leq Z_2(i)$, then $K_1(j) < i$.

Comparison of the optimal operating policy of an independent dealer with the optimal policy under the centralized system yields interesting insights. In the decentralized system D1 takes its best response actions given the operating policy of D2, whereas in the centralized system both dealers act to maximize the total profit of the system. Observation 3.2.2 and Theorem 3.3.1 show how the competitive nature of the decentralized system inversely affect the operating behavior of the dealers. In the centralized system, if the stock level at D2, j, decreases, then D1 shares its stock more by decreasing the rationing level K_1, and lowers down the amount of transshipment requests by increasing the transshipment level Z_1 considering the increased need of D2 for the available stock. On the other hand, in the decentralized system, if stock level at D2 decreases, then D1 decreases the proportion of requests that he accepts from D2 by increasing the rationing level. The reason is, when stock level decreases, D2 is more likely to request parts from D1 and meeting the increased transshipment requests (which bring lower revenues) may reduce the profit of D1. Similarly, observing a decrease in the stock level of D2, D1 increases the proportion of requests made by increasing the transshipment level, anticipating that D2 will share its stock less frequently. Figure 3.8 below shows the difference of optimal policies for decentralized and centralized system settings.



(a) Optimal Policy of D1 under Centralized System

(b) Optimal Policy of D1 under Decentralized System

Figure 3.8: Comparison of Optimal Policies of D1 under Centralized and Decentralized Systems for $(S_2, K_2, Z_2) = (3, 0, 0)$, $r = 3$, $c_h = 0.5$, $\lambda_1 = 0.5$ and $\lambda_2 = 0.5$

3.4 COMPUTATIONAL ANALYSIS

In Section 3.2, it is pointed out that optimal operating policy of D1 cannot be shown to be monotone on status of D2 using sufficient conditions, but observations indicate that monotonicity exist for most of the numerical instances. Some numerical models are solved and D1's policy for changing model parameters are analyzed. Solutions are obtained by using a linear programming model for the Continuous-Time Markov Decision Process (CTMDP) which is adopted from Puterman [1994] (p. 224). Objective function of the model is the maximization of **total discounted expected profit** of D1 subject to initial-state distribution, where discounting is applied with respect to time.

The computational study is performed for two purposes: (i) to analyze how the total discounted expected profit of an independent after-sales service provider is affected by the system parameters, (ii) to assess the benefit of collaboration by analyzing various pooling strategies. Specifically, four scenarios that define different levels of collaboration is designed to identify the effect of information sharing and inventory and service pooling on profitability. In terms of system parameters, the conditions that are most and least favorable to D1 are identified. To determine the total discounted expected profit of D1, the optimal operating policy of D1 under the total discounted expected profit criterion is considered by solving the following Linear Programming Model:

$$
\begin{aligned}
&\textit{Maximize} \quad \sum_{\forall s}\sum_{\forall a}\rho(s,a)X(s,a) \\
&\text{subject to} \\
&\qquad \sum_{\forall a} X(s,a) - \sum_{\forall s'}\sum_{\forall a}\left(\frac{\beta}{\beta+\alpha}\right)\times p(s|s',a)X(s',a) = \delta(s) \quad \forall s \\
&\qquad X(s,a) \geq 0 \qquad\qquad\qquad\qquad \forall (s,a)
\end{aligned}
$$

where,

s denotes the state (i, j),

a is the three-dimensional action vector (a_1, a_2, a_3) and is a function of s,

α is the discount rate,

$\delta(s)$ is the probability that the system starts in state s ($\delta(s) > 0$, $\sum_s \delta(s) = 1$)

$\rho(s, a)$ is the one-step reward function,

$\rho(s, a) = \frac{[-H(i,j)-L(i,j)]}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}\Big[\frac{\lambda_1}{\beta} I_{[a_1=Accept]}R + \frac{\lambda_1}{\beta}(I_{[a_1=Accept\&ltr]}I_{[K_2<j\leq S_2]})$
$\times(R - r) + \frac{\lambda_2}{\beta} I_{[T_2<j\leq Z_2]}I_{[a_2=Accept]}r\Big]$

$p(s|s', a)$ is the transition probability function,

$X(s, a)$ is the discounted fraction of visits to state s under action a, (definition of $X(s, a)$ is used exactly as it is in Puterman [1994] (p. 226.))
$X(s, a) = \sum_{n=1}^{\infty} (\frac{\beta}{\beta+\alpha})^{n-1} P\{s_n = s, a_n = a\}$

where the equation elements are defined in Section 3.1.

3.4.1 PERFORMANCE MEASURES

Total discounted expected profit of D1 is not the only performance measure considered in the analyses. Service level is considered as a performance measure, which is the discounted expected number of customers waiting in D1. Flow rates between the dealers are also used as performance measures.

Total discounted expected profit of D1 is the objective function used in the mathematical model and is defined as:

$$\begin{aligned}\textbf{PROFIT} &= \sum_{\forall s,\forall a} \rho(s, a)X(s, a) = \sum_{\forall s}\sum_{\forall a} \frac{-c_h i^+ - c_l i^-}{\alpha + \beta} X(s, a)\\ &+ R\Big(\sum_{\{s:j\leq K_2\}}\sum_{\forall a} \frac{\lambda_1}{\alpha+\beta}X(s, a) + \sum_{\{s:K_2<j\leq S_2\}}\sum_{\{a:a_1=Accept\}} \frac{\lambda_1}{\alpha+\beta}X(s, a)\Big)\\ &+ r\sum_{\{s:j\leq Z_2\}}\sum_{\{a:a_2=Accept\}} \frac{\lambda_2}{\alpha+\beta}X(s, a) + (R - r)\sum_{\{s:K_2<j\}}\sum_{\{a:a_1=ltr\}} \frac{\lambda_1}{\alpha+\beta}X(s, a)\end{aligned} \tag{3.21}$$

Total discounted expected profit of D1 has three revenue components, viz. expected discounted revenue by own customers (R1), expected discounted revenue by sending parts to D2 (R2) and expected discounted revenue by lateral transshipment from D2 (RLT).

R1 is generated by unit revenue R per customer. The following equation defines this revenue component.

$$R1 = \sum_{\{s:j\leq Z_2\}} \sum_{\forall a} \Big[\frac{\lambda_1}{\alpha+\beta} RX(s,a)\Big] + \sum_{\{s:Z_2<j\leq K_2\}} \sum_{\forall a} \Big[\frac{\lambda_1}{\alpha+\beta} RX(s,a)\Big] + \sum_{\{s:K_2<j\leq S_2\}} \sum_{\{a:a_1=Accept\}} \Big[\frac{\lambda_1}{\alpha+\beta} RX(s,a)\Big]$$

R2 is generated by unit revenue r per sending. The following equation defines this revenue component.

$$R2 = \sum_{\{s:j\leq Z_2\}} \sum_{\{a:a_2=Accept\}} \Big[\frac{\lambda_2}{\alpha+\beta} rX(s,a)\Big]$$

RLT is generated by unit revenue $R - r$ per sending. The following equation defines this revenue component.

$$RLT = \sum_{\{s:K_2<j\leq S_2\}} \sum_{\{a:a_1=ltr\}} \Big[\frac{\lambda_1}{\alpha+\beta} (R-r)X(s,a)\Big]$$

Service level **SL** is the total discounted expected number of customers waiting in D1's queue and is defined as:

$$\mathbf{SL} = \sum_{\{s:i<0\}} \sum_{\forall a} \frac{i^-}{\alpha+\beta} X(s,a)$$

Probability $\mathbf{P}_{from}$ is the total discounted fraction of time part transshipment from D2 is made and is defined as:

$$P_{from} = \sum_{\{s:j>K_2\}} \sum_{\{a:a_1=ltr\}} \frac{\lambda_1}{\alpha+\beta} X(s,a)$$

Probability $\mathbf{P}_{to}$ is the total discounted fraction of time transshipment requests of D2 is met and is defined as:

$$P_{to} = \sum_{\{s:j\leq Z_2\}} \sum_{\{a:a_2=Accept\}} \frac{\lambda_2}{\alpha+\beta} X(s,a)$$

3.4.2 TRUNCATION OF THE STATE-SPACE

In order to perform the numerical runs, the original state-space should be truncated. For status of D1, which is composed of integers from $-\infty$ to $+\infty$, consider lower and upper bounds I^l and I^u, such that $I = \{i : i = I^l, I^l + 1, \ldots, I^u - 1, I^u\}$. For status of D2, as S_2 is the upper bound on j, only a lower bound J^l is considered and range becomes $J = \{j : j = J^l, J^l + 1, \ldots, S_2 - 1, S_2\}$. To simplify, $I^l = J^l$ is assumed. To determine the values of I^l and I^u, there are several methods used in the literature. Puterman [1994] proposes using a value function u after truncation limits instead of the original value function v, which satisfies several properties with respect to one-step profit and one-period return operator. According to the decentralized model, the value function satisfying $u(i,j) \geq \frac{(1+\alpha)(\lambda_1 R+\lambda_2 r)}{\alpha(\alpha+\beta)}$ can replace $v(i,j)$ out of truncation limits. After finding u, truncation limits are increased by one and the difference between maximum and minimum differences in value function of the two models are checked to be less than a-priori tolerance ϵ. Ha [1997b] identifies limiting behavior of value function differences with respect to holding and waiting costs and interest rate. With adding those limiting differences to the conditions set defined for characterization of optimal policy, new set of conditions are satisfied by optimal operator. Truncation is determined by checking the change in the value function at the origin to be less than 0.00001% for each increment in the truncation limits.

Neither of the above methods is used in this study. Instead, the probabilities from the truncation limits are to the next states out of the truncation limits added to the limiting states, which means actually using the value function at the limits instead of u. The former method is not used because it is believed to bring little benefit in truncation. The latter is not used as the proof of optimal behavior failed using sufficient conditions. Instead, change in the objective function value is considered while as much as large values are tried to be used which do not hamper to make the runs in reasonable times. Reasonable time is considered to be less than 2 minutes on a Pentium 3.0 Ghz. processor computer with 2 GB RAM. The parameters are set to following values: $S_2 = 1$, $r = 3$, $c_h = 0.5$, $\lambda_1 = 0.7$, $\lambda_2 = 0.7$. Lowest c_h and highest λ_1 and λ_2 are selected to extend the used range of I for D1, and other parameters are set to moderate values. Two models are compared with ranges $I^l = -45$, $I^u = 45$ and $I^l = -46$, $I^u = 46$. Objective value of the model with $-I^l = I^u = 45$ is 117.43227 and the change in objective value to the model with $-I^l = I^u = 46$ is 0.0000094%. Therefore, $I^l = -45$ and $I^u = 45$ are selected for the numerical runs.

Solutions are obtained on a Pentium 3.0 Ghz. processor computer with 2 GB RAM using GAMS Optimization Package Version 20.2. CplexPar is used as the GAMS solver for linear programming models.

3.4.3 NUMERICAL SETTING OF PARAMETERS FOR THE ANALYSIS

In order to be more realistic for the values of parameters for the numerical setting, expert opinions are taken into consideration. Experts are selected from both industry and academia (based on suggestions of van Houtum [2008]).

For comparison purposes, three other pooling strategies other than dynamic-pooling strategy are considered. In full-pooling strategy, D1 and D2 share their resources fully. For this purpose, the available action $a_2 = Reject$ when D2 makes a lateral transshipment request is eliminated and $a_2 = Accept$ became the only available action. K_2 is set to the lowest j value, so that D2 becomes available to perform a lateral transshipment for D1 in any case. $Z_2 = 0$ is not changed meaning that D2 asks (and makes as $a_2 = Reject$ eliminated) for lateral transshipment in case of stock-out. Table 3.4 shows the setting of the numerical study for decentralized system for both dynamic-pooling and full-pooling settings. Full-factorial experiment is made for changing parameter values. For all numerical instances, $K_2 = Z_2 = 0$, $R = 10$, $c_l = 2$ and $\mu_1 = \mu_2 = 1$ are assumed. As a result, number of cases solved are 1800 for each setting, which is simply the multiplication of the numbers of different parameter values.

Table 3.4: Parameter Setting for the Numerical Study for Dynamic-Pooling and Full-Pooling Strategies

Parameters	S_2	r	c_h	λ_1, λ_2
Values	0,1,2,3	0,1,3,6,9,10	0.5,1,2	0.3,0.4,0.5,0.6,0.7

In no-pooling strategy, D1 operates isolated from D2. As there is not lateral transshipment between dealers, D1 makes no rationing. The only decision that D1 makes is his base-stock level. For the no-pooling strategy, parameters $S_2, K_2, Z_2, r, \lambda_2, \mu_2$ are irrelevant, as the model becomes an isolated system for D1. Table 3.5 shows the setting of the changing parameters for the numerical study. Total of 15 models are solved for no-pooling strategy.

Table 3.5: Parameter Setting for the Numerical Study for No-Pooling Strategy

Parameters	c_h	λ_1
Values	0.5,1,2	0.3,0.4,0.5,0.6,0.7

In static-pooling strategy, policy of D1 is characterized as (S_1, K_1, Z_1), which is independent of the status of D2, j. For this comparison, integer values of (S_1, K_1, Z_1) are searched in a range, where $0 \le S_1 \le 4$, $-2 \le K_1 \le 1$, $-2 \le Z_1 \le 1$. For a given combination of (S_1, K_1, Z_1), an evaluation of this given policy of D1 is made. In other words, based on the policy of D1 as (S_1, K_1, Z_1), unavailable actions are eliminated and the objective function value is evaluated. For example, for status of D1 equal to or greater than S_1, $a_1 = Accept$ is eliminated as production should not happen, and $a_1 = Reject$ is used as the only available action. Table 3.6 shows the setting of the numerical study for static-pooling model. Total of 32 cases are solved. For each case, different combinations of (S_1, K_1, Z_1) are considered. As explained above, there are 5 levels of S_1 and 4 levels of K_1 and Z_1, which constitute 80 combinations. The combinations are eliminated where $S_1 \geq K_1 \geq Z_1$ is not satisfied, and for each case, 46 combinations are used. Number of instances solved are 1472, and for each of the 32 cases, best solutions are selected.

Table 3.6: Parameter Setting for the Numerical Study for Static-Pooling Strategy

Parameters	S_2	r	c_h	λ_1	λ_2
Values	0,1,2,3	1,9	0.5	0.3,0.6	0.3,0.6

In the following section, the effect of the parameters λ_1, λ_2, r, S_2, and c_h on the total discounted expected profit of D1 are analyzed.

3.4.4 ANALYSIS OF PARAMETER EFFECTS ON SYSTEM PERFORMANCES

In the following subsections of this section, **profit** will refer to the average of the total discounted profits of D1, where average is taken with respect to all other parameters except the parameter in consideration. For example, in subsection 3.4.4.1, the effect of λ_1 on total discounted profits is analyzed and the term **profit** refers to the average of total discounted profits of D1, where average is taken for all other remaining parameters.

3.4.4.1 THE EFFECT OF ARRIVAL RATE TO D1 (λ_1) ON PROFIT

For the experimental setting considered, it is observed that as the arrival rate to D1 increases, the profit of the dealer increases for low values of λ_1 and then possibly decreases. D1 accepts all arriving customers and decides how to serve them, whether by own resources or by lateral transshipment. Numerical results show that R1 constitutes more than three-fourths of all revenue, on average, indicating that D1 serves his own customers by his own resources. R1 increases in λ_1, which is natural. R2 and

RLT are a small part of the revenue and their sum is nearly constant over λ_1, while the former decreases and the latter increases in λ_1.

In the experiments, it is observed that when $S_2 = 0$ and when revenue obtained from transshipment, r, is high, for high values of λ_2, the profit decreases as arrival rate increases (see Figure 3.9). The reason is, under $S_2 = 0$ upon each demand arrival, D2 places a transshipment request to D1 and therefore D1 faces the demand of customers arriving to both dealers. Especially when r is high, D1 meets a higher frequency of requests from D2 (i.e., K_1 is lower) and as a result effective arrival rate to D1 is high. An increase in the arrival rate results in an increase in customer waiting cost, and to avoid high waiting costs D1 keeps higher inventory. As a result, both inventory holding and waiting costs increase as arrival rate increases. Under heavy traffic, waiting cost and inventory holding cost dominates the revenue obtained, and the profit decreases as arrival rate increases. When $S_2 = 0$, for $\lambda_2 = 0.7$, D1 is as if operating under heavy traffic and thus the profit deteriorates with increasing arrival rate, while for $S_2 > 0$ the effective arrival rate is lower and the arrival rate at which the profit starts decreasing is higher (i.e., for $\lambda_1 = 0.7$ profit still increases with arrival rate). For other instances (when effective arrival rate to D 1 is not high) as λ_1 increases profit increases.

3.4.4.2 THE EFFECT OF ARRIVAL RATE TO D2 (λ_2) ON PROFIT

The analysis on λ_2 based on average profit may be insufficient or misleading if other parameter effects are not considered simultaneously. As the relation between profit and λ_2 highly depends on R2 and RLT, the relation is also considered under different r and S_2 values. Figure 3.10 shows the relation under r values, for example. The behavior of profit changes with respect to r level.

As rate of arrivals to D2 increases, the profit of D1 may or may not increase depending on the base-stock level of D2, and the amount of payment, r. As arrival rate to D2 increases, the percentage of the time D2 places transshipment requests increase, while the percentage of time the requests of D1 are met, decreases. If the amount

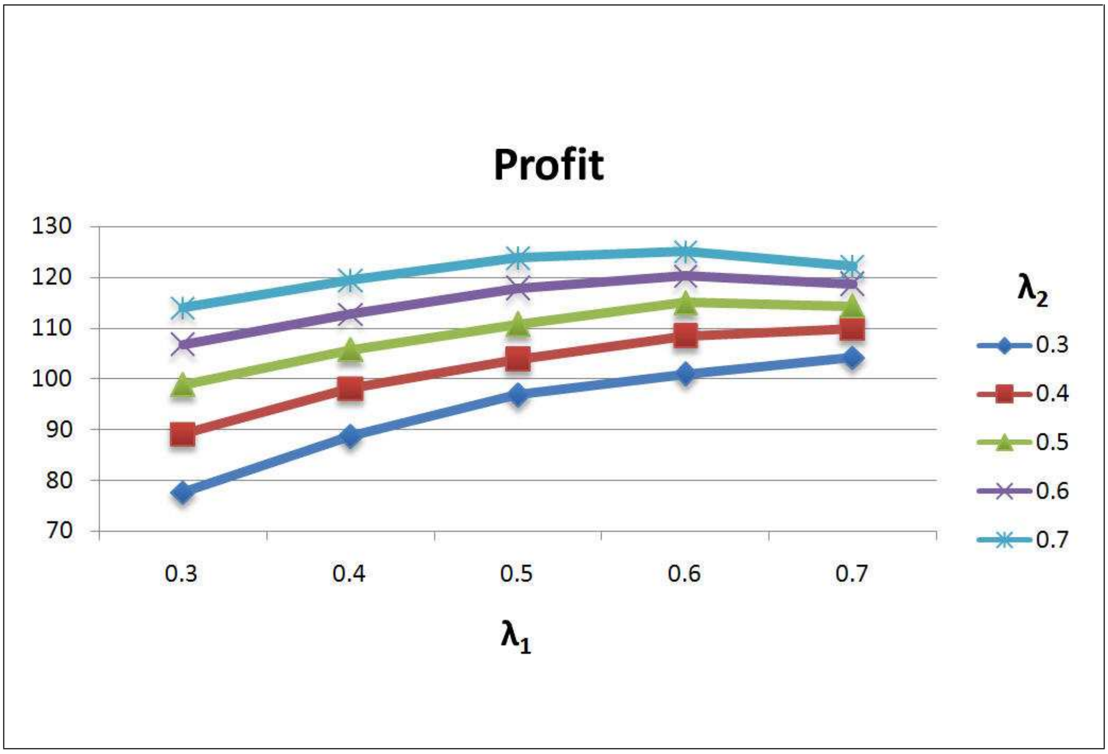


Figure 3.9: Profit versus λ_1 under λ_2 for the Dynamic-Pooling Strategy for $r = 10$, $S_2 = 0$ and $c_h = 2$

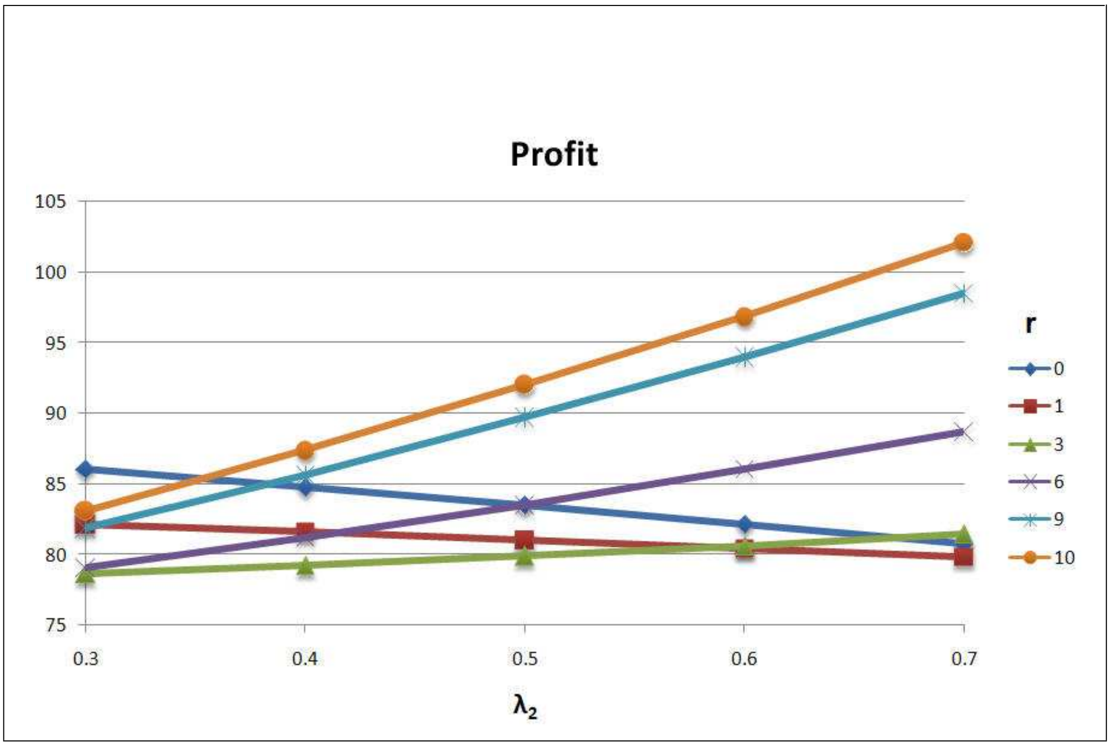


Figure 3.10: Profit versus λ_2 under Payment Amounts for the Dynamic-Pooling Strategy

of payment upon request is high, then D1 benefits from an increase in arrival rate to D2. Otherwise, if the payment amount is low, then D1 prefers lower rates of arrival to D2. For instance, when $r = 0$ optimal rationing level is equal to optimal base-stock level ($S_1(j) = K_1(j)$) and D1 does not meet the lateral transshipment requests of D2, while continues placing requests to D2. As λ_2 increases, D2 is less likely to meet the requests of D1, and more likely to request parts from D1, which results in a decrease in the profit of D1. Results show that there exists a threshold level for r, above which D1 benefits an increase in arrival rate to D2.

For high payment amount, a more detailed analysis on revenue components and probabilities is as follows. As seen from Figure 3.10, profit increases in λ_2 for high r. As D1 has the available actions of rejecting lateral transshipment request of D2 and asking for lateral transshipment from D2 (where realization depends on status of D2, j, of course), D1 finds more room to exploit D2. With increasing λ_2, D1 can use his available actions of rejecting lateral transshipment request of D2 and asking for lateral transshipment from D2 more effectively and exploit D2's resources by controlling the related revenues, R2 and RLT. Numerical results also support this and show that the relation between profit and λ_2 highly depends on R2 and RLT. For the lowest λ_2, RLT is more than twice of R2, which is reversed at the highest λ_2 level, i.e. R2 becomes almost twice of RLT. At the same time, R1 increases with λ_2. Combining all together, D1 serves more customers by getting less lateral transshipment from D2 with increasing λ_2, but also finds it profitable and sends more to D2. This course of D1 to exploit D2 is also supported by discounted probabilities below.

The followings are observed for probabilities, again for for high payment amount. For higher λ_2 values, D2 can send fewer parts to D1, but asks for more lateral transshipment from D1. Although D1 rejects more requests, number of accepted requests also increases. By increasing operational intensity, D1 can ask for fewer parts as D2's net inventory level is rarely available. The situation is just reverse for lower λ_2 values.

When combined effect of both λ_1 and λ_2 on profit are considered simultaneously, the system's operation intensity is beneficial for D1 and profit increases with increasing λ_1 and/or λ_2. Note in Figure 3.10 that under changing λ_2 values, highest profit is either achieved under $r = 1$ or $r = 10$, which is a result in line with Proposition 3.2.3.

3.4.4.3 THE EFFECT OF PAYMENT AMOUNT (r) AND BASE-STOCK LEVEL AT D2 (S_2) ON PROFIT

A strong correlation between the payment amount and base-stock level at D2 is observed. Therefore, the analysis on r and S_2 is performed simultaneously.

In the numerical study, four levels of S_2 are considered. $S_2 = 0$ (together with $K_2 = Z_2 = 0$) means D2 never shares his inventory with D1 while asks for lateral transshipment from D1 for all customer arrivals. $S_2 = 1$ means D2 only shares his inventory with D1 at his base-stock level, and asks for lateral transshipment from D1 for customer arrivals when he is out of stock. The remaining S_2 levels can be interpreted similarly.

An increase in base-stock level at D2 may or may not increase the profit of D1, depending on the payment amount. As S_2 increases, D2 is more likely to meet the transshipment requests of D1, and is less likely to place lateral transshipment requests. If the payment amount is low, then D1 benefits from an increase in S_2 and prefers high S_2 levels. On the other hand, under high payment amounts, an increase in S_2 results in a decrease in profit of D1. See, for instance in Figure 3.11 that for $6 \leq r$ the profit of D1 decreases, while for $r < 3$ the profit increases in S_2.

Similarly, whether an increase in the payment amount r increases or decreases the profit of D1 depends on the base-stock level, S_2. If S_2 is high (low), then transshipments to D1 is likely to dominate the transshipments met by D1 and the after-sales service provider prefers low (high) payment levels. Figure 3.12 below shows this situation. If S_2 is at moderate levels (for our setting, if $S_1 = 1$), then it is not clear whether low or high payment levels are preferred, and analysis shows that λ_1 determines the preferred r level. For low λ_1 levels, transshipments met by D1 dominate the transshipments placed by D1, and high payment levels are preferred.

3.4.5 COMPARISON OF POOLING STRATEGIES

In this section, first the four pooling strategies that will be analyzed are introduced, namely full-pooling, static-pooling, dynamic-pooling and no-pooling, then compari-

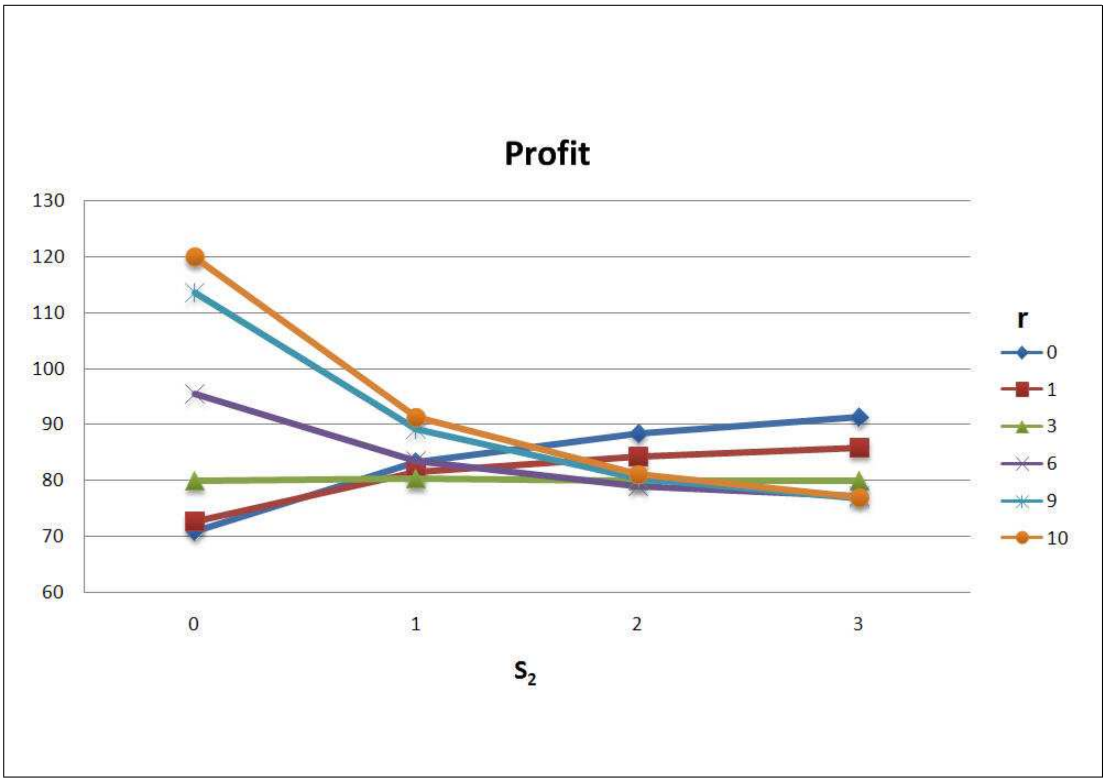


Figure 3.11: Profit versus S_2 under Payment Amounts

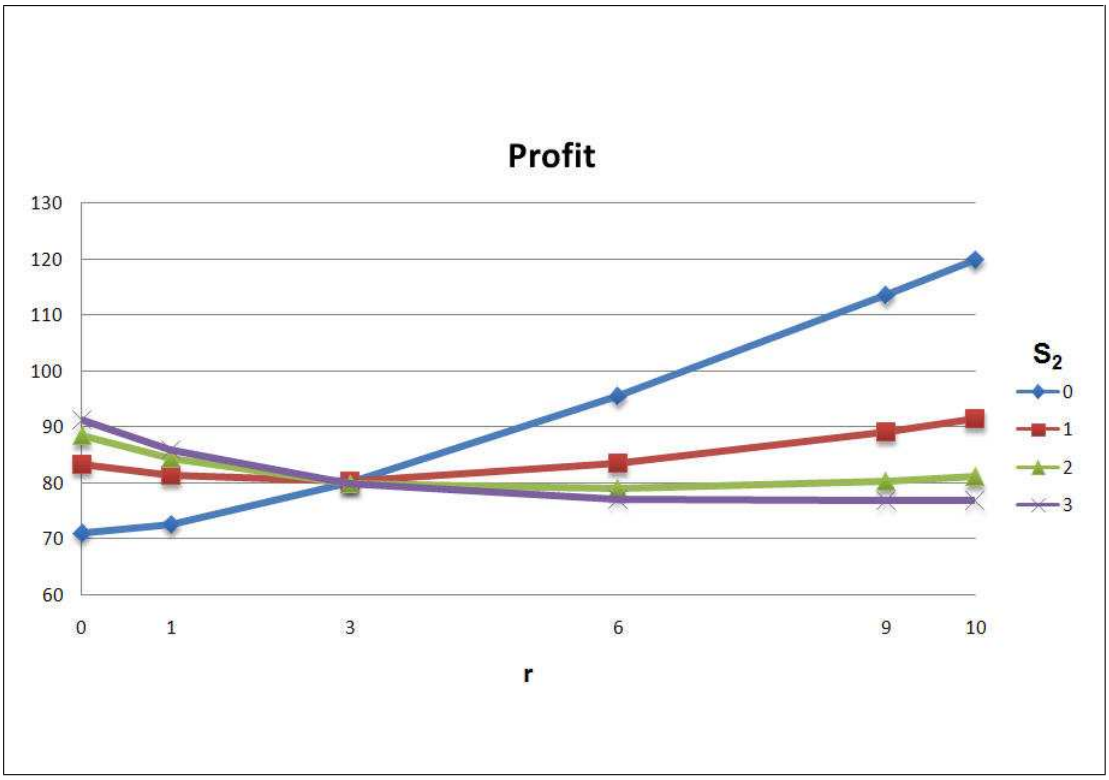


Figure 3.12: Profit versus r under S_2 for Dynamic-Pooling Strategy

son of the four pooling strategies is presented.

1. In the *no-pooling* strategy, D1 and D2 do not collaborate, i.e., neither share information on inventory status, nor share items or demand through transshipment. In other words, the existence of D2 does not affect the operations of D1. Under no-pooling strategy, D1 maximizes the total discounted reward, and determines the optimal base-stock level.

2. In the *full-pooling* strategy, a case where dealers share their resources fully is modelled, i.e., place and/or meet transshipment requests non-optimally. This scenario is designed to address the question of whether collaboration through transshipment is better than no collaboration, even if the collaborative agreement is not designed optimally. In the full-pooling strategy, rationing level of D1 does not exist, and D1 always meets transshipment requests of D2. Similarly, D2 always meets transshipment requests placed by D1, but places requests only when it stocks out, which implies $Z_2 = 0$. D1 maximizes the total discounted reward, and determines the optimal base-stock level and the optimal transshipment request level.

3. In the *dynamic (or optimal) pooling* strategy, D1 operates under the optimal operating policy by determining the base-stock, rationing and transshipment request levels. As the analysis in Section 3.1 shows that these control levels dynamically change with the inventory level at D2.

4. In the *static pooling* strategy, the case where the dealers do not share inventory status information but only collaborate through transshipments is mimicked. Since D1 is not updated about the inventory level at D2, optimal base-stock, rationing and transshipment levels are static with respect to inventory level at D2. Note that such a static policy is not necessarily the optimal operating policy of D1, since the optimal policy could be a randomized policy, which will be considered in Chapter 4. D1 determines the base-stock, rationing and transshipment request levels that maximize the total expected discounted profit, under the restriction that these levels do not change with the inventory level at D2.

3.4.5.1 COMPARISON OF NO-POOLING AND DYNAMIC-POOLING STRATEGIES

No-pooling and dynamic-pooling strategy are compared for λ_1. For the comparison, the percentage increase in profit of D1 under dynamic-pooling $\frac{\overline{\Pi}_{DP}-\overline{\Pi}_{NP}}{\overline{\Pi}_{NP}} \times 100\%$ will be used, where $\overline{\Pi}_{DP}$ and $\overline{\Pi}_{NP}$ denote the profit of D1 for dynamic-pooling and no-pooling strategies. The ratio is always expected to be positive since no-pooling is a feasible solution of the pooling problem. The comparison shows that if pooling is done optimally then collaboration through information and transshipment is always to the benefit of the dealer, and the amount of benefit can be substantial. A further observation is that optimal base-stock level under pooling can be higher or lower than the optimal base-stock level under no-pooling, depending on the effective arrival rate to D1.

Figure 3.13 shows how the benefit of pooling changes as arrival rate to D1 changes. The results are obtained under $S_2 = 2$ and are averaged over unit holding cost and arrival rate to D2. As arrival rate to D1 increases: (i) if payment level is low, the benefit of pooling increases, while (ii) if payment level is high, the benefit of pooling has a tendency to decrease. A higher arrival rate to D1 results in a higher fraction of transshipment requests placed to D2 and a lower fraction of satisfied requests placed by D2. This results in an increase or decrease in D1's profits depending on the payment level.

Consider $r = 6$ in Figure 3.13. The benefit of dynamic-pooling is at least 9% compared to the D1's profit of no-pooling strategy. The ratio is convex and is higher at extreme λ_1 values. This shows that D1 can exploit D2 more compared to no-pooling for extreme λ_1 levels, D1 can exploit D2 more compared to no-pooling strategy. For higher λ_1, D1 can meet more of his customers via D2. For lower λ_1, D1 can be better off by sending more to D2 and by this way utilizing his production capacity.

The benefit of pooling is most significant when the system operates under parameter values under those extremes for $r \geq 6$. When S_2 is very low, payment level is very high, arrival rate to D2 is high, and arrival rate to D1 is low, the benefit of pooling can be as high as 180%. In the opposite case, when S_2 is high, payment level is very low,

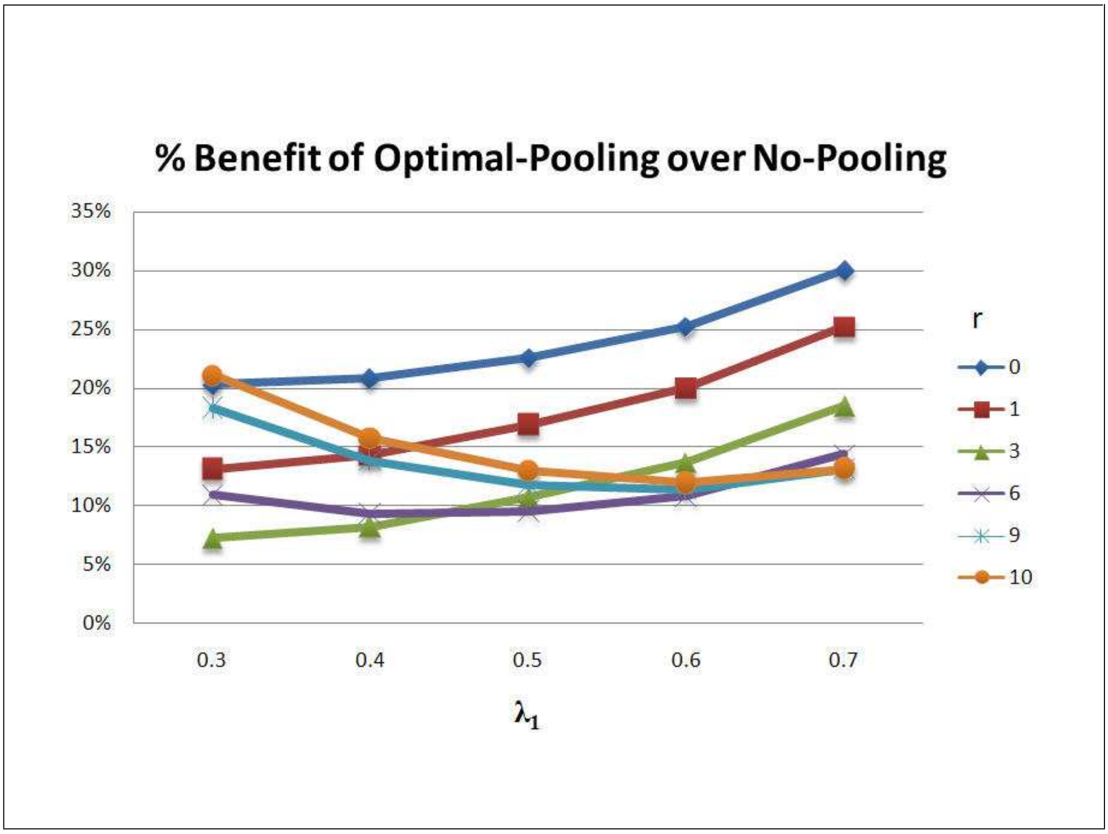


Figure 3.13: Profit Increase in Percentage versus λ_1 under r for $S_2 = 2$

arrival rate to D2 is low, and arrival rate to D1 is high, then the benefit of pooling is again relatively high at 55%. In the former case, D2 is totally dependent on D1 and totally exploits the capacity of D1, while in the latter case D1 exploits the capacity of D2.

Comparison of expected discounted number of waiting customers in the system (service level, SL) under no-pooling and optimal-pooling strategies show that in 1357 out of 1800 instances, expected number of customers in the system are higher under optimal-pooling. As pooling becomes more beneficial, the number of customers waiting in the system increases. There are instances where although benefit of pooling is negligible, number of customers in the system may increase as much as 245%. The reason is for higher number of customers under pooling is the instances where the other dealer does not meet orders placed by D1, but places transshipment requests (when S_2 is low). In that case, D1 operates under higher effective arrival rates and although profits improve, service level may deteriorate at D1, while it improves in D2. A final note is that for a given instance, higher expected number of waiting customers under pooling does not necessarily imply that expected waiting time of an arriving customer will be higher. An arriving customer may wait shorter under pooling, since transshipment of a part or redirecting the customer to D2 may result in lower waiting time.

3.4.5.2 COMPARISON OF NO-POOLING AND FULL-POOLING STRATEGIES

Comparison between no-pooling and full-pooling strategies is made. Although full-pooling might be considered better than no-pooling strategy, there are cases where no-pooling is better for D1. The percentage increase in profit of D1 is defined as $\frac{\overline{\Pi}_{FP}-\overline{\Pi}_{NP}}{\overline{\Pi}_{NP}} \times 100\%$, where $\overline{\Pi}_{FP}$ and $\overline{\Pi}_{NP}$ denote the profits od D1 for full-pooling and no-pooling strategy. For example, consider the no-pooling strategy with parameters $c_h = 0.5$ and $\lambda_1 = 0.3$. Profit of D1 for this setting is 48.17. Full-pooling setting with the same parameter values in addition with $S_2 = 0$, $r = 0$ and $\lambda_2 = 0.7$ has a profit of 22.00, which is less than half of the no-pooling strategy. For $c_h = 2$, the difference becomes enormous and the profits are 44.41 and 3.46, respectively for the same parameters. With increasing λ_1 value, such rare cases disappear and full-pooling always become better than no-pooling. Maximum difference occur for parameters $S_2 = 0$, $r = 0$ and $\lambda_2 = 0.7$, but there is no commonality among the parameters which lead to minimum difference. In addition, in decreasing order of the difference, the corresponding parameters do not have a lexicographical ordering. The summary of comparison results are in Table 3.7 below.

In Table 3.8, all the problem instances under full-pooling strategy that are worse than no-pooling strategy for parameters $c_h = 0.5$ and $\lambda_1 = 0.3$ is given using the above ratio. Similar to Table 3.7, the parameters in the ordered difference ratio does not have a lexicographical ordering.

The results imply that it may or may not be beneficial to pool inventory if the pooling strategy is not determined optimally. Figure 3.14 shows the parameters under which no-pooling strategy results in higher profits for D1 than full-pooling. Considering Tables 3.7 and 3.8, and Figure 3.14 together, the following conclusions can be made about parameter effects. When S_2 is low, under low levels of payment full-pooling

Table 3.7: Problem Instances Having Lower Profits under Full-Pooling Strategy Compared to No-Pooling Strategy

c_h	λ_1	# of instances	Average Difference	Minimum Difference	Maximum Difference
0.5	0.3	12	17.0 %	1.5 %	54.3 %
0.5	0.4	8	13.4 %	1.8 %	34.7 %
0.5	0.5	4	11.5 %	3.0 %	20.9 %
0.5	0.6	2	8.7 %	6.6 %	10.9 %
1.0	0.3	13	23.0 %	1.5 %	71.8 %
1.0	0.4	9	16.0 %	1.8 %	44.4 %
1.0	0.5	6	12.1 %	0.8 %	28.1 %
1.0	0.6	2	11.6 %	9.5 %	13.7 %
2.0	0.3	21	24.3 %	1.1 %	92.2 %
2.0	0.4	11	22.2 %	0.9 %	58.4 %
2.0	0.5	5	19.6 %	9.6 %	34.4 %
2.0	0.6	3	10.2 %	0.7 %	16.3 %

Table 3.8: Problem Instances Having Lower Profits under Full-Pooling Strategy Compared to No-Pooling Strategy for Parameters $c_h = 0.5$ and $\lambda_1 = 0.3$

S_2	r	λ_2	Difference
0	0	0.7	54.3 %
0	1	0.7	35.7 %
0	0	0.6	33.6 %
0	0	0.5	20.2 %
0	1	0.6	19.1 %
1	0	0.7	12.8 %
0	0	0.4	7.9 %
0	1	0.5	7.7 %
1	1	0.7	5.1 %
1	0	0.6	4.9 %
0	1	0.4	1.7 %
2	0	0.7	1.5 %

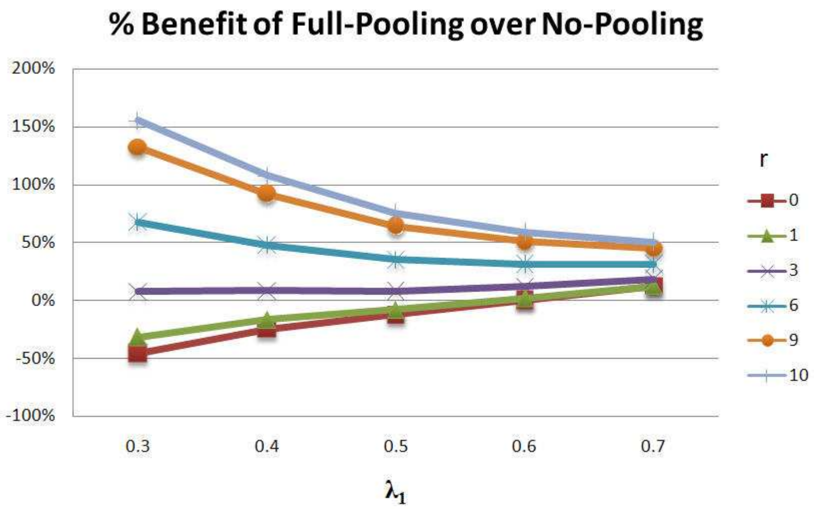
gives lower profits for D1 than no-pooling, especially if arrival rate to D1 is low. This constitutes a least favorable setting for D1, since the low payment level does not make up for the high frequency of requests by D2, and thus full-pooling results in a decrease in D1's profit compared to no-pooling strategy. As arrival rate to D1 increases, D1 starts effectively using service capacity of D2 and full-pooling is preferable to no-pooling. If payment level increases, this also makes full-pooling strategy a more preferable choice (see Figure 3.14.a).

For low levels of S_2, especially if arrival rate to D2 is high, full-pooling strategy gives lower profits of D1 than no-pooling strategy. This is the least favorable setting for D1, since with low S_2 and high λ_2 levels it is D2 that mostly exploits resources of D1 (see Figure 3.14.b). Increase in benefit due to full-pooling can be as high as 168% and as low as -54%, both achieved under $S_2 = 0$.

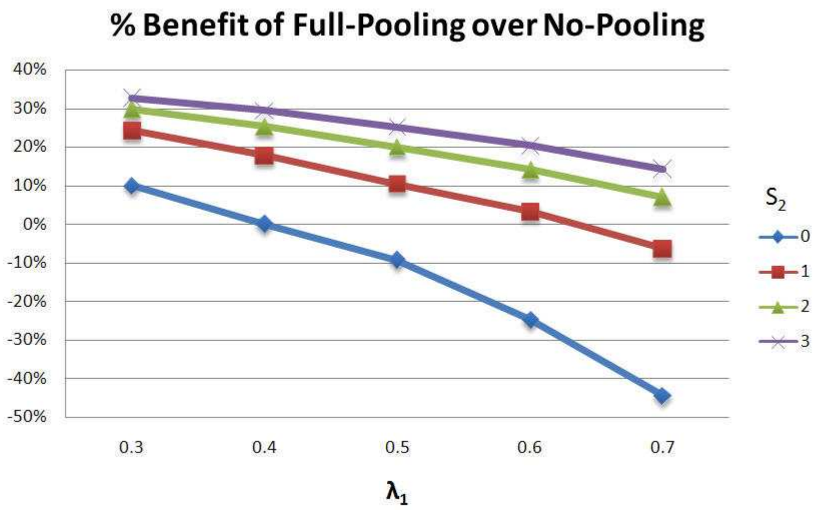
Full-pooling, no-pooling and dynamic-pooling strategies are considered together for the behavior of their profits of D1 with respect to λ_1 under parameter values $S_2 = 0$ and $r = 1$. In Figure 3.15, D1's profit of dynamic-pooling lies always above no-pooling, indicating dynamic-pooling strategy is dominant to no-pooling strategy. Profit of D1 for full-pooling strategy is less than both no-pooling and dynamic-pooling strategies for lower λ_1 values, but becomes higher than both with increasing λ_1. Full-pooling may give better results compared to optimal-pooling, because D1 can use D2's resources without a limit under full-pooling strategy, which is not possible under optimal-pooling strategy where D2's policy is given. For higher λ_1 values, D2's resources become relatively important for D1, which explains this behavior.

3.4.5.3 COMPARISON OF STATIC-POOLING AND DYNAMIC-POOLING STRATEGIES

The comparison between dynamic-pooling and static-pooling strategies is made as follows. For each dynamic-pooling setting, the best static-policy of D1 as (S_1, K_1, Z_1) in terms of the objective function value is selected and compared with the corresponding dynamic-pooling solution. For example, for the dynamic-pooling parameters $S_2 = 2, K_2 = 0, Z_2 = 0, r = 9, R = 10, c_h = 0.5, c_l = 2, \lambda_1 = 0.6, \lambda_2 = 0.6, \mu_1 = 1, \mu_2 = 1$, static-pooling policy $(S_1 = 3, K_1 = -1, Z_1 = -2)$ is the best within the nu-



(a) Percentage Benefit versus λ_1 for r



(b) Percentage Benefit versus λ_1 for S_2

Figure 3.14: Profit Increase in Percentage under Full-Pooling Strategy with respect to No-Pooling Strategy

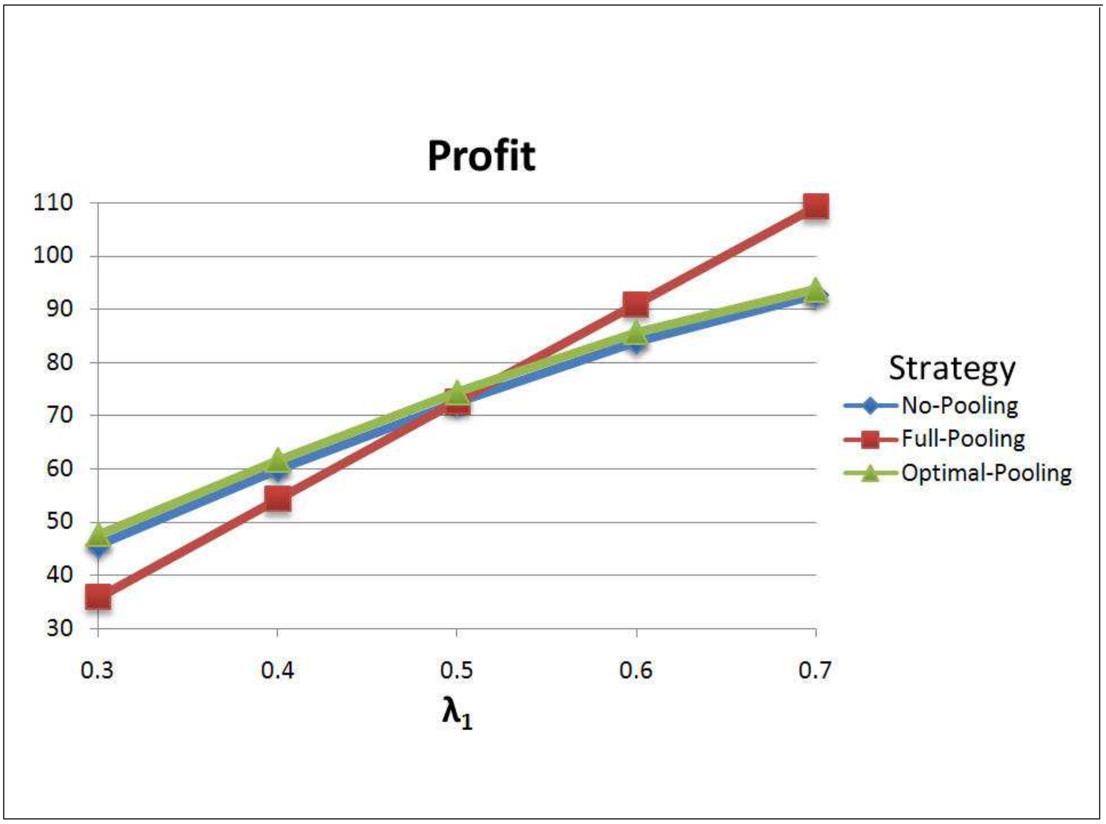


Figure 3.15: Comparison of Full-Pooling, No-Pooling and Dynamic-Pooling Strategies with respect to Profit versus λ_1 for $S_2 = 0$ and $r = 1$

merical setting. The percentage increase in profit of D1 is defined as $\frac{\overline{\Pi}_{DP}-\overline{\Pi}_{SP}}{\overline{\Pi}_{SP}} \times 100\%$ is used, where $\overline{\Pi}_{DP}$ and $\overline{\Pi}_{SP}$ denote the D1's profits for dynamic-pooling and static-pooling strategies. The comparison of D1's profits obtained under static and optimal-pooling shows that if the dealers already collaborate through pooling, then the additional benefit of information sharing on inventory status is insignificant. For the 1,684 instances (out of 1,800) the benefit of dynamic policy over static policy is less than 1.5%, and in total the benefit obtained under dynamic policy is limited by 4%.

Please remember that, static-pooling policy is found by evaluation, not by optimization to find (S_1^*, K_1^*, Z_1^*). Limited search over D1's policies is performed and the results indicate that, using even a static-policy, which is not necessarily optimal, a D1's profit value for the static-pooling strategy can be found which is very close to the D1's profit of optimal dynamic-pooling strategy.

The parameter settings where benefit is in the range of 0% – 1.5% versus in the range of 1.5% – 4% is analyzed and it is observed that when $S_2 = 0$ the optimal static base-stock level at D1 can be as high as $S_1 = 9$ (especially if holding cost is low and arrival rate is high), while for $S_2 = 2$ or $S_2 = 3$ optimal base-stock level is at most $S_1 = 3$. When S_2 is low, dynamic-pooling policy operates very close to a static policy (i.e., base-stock level, rationing and transshipment levels do not change with inventory status at D2) and the benefit of dynamic-pooling compared to static-pooling is very low. Under low S_2 levels, D2 almost always places requests to D1 upon customer arrival. In this case observing the inventory status at D2 does not bring any benefits over not observing, and therefore dynamically changing the control variables does not bring benefit over static levels. On the other hand, under high S_2 levels, information on inventory status is relatively more beneficial and the optimal control variables at D1 necessarily change with the inventory level at D2. In this case the optimal base-stock level at D1 is quite low, therefore keeping even one unit of stock affects the profit of D1. To set the base-stock level totally to 0 or to 1 deteriorates the profits of D1 compared to dynamically changing base-stock level. A final note is that it is not possible to characterize the amount of benefit changes with the system parameters, since the nature of the system is complex, and the benefit depends on the specific structure of the dynamic policy.

Comparing the expected discounted number of waiting customers in the two systems, there are instances where static policy results in lower number of customers or higher number of customers in the system. The difference in the number of waiting customers could be as low as -55% or as high as 129%. Although profit of D1 obtained under dynamic policy is slightly higher than profit of D1 under static policy, it is not possible to observe a pattern in the waiting number of customers.

The result that the D1's profits of optimal-pooling and static-pooling strategies are close to each other (even the policy found for static-pooling strategy is not necessarily optimal) is very important in practical meaning. It may not be very easy to implement a dynamic-pooling policy $(S_1(j), K_1(j), Z_1(j))$ which depends on status of D2, j, in practice. Therefore, in practical applications, static-pooling strategy might be adopted with a negligible sacrifice in profit, but with a significant gain in easiness of implementation.

CHAPTER 4

ANALYSIS OF BENEFITS OF INFORMATION AVAILABILITY IN THE PRESENCE OF POOLING

In this chapter, benefits of information availability in the presence of pooling is analyzed. A system consisting of two independent dealers in a decentralized system is assumed, where both dealers are rational, i.e. they make the same decisions under the same conditions. The spare parts system is assumed to be stable and the conditions defining the system (economic, structural, etc.) are not changing over time. One of the dealers is assumed to have a known policy and the other dealer's profit maximization problem is considered. Three different information availabilities are analyzed and the models used for the analyses are explained in Section 4.1. Optimal decisions taken under different information availabilities are compared in Section 4.2. Finally, the method and solution strategy used for computational analysis, numerical setting of parameters and the results of the analysis in terms of parameter effects are discussed in Section 4.3.

4.1 REDEFINING THE SYSTEM UNDER PARTIAL INFORMATION

In partial information system setting, the same decentralized spare part service network is considered as in the full-information setting. However, the dealer under consideration is assumed to have incomplete information on the net inventory levels and makes his decisions under partial information availability. Three different levels of information availability for the dealer under consideration is assumed, viz. incomplete information on the other dealer's status, on the number of waiting customers in his queue, and on the amount of on-hand inventory.

The following additional components are introduced to the model compared to the full-information model in Chapter 3: customer rejection decision, ordering cost, production cost, variable payment tariff, customer rejection cost, transshipment cost. These components are explained below.

Decisions of the dealer are similar to the ones in decentralized system of Chapter 3, except another control mechanism, namely the **customer rejection decision** upon customer arrival. Whenever a customer arrives to the dealer, the dealer may give service to the customer using his own resources, may reject giving service to the arriving customer or may ask the other dealer to use his resources, where the first action decreases status of the dealer under consideration by one unit, the second action decreases net inventory level of the other dealer by one and the last action conserves the net inventory level of the dealer under consideration. The dealer can decide to continue or stop production in order to control his inventory level, as in Chapter 3. Information availability has effects on these decisions.

Consider the lateral transshipment request decision under incomplete information on the other dealer's net inventory level. As the dealer under consideration has incomplete information on the other dealer's inventory level, he may not be sure whether his request will be fulfilled or not. No-information case is a special case of incomplete information on the other dealer's status, where the dealer cannot know whether a lateral transshipment request will be fulfilled or not. In existence of a fixed ordering cost, full-information allows him to know fulfillment status and therefore pay the ordering cost only when the request is certain to be fulfilled. Partial-information or no-information may lead to *pay for nothing* situations, i.e. ordering cost is paid although request is not fulfilled.

For customer rejection decision, if the dealer has incomplete information on the number of waiting customers in his queue, his decision quality might be poor and his profit may deteriorate. He can control his service queue under full-information availability, but incomplete information on waiting customers does not allow him to control his queue.

As a last example, consider incomplete information on the amount of on-hand inventory. Upon a lateral transshipment request by the other dealer, the dealer is free to accept such a request and send a part or reject the request, but this decision is affected by the information availability. Under a variable payment situation, where the payment depends on the stock level, this information availability effect might be larger. Also the continue or stop decision about production of the dealer depends on the information on amount of on-hand inventory. In presence of unit production cost, the effect of the production decision's quality on the dealer's profit is larger.

The decision of customer rejection makes the operation strategy of the dealers different from the one in Chapter 3, which is assumed to be a threshold policy characterized by four levels as (S, K, Z, T), where S, K, Z and T are base-stock, rationing, transshipment and rejection levels, respectively.

For each lateral transshipment request, the dealer under consideration pays a fixed ordering cost c_o. For partial or no information availability, the dealer may not know whether a lateral transshipment request will be fulfilled by the other dealer or not upon a customer arrival, whereas he knows the result of his request in full-information case. Due to the fixed ordering cost, intuitively, information on status of the other dealer is expected to be more valuable compared to no fixed ordering cost. This intuition is supported by a small numerical study. Value of information availability has increased significantly with the existence of ordering cost. The small numerical study shows that information on the other dealer's net inventory level is more valuable when fixed ordering cost exists.[1]

The dealer under consideration is assumed to pay a unit production cost, c_p, per unit time for a continuing production. The decisions $a_3 = Accept$ and $a_3 = Reject$ cause the dealer to pay or not to pay c_p.

Consider a case where there is inventory at the other dealer, more specifically assume inventory level at the other dealer is greater than his rationing level. In this situation, relative value of an item is lower for the other dealer and he will be willing to sell on-

[1] To measure the value of information availability, average of percent profit deviation is used, which is $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{no}}{\overline{\Pi}_{full}} \times 100\%$, where $\overline{\Pi}_{full}$ and $\overline{\Pi}_{no}$ denote the profit for full-information and no-information availabilities, respectively. For a small numerical study with 128 instances, percent profit deviation averaged 2.99% and 9.17% for $c_o = 0$ and $c_o = 3$, respectively.

hand item(s). In this case, the payment that the other dealer may accept could be low. Low payment is of course preferable for the dealer under consideration when lateral transshipment request is made. In the opposite case, specifically when customers are waiting at the other dealer, i.e. when inventory level at the other dealer is less than or equal to his transshipment level, the other dealer will be willing to pay a high r. Higher payment improves the profit of the dealer under consideration while meeting the other dealer's requests. The value of information is analyzed under this variable payment setting in addition to the constant payment setting. One may argue that, in practice, there will be a negotiation taking place between the dealers and for each (i, j) an equilibrium payment would be determined. However, those cases are not incorporated to the model for the time being. Variable payment is assumed to be decreasing in increasing status level of the other dealer and to be independent of the status of the dealer under consideration.

For each rejected customers, the dealer under consideration pays a fixed amount of rejection cost, c_r. The decisions $a_1 = Accept$, $a_1 = ltr$ and $a_1 = Reject$ cause the dealer to pay or not to pay c_r, where the cost is paid for the latter decision only.

Fixed ordering cost c_o is paid for each lateral transshipment request, but transshipment cost c_t is only paid when lateral transshipment is realized. Therefore, for a successful lateral transshipment, both c_o and c_t are paid, but when lateral transshipment request is not met, only c_o is paid.

In the following sections, the mathematical model for solution and partial information models used for different information incompleteness situations are given. The assumptions used for all the models are the same with the ones used in Chapter 3.

4.1.1 THE MATHEMATICAL MODEL UNDER PARTIAL INFORMATION SYSTEM

The following relation is the key of forming the non-linear mathematical programming model.

$$\alpha(a|k) = \frac{X(s,a)}{X(s)} \quad \forall s \in S_k, a \in A, k \in K$$

where $\alpha(a|k)$ is the conditional probability of choosing action a given that the observation process is at state k, $X(s,a)$ is the probability of choosing action a given that the internal process is at state s, $X(s) = \sum_{a \in A} X(s,a)$, $\forall s \in S$. This relation is actually putting a restriction on actions for a specific observation process state k and forces to take the same action at corresponding internal process states, $s \in S_k$.

Objective function of the model is maximization of **average expected profit under incomplete information** of the dealer under consideration (D1). To determine the average expected profit of D1 and the optimal operating policy of D1, the following Non-Linear Programming Model is introduced:

$$Maximize \quad \sum_{\forall k} \sum_{s \in S_k} \sum_{\forall a} \rho(s,a)\alpha(a|k)X(s)$$

$$subject \quad to$$

$$X(s) - \sum_{\forall k} \sum_{s' \in S_k} \sum_{\forall a} p(s|s',a)\alpha(a|k)X(s') = 0 \quad \forall s$$

$$\sum_{\forall a} \alpha(a|k) = 1 \quad \forall k$$

$$\sum_{\forall s} X(s) = 1$$

$$X(s) \geq 0 \quad \forall s$$

$$\alpha(a|k) \geq 0 \quad \forall k, \forall a$$

where one-step reward function $\rho(s,a)$ is as follows:

$$\rho(s,a) = \frac{[-H(i,j)-L(i,j)]}{\alpha+\beta} - \frac{c_p}{\alpha+\beta} I_{[a_3=Accept]} + \frac{\beta}{\alpha+\beta}\Bigg[(\frac{\lambda_1}{\beta} I_{[a_1=Accept]} + I_{[j\leq K_2]} I_{[a_1=ltr]})R + \frac{\lambda_1}{\beta} I_{[K_2<j\leq S_2]} I_{[a_1=ltr]}(R - r(j) - c_t) + \frac{\lambda_2}{\beta} I_{[T_2<j\leq Z_2]} I_{[a_2=Accept]} r(j) - \frac{\lambda_1}{\beta} I_{[a_1=ltr]} c_o - \frac{\lambda_1}{\beta} I_{[a_1=Reject]} c_r \Bigg]$$

The model is a non-linear model as both $X(s,a)$ and $\alpha(a|k)$ are decision variables. $p(s|s',a)$ is the transition probability from state s' to state s under action a.

The above mathematical model is used for both partial information model under incomplete information on D2's status and D1's status. In the following two subsections, these models are explained in order.

4.1.2 PARTIAL INFORMATION MODEL UNDER INCOMPLETE INFORMATION ON D2'S STATUS

The system under consideration consists of two dealers, Dealer 1 (D1) and Dealer 2 (D2), and optimal policy for D1 is determined in the presence of of partial or no information availability on D2's net inventory level. To model the system, two processes are defined. Due to the information incompleteness, an **observation process** is defined that evolves in continuous time that represents the available side of the the process as observed by D1. Considering the real process, an **internal process**[2] can be defined to represent the real dynamics of the system. For these two processes, states are defined as follows.

States of the internal process are the same with the states of the decentralized system of Chapter 3 and consists of status of D1 and D2, where set of all states is $S = \{s : s = (i,j) | i \in I, j \in J\}$. Internal process is represented by $\{X_t, t \geq 0\}$, where X_t denote the state of the internal process at time t.

Let $\{S_1, S_2, \ldots, S_K\}$ be a given partition of the state space, such that $S = S_1 \cup S_2 \cup \ldots \cup S_K$, $S_k \cap S_l = \emptyset$ for $k \neq l$. At time t, it is assumed that the state of the internal process

[2] Real process, core process or underlying process are used as synonyms of internal process in the literature.

X_t cannot be observed, but the subset of S, say S_k, that X_t belongs to is observed. These partitions define the states of the observation process. Observation process is defined as $\{Y_t, t \geq 0\}$, where Y_t denotes the state of the observation process at time t and $Y_t = k$ if and only if $X_t \in S_k$. Please note that internal process $\{X_t, t \geq 0\}$ is a Markov Process, but observation process $\{Y_t, t \geq 0\}$ is not. The reason is that having the information that $Y_t = k$ is not sufficient to describe the next state. The new state may be an element of S_k or may belong to another partition of the internal process' state.

For information availability on D2's status, definition of partition of the state space depends on j. In order to give an idea, consider the following example. As explained in Chapter 1, an authorized service of TOFAŞ can locate a specific part at another dealer with the help of the information system. However, the information system allows the authorized service only to observe whether the part exists or not at the other dealer. The authorized service performing the search observes $j > 0$ or $j \leq 0$. In this case, only two levels of information is available in the observation process. The observation variable Y_t takes values from the observation set K, which is actually the states of the partially observable system, where $K = \{k : k = (i, \widetilde{j}) | i \in \mathbb{Z}, \widetilde{j} \in \{0, 1\}\}$. For each i, state of the internal process s implies state of the observation process k to be either $k = (i, 1)$ or $k = (i, 0)$. For example, whenever D1's and D2's net inventory levels are i and j at time t, the real system (or internal system) is at state $s = (i, j)$. If $j > 0$, D1 observes $Y_t = (i, 1)$ and the state of the observation process is $k = (i, 1)$. Figure 4.1 depicts the relation between set S and set K. It can be seen that several states of full-information model belongs to only one state in partial-information model, depending on j and $\widetilde{j}$.

Consider the case that D1 has no information on D2's status. This situation is a special case of partial-information and is called as **no-information case**. No-information availability for D1 on D2's status results in a single level of information availability in the observation process. The observation set K is $K = \{k : k = (i, \widetilde{j}) | i \in \mathbb{Z}, \widetilde{j} = \{0\}\}$ for no-information case, i.e. there exists a single value for $\widetilde{j}$.

It is previously explained that D2's operating policy is assumed as (S_2, K_2, Z_2, T_2). Table 4.1 presents this operating policy.

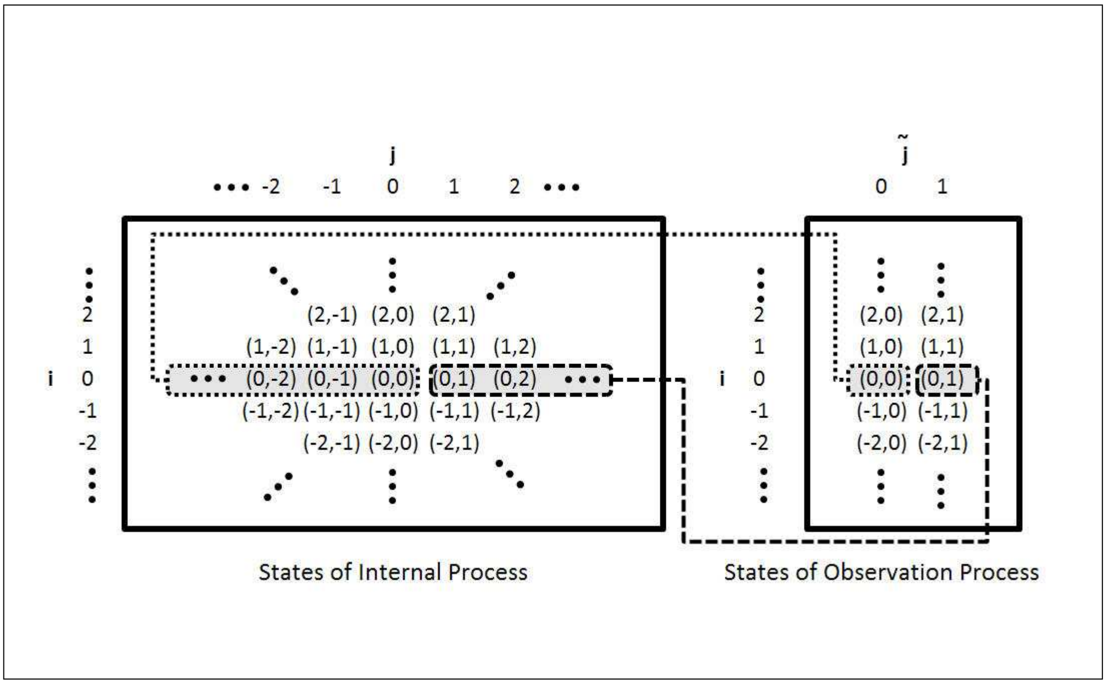


Figure 4.1: States of Full-Information and Partial-Information Models

D1's action space is defined by D2's operation strategy and D2's inventory/queue status. For example, for $Z_2 \geq j > T_2$, D2 asks for lateral transshipment upon a customer arrival and D1 has the options to accept this request or reject it, whereas for $j = T_2$, D2 rejects arriving customers and does not ask for lateral transshipment, and therefore no is available to D1. Table 4.2 presents the available actions of D1.

Under partial information availability, D1 may not have the information on D2's status, but the events occur according to the available actions defined in Table 4.2. For example, upon a customer arrival, D1 may place a lateral transshipment request without having the information on j with respect to K_2. If D2 meets this request, $j > K_2$ is the case, which might not be not known previously by D1. Otherwise, D2 does not meet the request, indicating $K_2 \geq j$.

Considering the above situations, depending on status of D2, D1 makes his decision on a customer arrival or lateral transshipment request based on payment amount being high or low. To reflect this situation, a variable payment amount as a function of j is used, where $r(j)$ is decreasing in j. $r(j)$ is set to be minimum at S_2, and maximum at T_2.

Table 4.1: Policy of D2 for Partial Information Model

	Inventory/Queue Status of D2				
EVENTS	$j \geq S_2$	$S_2 > j > K_2$	$K_2 \geq j > Z_2$	$Z_2 \geq j > T_2$	$T_2 \geq j$
part arrival	reject	accept			
ltr from D1	accept		reject		
customer arrival	accept			ltr	reject

4.1.3 PARTIAL INFORMATION MODELS UNDER INCOMPLETE INFORMATION ON D1'S STATUS

It is not uncommon that the companies have inaccurate information about their inventory levels. An example is about Yıldırım Engineering, a mid-size Turkish crushing and screening plants manufacturer located in OSTİM Organized Industrial Zone, Ankara.[3] While production of a screening plant, two ball bearings and three muffs are required. Company ordered these inventory items because both the items were seen to be out-of-stock over the inventory database. After the arrival of items, by scrutinizing the warehouse, two ball bearings and three muffs were found in the inventory and it is found that they have been in the inventory for more than two years. In Chapter 1, a large local plastics manufacturer's high quantity mismatch ratio compared to the inventory database (20% for a random sample of items) is given as another example. Inaccurate inventory information is a problem of companies independent of their sizes. From this motivation, a system where the existence of inventory is the only available information is considered-without the amount information. Other studies in the literature also use inaccurate inventory information models. For example, Bensoussan et al. [2008] assume partially observed inventory due to transaction errors, misplaced inventories, spoilage or production yield. Although it is not common, problems with the information about number of waiting customers may exist,

[3] The observation on Yıldırım Engineering's inventory system was made during a system improvement project.

Table 4.2: Available Actions of D1 for Partial Information Model

	Inventory/Queue Status of D2			
EVENTS	$S_2 \geq j > K_2$	$K_2 \geq j > Z_2$	$Z_2 \geq j > T_2$	$j = T_2$
part arrival	accept, reject			
ltr from D2	naa*		accept, reject	naa
customer arrival	accept, ltr, reject	accept, reject		

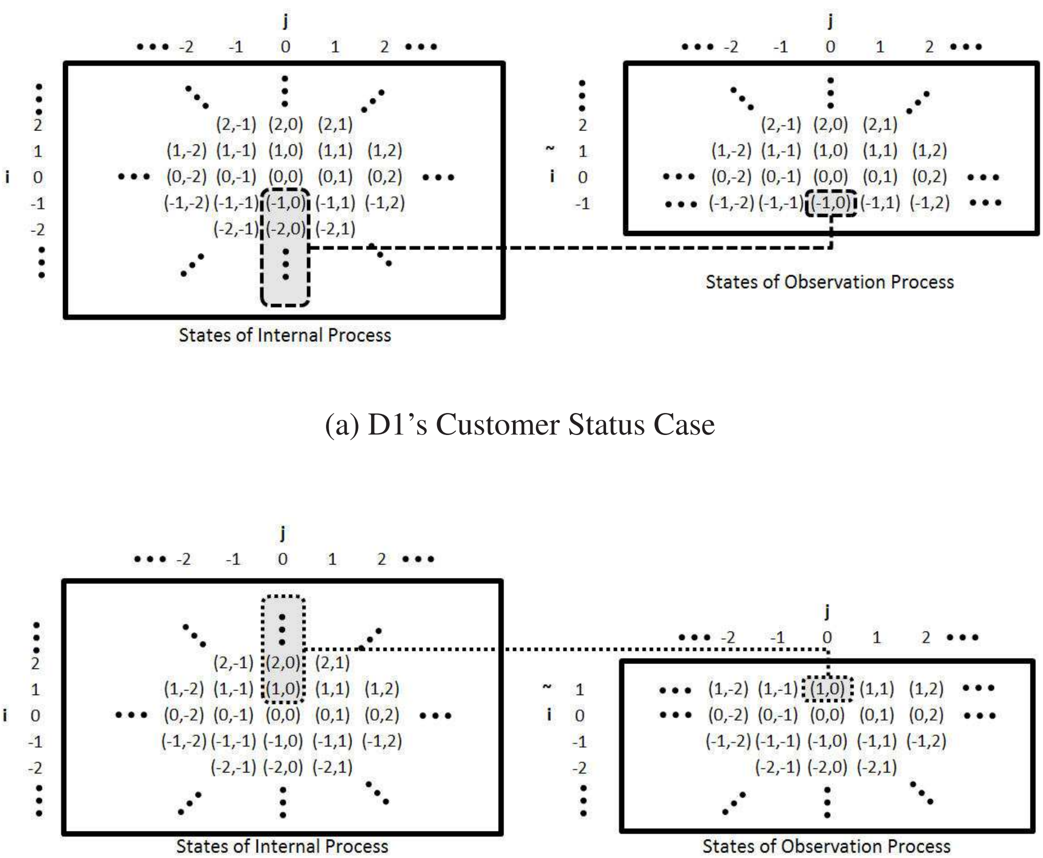
(*naa: no available action)

as well. Lack of recording system of waiting customers or transaction errors may be the reasons for such a case.

Consider the case that a dealer has incomplete information on his inventory level. This situation is called as **D1's inventory status case** in this study. Similarly, **D1's customer status case** is called as the case of incomplete information on the number of customers in the queue.

For D1's inventory status case, D1 just has the information on existence of inventory, but is not able to distinguish between the inventory levels. This implies D1 exactly knows the number of waiting customers in the queue, a stock-out situation or whether there exist items on-hand. Bensoussan et al. [2008] also use a similar model where either zero inventory level is observed for out-of-stock situation or positive inventory is observed for any level of inventory. In D1's customer status case, D1 cannot distinguish between number of waiting customers in the queue, but exactly knows the amount of inventory on-hand. Figure 4.2 shows these cases.

Using the definitions of Subsection 4.1.2, **observation process** is redefined as $\{Z_t, t \geq 0\}$ for D1. For decision epoch n, $Z_n = g$ if and only if $X_n \in S_g$, where $\{S_1, S_2, \ldots, S_G\}$ be a given partition of the state space S and G is defined as follows. For D1's inventory status case, $G = \{g : g = (\tilde{i}, j) | \tilde{i} \in \{-\infty, \ldots, -1, 0, \mathbf{1}\}, j \in \mathbb{Z}\}$ and for D1's



(a) D1's Customer Status Case

(b) D1's Inventory Status Case

Figure 4.2: States of D1's Inventory Status Case and D1's Customer Status Case

customer status case $G = \{g : g = (\widetilde{i}, j) | \widetilde{i} \in \{\mathbf{-1}, 0, 1, \ldots, \infty\}, j \in \mathbb{Z}\}$.

4.2 COMPARISON OF THE OPTIMAL POLICIES UNDER FULL INFORMATION AND PARTIAL INFORMATION MODELS WHEN D2'S STATUS INFORMATION IS INCOMPLETE

Consider the case where D1 has incomplete information on D2's status. D1 is informed of D2's status implicitly with respect to the region that j belongs to, such as $j \leq Z_2$, $Z_2 < j \leq K_2$ or $K_2 < j \leq S_2$. Then, for these three intervals of j, $\widetilde{j} \in \{0, 1, 2\}$ can be defined to indicate those state-partitions, respectively. This situation is called as **partial-information case** in this study and denotes partial-information availability on D2's status. Combining i with $\widetilde{j}$, the state definition for partial information model becomes $k = (i, \widetilde{j})$. Optimal decisions taken under partial-information and no-information models are found using the models proposed in Subsection 4.1.1. From the initial numerical study given in Appendix D, a sample instance is selected for the explaining the policies under different information availabilities and the value of information.

Policies obtained for the sample problem instance are given in Figure 4.3. The parameters of the instance are as follows: $(S_2, K_2, Z_2, T_2) = (4, 0, -1, -4)$, $r = 9$, $R = 10$, $c_h = 2$, $c_l = 2$, $c_r = 0$, $c_t = 0$, $c_d = 2$, $\lambda_1 = 1$, $\lambda_2 = 1$, $\mu_1 = 1$, $\mu_2 = 1$. Figure 4.3(a) shows the policy under full-information. Policy of D1 depends on j ($S_1(j)$ and $T_1(j)$ are defined for all j values, but $K_1(j)$ and $Z_1(j)$ are not defined for all j values due to D2's policy). Non-increasing behavior of policy with respect to j is observed. Figure 4.3(b) shows the policy under partial-information availability. $K_1(\widetilde{j})$ is only defined for $\widetilde{j} = 0$ where D2 makes lateral transshipment requests, and $Z_1(\widetilde{j})$ is only defined for $\widetilde{j} = 2$ where D2 meets D1's requests. Figure 4.3(c) has no horizontal axis as it belongs to no-information case. At $i = -1$, the decision upon a customer arrival is random between rejecting the customer (with probability 0.722) and making a lateral transshipment request to D2 (with probability 0.278).

To describe the differences between the policies under full-information and partial-information, a comparative sketch is made in Figure 4.4. An optimal pooling policy

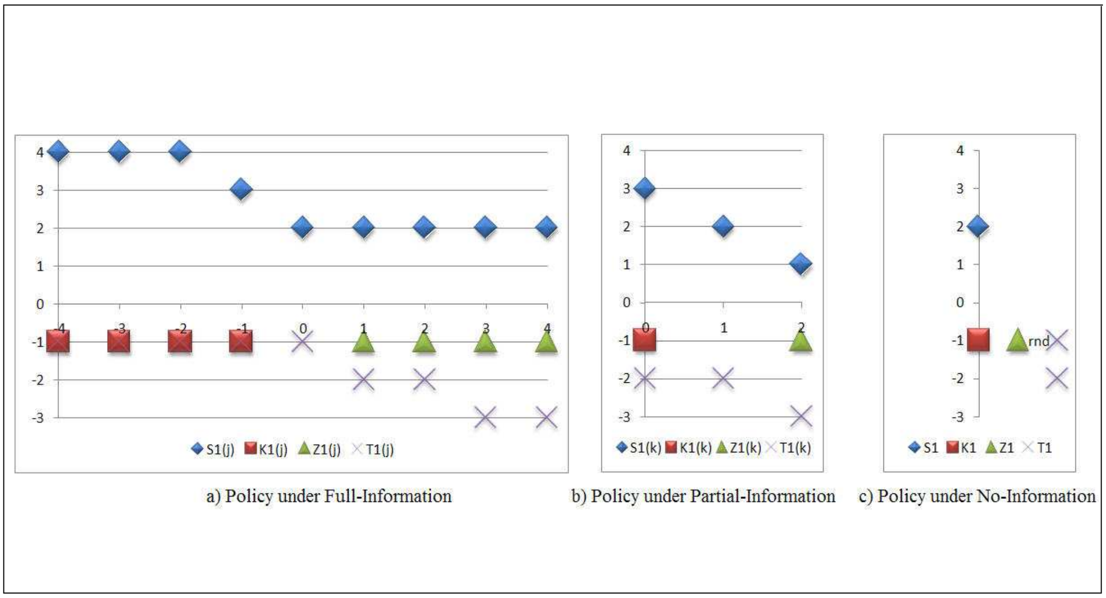
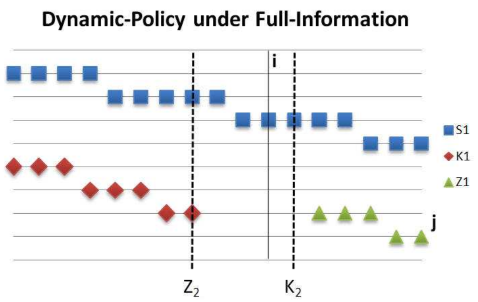


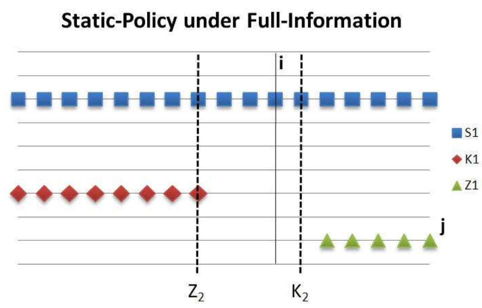
Figure 4.3: Optimal Decisions under Full-information, Partial-information and No-information Availabilities for $(S_2, K_2, Z_2, T_2) = (4, 0, -1, -4)$, $r = 9$, $R = 10$, $c_h = 2$, $c_l = 2$, $c_r = 0$, $c_t = 0$, $c_d = 2$, $\lambda_1 = 1$, $\lambda_2 = 1$, $\mu_1 = 1$, $\mu_2 = 1$

is given in Figure 4.4(a), which is dynamically adjusted with respect to j, and rationing and transshipment levels are defined only for the states where D2's policy is suitable (i.e. $K_1(j)$ is defined for $j \leq Z_2$ and $Z_1(j)$ is defined for $K_2 < j$). Under full-information, a static-policy has constant levels for base-stock, rationing and transshipment levels and the policy is not adjusted according to j (see Figure 4.4(b)). Under partial-information case, information on D2's net inventory is available for D1 as $j \leq Z_2$, $j \in (Z_2, K_2]$ or $K_2 < j$ in Figure 4.4(c) and the policy is defined with base-stock, rationing and transshipment levels accordingly. Under partial-information, randomized decisions may occur and base-stock level is a randomized decision for $j \leq Z_2$ in this example. Under no-information on D2's status, D1 use the same policy for all j values, which may possible be randomized (see Figure 4.4(d)).

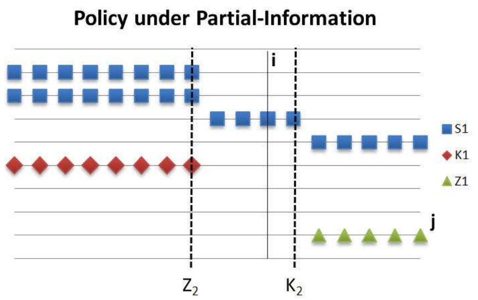
In the following parts of the study, a single information availability level is assumed on D2's status other than full-information, which is no-information availability. No-information implies D1 has no information on j (i.e. $\widetilde{j} = 0$ only), therefore observation set is $K = \{k : k = (i, 0)\}$. TOFAŞ case is an example for an intermediate information availability level between partial-information and no-information availability, where the observation set is $K = \{k : k = (i, \widetilde{j}) | i \in \mathbb{Z}, \widetilde{j} \in \{0, 1\}\}$. Another intermediate information availability level is the case where D1 is informed of D2's



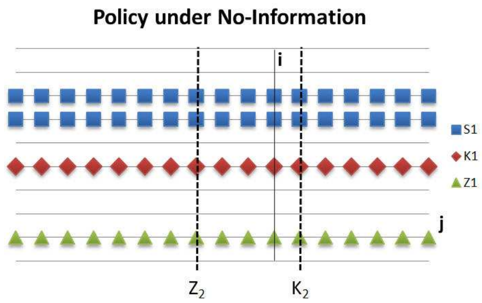
(a) Dynamic-Policy under Full-Information



(b) Static-Policy under Full-Information



(c) Policy under Partial-Information



(d) Policy under No-Information

Figure 4.4: Comparison of Policies under Full-Information, Partial-Information and No-Information

status according to D2's policy, where the observation set becomes $K = \{k : k = (i, \widetilde{j}) | i \in \mathbb{Z}, \widetilde{j} \in \{0, 1, 2\}\}$, as described earlier. All the cases where information is not fully available can be described as partial-information (i.e. TOFAŞ case or D2's policy cases are all examples of partial-information). In addition, no-information case can also be assumed as a special case of partial-information.

In this study, only no-information case is considered as partial-information on D2's status. The reason is problems faced with at the solution phase. As the models used for partial-information are complex, solver performances are low and solution quality is negatively affected. In Subsection 4.3.1, the details of those solution-phase problems are described. For partial-information on D1's status, two different information availability levels are used. Therefore, **no-information case** and **partial-information on D2's status**, and **partial-information case** and **partial-information on D1's status** can be used interchangeably.

4.3 COMPUTATIONAL ANALYSIS

Computational study is performed for two purposes: (i) to assess the benefit of information availability in terms of average expected profit by comparing different information availability cases with full-information availability case, (ii) to analyze how the benefit is affected by the system parameters. Specifically, four different levels of information availability on D1's status and on D2's status is considered, which are as follows. On D2's status, only no-information availability is considered; on D1's status, incomplete information on D1's inventory status and incomplete information on D1's customer status cases are considered; and finally full-information availability is considered. In terms of system parameters, the conditions that are most and least favorable to D1 are identified.

Before starting the details of computational analysis, note that the solution of the mathematical model given in Subsection 4.1.1 might give multiple positive $\alpha(a|k)$ values for an observation process state k, which indicates randomized decision. Conditional probability $\alpha(a|k)$ can be written as $\alpha(a|k) = P\{A_t = a | Y_t = k\}$. It is clear that an internal state $s = (i, j)$ implies observable state $k = (i, \widetilde{j})$ at which decision a is chosen with probability $\alpha(a|k)$.

In the solution phase of the non-linear models, instability problems observed with non-linear solvers. To cope with this situation, a neighborhood search algorithm is developed, which can run with binary variables only. This neighborhood search algorithm used only for incomplete information on D2's status case. In order to have more consistent results, this is required. The payoff by using binary $\alpha(a|k)$ variables instead of continuous ones is observed to be negligible. Details of the usage of binary variables and the effects on quality of results are explained in detail in the following parts of this chapter.

4.3.1 SOLUTION STRATEGY FOR NO INFORMATION MODEL

For the analysis of partial-information on D2's status (i.e. no-information case), solutions for numerical instances are obtained by using a search strategy, which uses a non-linear programming model and a neighborhood search algorithm. Non-linearity of the mathematical model used is one of the strong difficulties faced with in computational analysis. Available non-linear solvers are used with different solution options (time and memory limits options, non-linear solution methods, etc.) and in interaction with each other (using solutions of one solver as starting solution in another solver), but the results were not satisfactory. A search strategy is developed to cope with this difficulty. The number of states and the number of parameters in consideration is the other most significant difficulty, which is partially resolved.

From the initial numerical study (in Appendix D), an important observation about randomized decisions is made. Although randomized decision is allowed in the formulation, randomization is rarely observed in results and $\alpha(a|k)$, the conditional probability of choosing action a given that the observation process is at state k, is observed be either 1 or 0 for almost all instances in the initial runs. Therefore, limiting $\alpha(a|k)$ to assume binary values in the revised model to give deterministic policies (instead of continuous $\alpha(a|k)$ values that give randomized policies) is not considered to have a significant effect on the results. After this decision, the mathematical model proposed becomes a non-linear mixed-integer model.

The mathematical programming model proposed in Subsection 4.1.1 is a non-linear model. In solution phase, the non-linearity brings the risk of ending up with a local optimal solution that might be different than global optimal. In the initial numerical study (see Appendix D), this situation is observed with different local optimal solutions with the non-linear solvers for the problem instances. In order to cope with the problems in the solution of the non-linear programming model, a solution strategy is developed. Figure 4.5 sketches the solution strategy.

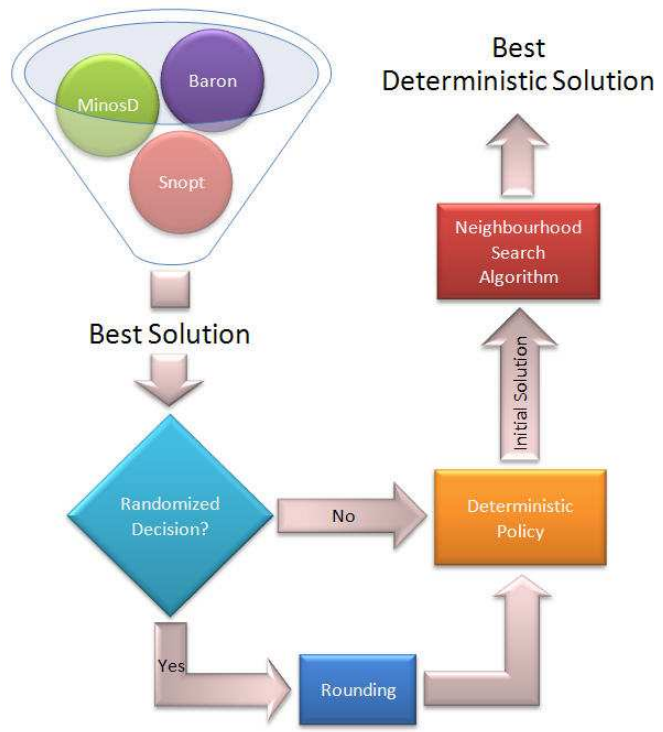


Figure 4.5: Solution Strategy Used with Non-Linear Solvers and Matlab

Basically, solution strategy has three main steps, namely obtaining non-linear solver solutions, rounding the randomized decisions to deterministic ones (if required) and a neighborhood search algorithm to improve the best solution obtained so far. To clarify the solution strategy, the pseudo-code of the strategy is given in Figure 4.6. Three main steps is seen in the the pseudo-code. Each step is explained as follows.

Step 0 Solution strategy starts with linear mathematical programming model solu-

S0. **LP Solutions**

Solve linear programming model for full-information.

S1. **Solver Solutions**

i. Solve non-linear programming models for no-information.
Do not use any initial solution.

ii. Solve non-linear programming models for no-information.
Use S0 solutions as initial solutions.

iii. Solve non-linear programming models for no-information.
Use S1-ii solutions as initial solutions.

iv. Select the solution with the best objective of S1-i, S1-ii and S1-iii and set *solution*.

S2. **Rounding**

If solution of S1-iv gives randomized policy, use rounding, else do not.
Set *deterministic solution* and go to S3.

S3. **Neighborhood Search Algorithm**

Perform neighborhood algorithm using the deterministic solution and set *best deterministic solution*.

Figure 4.6: Pseudo-Code of the Solution Strategy for No-Information Model

tions. Solutions are obtained on a Pentium 3.0 Ghz. processor computer with 2 GB RAM using GAMS Optimization Package Version 22.6. Problem instances are solved for full-information case as explained in Section 3.4, i.e. with a single run for each instance using GAMS with CplexPar as the solver. A similar model to the one given in Section 3.4 is used with the difference of average expected profit as the objective function.

Step 1 Solutions to non-linear mathematical programming model are obtained on a Pentium 3.0 Ghz. processor computer with 2 GB RAM using GAMS Optimization Package Version 23.0. As the models for partial-information case are non-linear, multiple non-linear solvers are used interacting with each other. The non-linear solvers used are Baron, Snopt and MinosD. All those three solvers are used with default options. Baron does not make use of initial point, therefore a single run is made for each instance. Snopt and MinosD are used with no initial solution (*S1.i*), initial solution derived from full-information solution (*S1.ii*) and initial solution provided from each other (*S1.iii*). Figure 4.7 explain the strategy used in non-linear solution. As a result, 12 runs are made for each instance using Snopt and MinosD. Including Baron, 13 runs are made for each instance for the solution. Best solution in terms of profit is selected among those results as the non-linear solution.

The non-linear solvers Snopt and MinosD work as explained in Figure 4.7. These two solvers might provide integer decisions indicating a deterministic policy or probabilities for decisions indicating a randomized policy. Baron is used to solve non-linear mixed-integer model, therefore it gives only binary decision variable values. The solution with the best objective function value is selected among non-linear solver outputs (*S1.iv*) and used in the next steps.

Step 2 The solution of *S1.iv*) is checked for the randomization on policy. In case of solution belongs to Snopt and MinosD solvers, the solution might have a randomized decision. In order to continue to the neighborhood search algorithm, deterministic solution is required and the randomized decisions are rounded to deterministic decisions. If there exists randomization at a particular state for a particular decision, rounding is made and the decision with the highest steady-state probability is set to one, implying that only that decision is certainly made

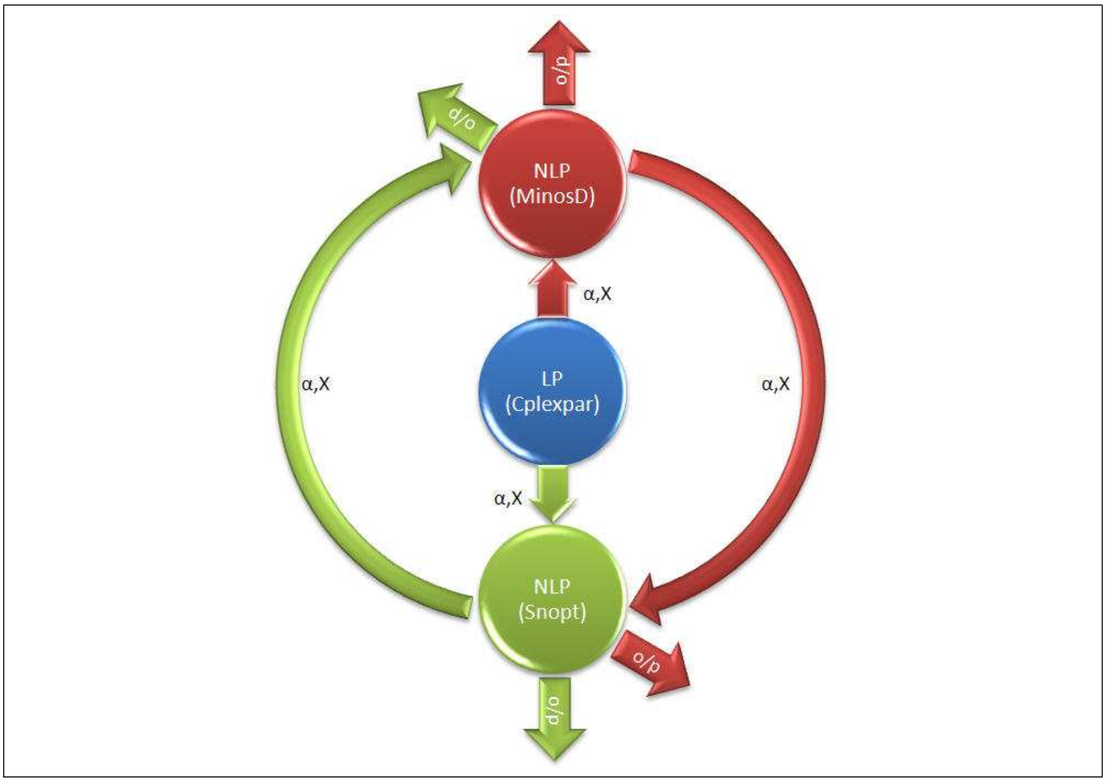


Figure 4.7: Method used for Obtaining Solution to Non-Linear Mathematical Programming Model with Interactive use of Non-Linear Solvers

at that state. For some solutions of Snopt and MinosD solvers where the solution is deterministic and for all solutions of Baron solver (the solution is always deterministic), rounding is not required.

Step 3 The deterministic policy obtained in *S2* is improved through a a neighborhood search algorithm. The details of the algorithm are explained in the following.

Coding of the neighborhood search algorithm is made with Matlab Version 7.5.0.342. Matlab performs the neighborhood search algorithm as follows. At a specific solution of a particular problem instance, D1 has decisions on base-stock, rationing, transshipment and rejection levels. For each level i, $i-1$ and $i+1$ values are considered and all the neighbor solutions to the specific solution are evaluated. This one-step neighborhood search makes at most 3^4 evaluations for each specific solution (since the solutions that do not satisfy $S_1 \geq K_1 \geq Z_1 \geq T_1$ are eliminated). When the improvement in profit is less than 0.1% in relative terms, the algorithm stops and gives the best solution obtained so far.

Considering the overall approach in the numerical study, no-information case is solved as follows. For each instance, 13 runs are made with non-linear solvers and best solution in terms of profit is selected among the results, which forms the initial solution of neighborhood search algorithm. For each instance, Matlab performs at most 81 evaluations, which is the number of possible solutions in one-unit neighborhood of the initial solution. Total number of evaluations depend on the number of iterations that bring improvement for each instance.

Initial solutions provided to the neighborhood search algorithm has an effect on the performance. For the total of 576 instances, the number of initial solutions provided by Snopt or MinosD and Baron are given in Table 4.3. Some of the Baron solutions are improved by the neighborhood search algorithm. 142 out of 432 solutions of Snopt or MinosD are improved, while profit decreased in only five solutions. In these few instances, randomization provides advantage and forcing policy to be deterministic has a negative effect on profit. Such a case is not possible with initial solutions provided by Baron (since they are already binary variables).

Table 4.3: Initial Solutions Provided to Neighborhood Search Algorithm by Non-Linear Solvers

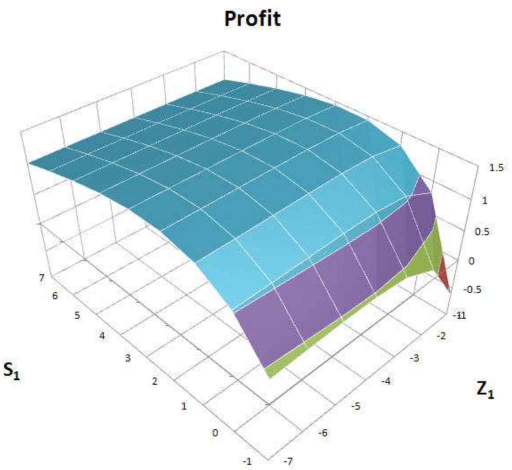
GAMS Solver	# of Initial Solutions	Profit Increased	Profit Decreased
Baron	144	33	0
Snopt or MinosD	432	142	5

From the results of Table 4.3, the following comments can be made. Firstly, Baron does not yield optimal solutions and is stuck at local optima. Neighborhood search algorithm improved many of the solutions provided by Baron. Secondly, many Snopt or MinosD solutions' profits are increased, which is due to local optimality. This improvement is interesting because even rounding is used on randomized decisions, improvements observed. Profit decreased for few cases where randomized decision makes difference, which indicates deterministic policy change proposed has no significant effect on results.

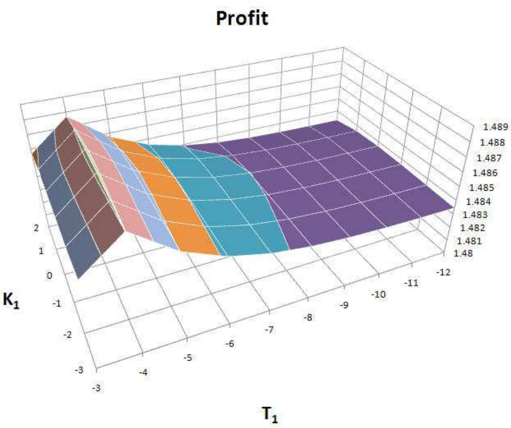
Thirdly, Matlab performed well and the neighborhood algorithm improved some of the solutions, but not deteriorated any of the solutions. A particular problem instance is selected with the following parameters: $(S_2, K_2, Z_2, T_2) = (6, 2, -4, -10)$, $r = 9$, $R = 10$, $c_h = 0.5$, $c_l = 2$, $\lambda_1 = 0.6$, $\lambda_2 = 0.5$, $\mu_1 = 1$, $\mu_2 = 1$, $c_r = 3$, $c_t = 0$, $c_p = 0$, $c_o = 3$. The best policy of D1 obtained for the instance using the search strategy is $(S_1, K_1, Z_1, T_1) = (2, -1, -3, -7)$. Under no-information, a deterministic policy for D1 is a static policy as D1 uses the same policy at all j values. In order to understand the behavior of profit around the best solution, a search in terms of base-stock and transshipment, and rationing and rejection levels is made. The results are given in Figure 4.8. From the best policy obtained using the search strategy, base-stock and transshipment levels are $(S_1, Z_1) = (2, -3)$. This is the maximum point of (a). Around this point, a concave profit function is observed, which indicates the decision on base-stock and transshipment levels are the global optimal values. For (b), rationing and rejection levels are $(K_1, T_1) = (-1, -7)$, but the maximizer is $(K_1, T_1) = (0, -4)$. With changing rationing and rejection levels, profit might improve, but neighborhood search algorithm stops at $(K_1, T_1) = (-1, -7)$ as the neighborhood does not provide necessary amount of improvement. When the scales of profit for (a) and (b) are considered, this can be seen from the figure. The profits at (b) for D1's decision and maximizer, profit at $(K_1, T_1) = (-1, -7)$ is 1.484, whereas it is 1.488 at $(K_1, T_1) = (0, -4)$. For the other combinations of S_1, K_1, Z_1, T_1, the same profit graphs are drawn and a concave shape or close to it is observed. As a conclusion, neighborhood search algorithm makes the search on a concave or alike profit function and performance of the algorithm is satisfactory for the no-information case.

4.3.2 SOLUTION STRATEGY FOR PARTIAL INFORMATION MODEL

For the analysis of partial-information on D1's status (i.e. partial-information case), solutions strategy used for no-information case (defined in Figure 4.5) could not be used. Only the same strategy used in non-linear solution for no-information case (given in Figure 4.7) is used, which is only a part of the solution strategy. The reason is that neighborhood search is not possible for partial-information case due to the large number of states.



(a) S_1 *and* Z_1



(b) K_1 *and* T_1

Figure 4.8: Behavior of Profit versus D1's Policy under No-Information Around Local-Optimal of a Sample Instance

Total number of states for D1's customer status or D1's inventory status is 527, which is large compared to 31 in no-information case. When number of states is large, Baron cannot find a solution in reasonable amount of time. For this large number of states, neighborhood search algorithm would not provide results as the number of neighborhood solutions is $\left(3^4\right)^{17}$ (there are 4 decisions with 3 levels and 17 states for a j). For D1's customer status or D1's inventory status, total of 13,824 runs are made. Considering all information availability levels, total of 21,888 GAMS runs are made.[4]

Since the solution strategy is solely the strategy used in non-linear solution, it has a single main step, namely obtaining non-linear solver solutions. The pseudo-code of the strategy is given in Figure 4.9. The main step is seen in the the pseudo-code. Explanations of each step is similar to those for no-information case, except the solutions for non-linear programming models are obtained for partial-information case and the best solution is set by selecting the solution with the highest objective function value after obtaining the non-linear solutions.

S0. **LP Solutions**

Solve linear programming model for full-information.

S1. **Solver Solutions**

i. Solve non-linear programming models for partial-information.
Do not use any initial solution.

ii. Solve non-linear programming models for partial-information.
Use S0 solutions as initial solutions.

iii. Solve non-linear programming models for partial-information.
Use S1-ii solutions as initial solutions.

iv. Select the solution with the best objective of S1-i, S1-ii and S1-iii and
set *best solution*.

Figure 4.9: Pseudo-Code of the Solution Strategy for Partial-Information Model

[4] 576 instances are considered in the parameter setting. For each instance, 12 runs are made using Snopt and MinosD solvers for D1's customer status and D1's inventory status cases, which makes total of 13,824 runs. For no-information case, 13 runs are made using Snopt, MinosD and Baron solvers, and 1 run is made for full-information CplexPar solver, which makes total of 21,888 runs.

4.3.3 NUMERICAL SETTING OF PARAMETERS FOR THE ANALYSIS

Parameter setting for the numerical study is decided in the light of initial numerical study results. Base-stock and rejection levels of D2 are set as $(S_2, T_2) = (6, -10)$. Revenue amount is fixed as $R = 10$. Inventory holding and customer waiting costs are fixed at $c_h = 0.5$ and $c_l = 2.0$. Arrival rate of D2, λ_2, is fixed as 0.5. Production rates of D1 and D2 are used as $\mu_1 = \mu_2 = 1.0$. Table 4.4 shows the values for the remaining parameters.

Table 4.4: Parameter Values for the Numerical Study for Partial-Information Case

Parameters	r	(K_2, Z_2)	c_r	c_t	c_p	c_o	λ_1
Values	1,5,9,$r(j)$	(3,0),(2,-4),(-3,-5)	0,3	0,3	0,3	0,3	0.2,0.6,1.0

Variable payment amount $r(j)$ is a decreasing function in j is set to be minimum for largest j, i.e. $argmin_j\ r(j) = S_2$, and $r(j)$ is set to be maximum for smallest j, i.e. $argmax_j\ r(j) = T_2$. Explicitly, $r(j) = 9 - \frac{j+10}{2}$. Figure 4.10 shows the structure of $r(j)$.

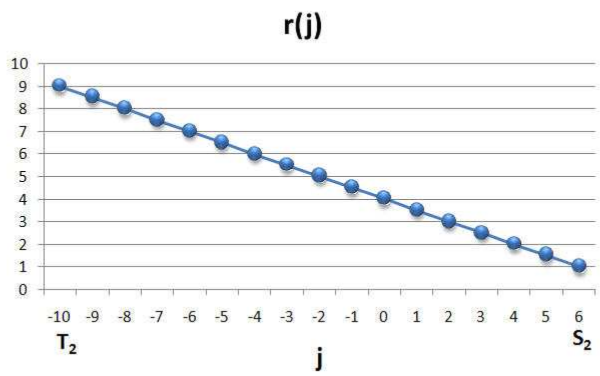


Figure 4.10: Variable Payment Amount $r(j)$ versus D2's Status

Information incompleteness on D2's status is modelled with no-information availability only. As mentioned before, there are several reasons for this. In many instances, no-information profits are close to that of full-information's. Any other information availability level between no-information and full-information (such as the TOFAŞ case) would provide a profit between no-information and full-information profits, which might even be closer to full-information profit. Another reason is the state-space size and run time. Although profits are closer between no-information and full-information cases after neighborhood search algorithm, the difference between profits for some instances that are inputs to the neighborhood search algorithm are very large. This indicates that non-linear solvers would provide local optimal solution which might be improved significantly for some instances. Even using an information availability of TOFAŞ case (this gives the smallest state-space size after no-information, as $k = \{0, 1\}$ only) would make the neighborhood search algorithm impossible to run in reasonable time, because for each iteration, the number of solutions in the neighborhood would squared up compared to no-information case, i.e. 6,561 solutions might be checked for a particular iteration. The number of steps is not only problem for neighborhood search algorithm, but also a problem for the non-linear solver Baron. Baron provides very good results in terms of profit compared to neighborhood search algorithm solution and also gives deterministic decisions. Being unable to use Baron for any partial-information model would bring another important disadvantage. Therefore, a partial-information model is not formed.

Numerical study is performed by using the parameter setting given in Table 4.4 and solution strategy given in Figure 4.5. Total of 576 cases are considered, which is the number of all possible parameter combinations.

4.3.4 ANALYSIS OF PARAMETER EFFECTS ON AVERAGE EXPECTED PROFIT UNDER INCOMPLETE INFORMATION ON D2'S STATUS

In the following subsections of this section, **profit** will refer to the average expected profit of D1. The Percent Benefit of Information Availability (PBIA) is used as the performance measure which is defined as $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{no}}{\overline{\Pi}_{no}} \times 100\%$, where $\overline{\Pi}_{full}$ and $\overline{\Pi}_{no}$ denote the profits of D1 for full-information and no-information availabilities, respectively. Basically, this measure is used for comparisons and called as PBIA in the remaining parts of this chapter. Profit is also used as a measure when required. The performances of the two systems are clearly dependent on their policies, and therefore for the comparison of two systems, there is a need to analyze these policies in order to understand the performance differences.

Referring to Table 4.4, three different D2 policies are under consideration. S_2 and T_2 values are kept constant at 6 and -10, respectively, while (K_2, Z_2) values are varied. For $j \in [S_2, K_2)$, D2 meets D1's lateral transshipment requests. $(K_2, Z_2) = (-3, -5)$ can be called as **high secondary supply** as D2 relatively meets more of D1's requests compared to other (K_2, Z_2) values. Similarly, for $j \in [Z_2, T_2)$, D2 makes lateral transshipment requests, which is relatively more frequent for $(K_2, Z_2) = (3, 0)$ compared to other (K_2, Z_2) values. Therefore, $(K_2, Z_2) = (3, 0)$ can be called as **high secondary demand**. Low secondary demand and high secondary supply can be used as synonyms, and vice-versa. Table 4.5 summarizes these cases.

Table 4.5: Properties of D2's Policies Used in the Numerical Study

(K_2, Z_2)	Property
(3,0)	high secondary demand / low secondary supply
(2,-4)	moderate secondary demand / moderate secondary supply
(-3,-5)	low secondary demand / high secondary supply

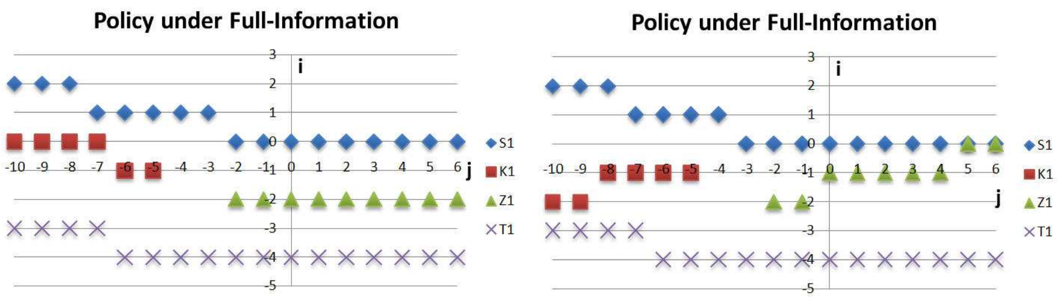
Before starting the comparisons with respect to parameters, two general observations are made. These observations have implications in the explanations of parameter effects.

Observation 4.3.1 *Under full-information availability, D1 has an optimal control policy that can be characterized with the control variables* $(S_1(j)^*, K_1(j)^*, Z_1(j)^*, T_1(j)^*)$ *for each j. Control variables have monotonic behavior in many instances and the monotonic behavior is as follows. For fixed payment amount r,* $S_1(j)^*$, $K_1(j)^*$, $Z_1(j)^*$, *and* $T_1(j)^*$ *are non-increasing in j, however; for variable payment amount r(j),* $S_1(j)^*$, *and* $T_1(j)^*$ *are non-increasing in j, whereas* $K_1(j)^*$, *and* $Z_1(j)^*$ *are non-decreasing in j.*

The behavior of rationing and transshipment levels under variable payment amount yields interesting insights. In the full-information model with fixed payments, D1 takes its best response actions given the operating policy of D2 in a competitive environment. As explained in Subsection 3.3.2 in detail, D1 increases the rationing level with decreasing stock level at D2 in order to meet fewer of D2's requests by increased number of lateral transshipment requests. Similarly, D1 increases the transshipment level with decreasing stock level at D2 in anticipation of a decrease in stock sharing tendency of D2 and increase the number of realized lateral transshipments from D2. Variable payment amount reverses this behavior. As stock level at D2 decreases, $r(j)$ increases. Consider the range for status of D2 as $T_2 \geq j \leq Z_2$. In this range, D2 makes lateral transshipment requests, and with decreasing j, meeting a request becomes more profitable for D1. Therefore, D1 decreases his rationing level with decreasing j in order to meet more of D2's requests at a higher payment amount. For the range $K_2 < j \leq S_2$, D2 meets all the lateral transshipment requests of D1. For increasing j, $r(j)$ decreases and D1 prefers using D2's resources more as they become relatively cheaper. Therefore, D1 increases his transshipment level with increasing j in order to place more lateral transshipment requests at a lower payment amount. Figure 4.11 shows this behavior for a sample instance.

Observation 4.3.2 *Percent benefit of information availability in terms of average expected profit is high for the variable sharing amount $r(j)$.*

Observation 4.3.2 is about variable payment amount. $r(j)$ is high when D2 has waiting customers, and D2 places lateral transshipment requests for arriving customers for those low j values. When D2 has on-hand inventory, $r(j)$ is low and D1 prefers to use D2's resources as they are relatively cheaper. D2 shares his inventory with D1 and meet more of the lateral transshipment requests of D1 for those high j values. These imply that variable payment amount is beneficial for D1 both in high secondary demand and high secondary supply conditions. The benefit highly depends on the information on j, because the decision of D1 is affected by the information on D2's actions and the corresponding $r(j)$ at a specific j. For no-information case, D1 cannot make use of this information and cannot adopt his policy to enjoy the advantage of $r(j)$. On the contrary, full-information case allows D1 to use the information of j,



(a) Policy under Full-Information for $r = 5$

(b) Policy under Full-Information for $r(j)$

Figure 4.11: Comparison of Policies under Full-Information for $r = 5$ versus $r(j)$ for $(K_2, Z_2) = (-3, -5)$, $\lambda_1 = 0.2$, $c_r = 0$, $c_t = 0$, $c_p = 0$ and $c_o = 0$

which increases the PBIA.

In the following parts of this subsection, parameters except the one in consideration are selected in a way that PBIA values are the highest among all parameter combinations. For small values of PBIAs, it is not very meaningful to attempt interpretations of effects due to two reasons. Firstly, small effects are not very interesting to interpret and does not have practical implications for quantifying the value of information. Secondly, the effect of a possible local optimality for the solution of no-information model might result in absurd effects and the real effects might be shadowed or even completely vanished.

4.3.4.1 THE EFFECT OF ARRIVAL RATE TO D1 (λ_1)

PBIA is analyzed under varying λ_1 values and varying secondary demand and supply values (i.e. (K_2, Z_2) values). Observe that under $r(j)$, as λ_1 increases PBIA decreases. The impact of λ_1 is more pronounced under high secondary supply (i.e.

$(K_2, Z_2) = (-3, -5)$). Furthermore, PBIA is always higher under high secondary supply compared to other secondary supply levels (see Figure 4.12). Secondary supply/demand and traffic effects are analyzed as follows.

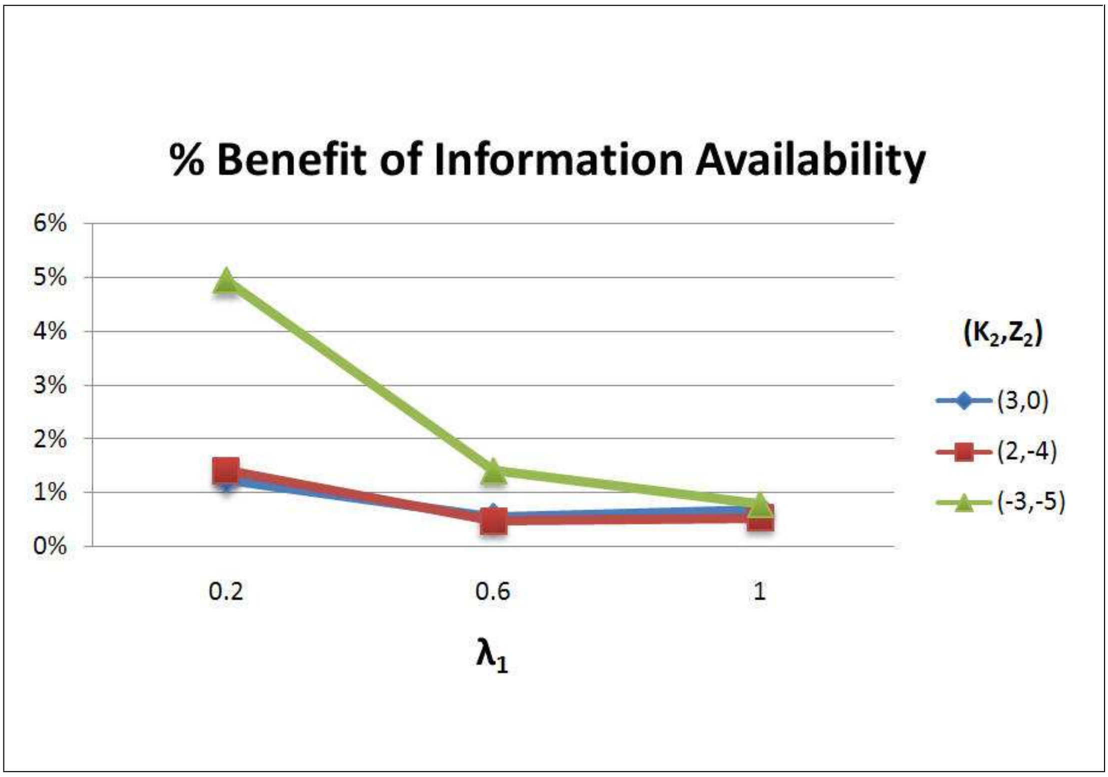


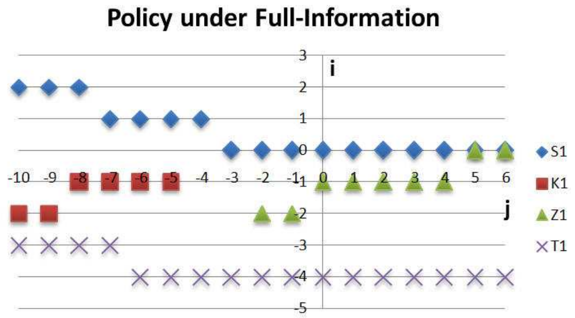
Figure 4.12: Percent Benefit of Information Availability versus λ_1 under D2's Policies for $r(j)$, $c_r = 0$, $c_t = 0$, $c_p = 0$ and $c_o = 0$

Under variable payment and low λ_1, highest PBIA is obtained for high secondary supply, but benefit is positive for moderate and low secondary supply situations also. Consider the low secondary supply ($(K_2, Z_2) = (3, 0)$) situation. The positive PBIA low secondary supply can be attributed to the dynamic policy adjustment under full-information. Under high j values, variable payment is low and since λ_1 is low, D1 can tolerate not keeping inventory and frequently delegates orders to D2. On the other hand, under low j values where variable payment is high, D1 keeps high inventory levels and sets rationing levels to low values to benefit from requests of D2. Under full-information, a majority of the profit is obtained due to delegating the requests to D2 without significant revenue losses, since the system spends most of the time at state $(0, S_2)$ due to low λ_1.

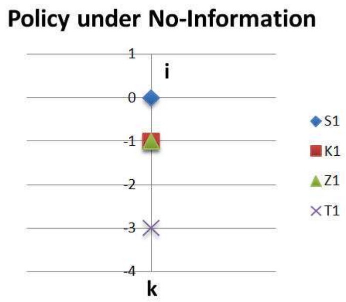
Under no-information, status of D2 cannot be observed, and therefore D1 sets base-stock, rationing and transshipment levels by considering the steady-state probabilities of the internal process. D1 adopts to high secondary demand by adjusting transshipment level to zero while keeping still no inventory, i.e. all the arriving customers are redirected to D2. This behavior is similar under full-information availability for the states that the system spends most of the time. Therefore, the profit is not significantly different between no-information and full-information cases, and information value is not apparent compared to high secondary supply situation, but the profit under no-information availability is lower than the profit under full-information availability and PBIA is positive.

The highest PBIA under variable payment and low customer arrival rate ($\lambda_1 = 0.2$) is obtained for high secondary supply ($(K_2, Z_2) = (-3, -5)$). The effect of dynamic policy adjustment under full-information is also valid at this situation. Under high j values, D1 does not keep inventory and frequently delegates orders to D2, and under low j values, D1 keeps high inventory levels and shares more with D2 by setting rationing levels to low values (see Figure 4.13). The reason that PBIA is larger at high secondary supply is the policy change under no-information. When secondary supply is high, secondary demand is low, which means D2 places lateral transshipment requests less frequently. For low secondary supply, D1 does not hold any inventory, redirects all customers under no-information, and does not meet and requests of D2 by equating base-stock, rationing and transshipment levels at zero. For low secondary demand, D1 still does not hold inventory (as for low secondary supply), but decreases his rationing to benefit from seldomly arriving secondary demand (see Figure 4.13). Compared to full-information, D1 only partly delegates the arriving customers to D2 and furthermore cannot benefit from meeting D2's requests under high variable payments. The payoff of this policy modification is the cost of increasing waiting customers since transshipment level is also decreased as its level cannot exceed rationing level. Although this policy change under no-information is more beneficial to D1 compared to all zero base-stock, rationing and transshipment levels, D1's profit decreases and the PBIA increases.

Traffic has an adverse effect on PBIA. As λ_1 increases, it is observed that the value of information decreases. Note that base-stock levels under both no-information and full-information availabilities are high (see Figure 4.14). The difference between



(a) Policy under Full-Information



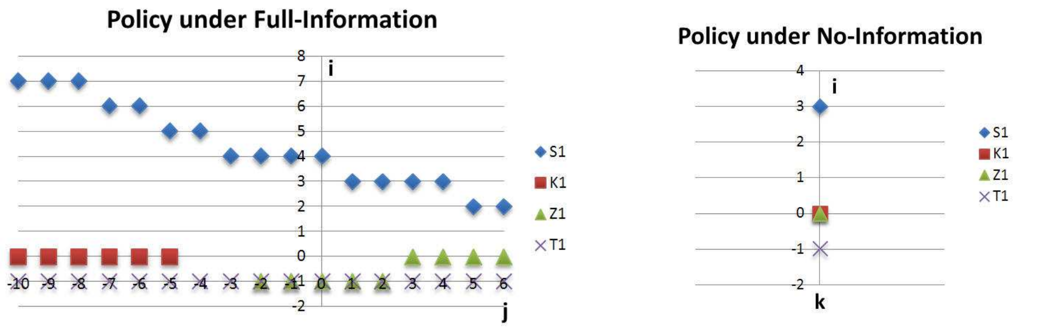
(b) Policy under No-Information

Figure 4.13: Comparison of Policies under Full-Information and No-Information for $(K_2, Z_2) = (-3, -5)$, $\lambda_1 = 0.2$, $r(j)$, $c_r = 0$, $c_t = 0$, $c_p = 0$ and $c_o = 0$

base-stock and transshipment levels are higher compared to low traffic situation under both no-information and full-information. This implies D1 prefers meeting the demand thorough stock rather than delegating it to D2, although D2 provides favorable conditions. In addition, under full-information availability, both $K_1(j)$ and $Z_1(j)$ levels are almost static with changing j levels. Therefore, as the behaviors are similar under both no-information and full-information and dynamic policy adjustment is not aggressive under full-information, PBIA decreases under high traffic.

Under static payment amount, an opposite trend is observed compared to variable payment $r(j)$. PBIA follows an increasing trend with λ_1, which is different than variable payment (see Figure 4.15). Note that PBIA values are smaller in Figure 4.15 compared to Figure 4.12. In Observation 4.3.2, variable payment amount effect on larger PBIA due to usage of information on j is explained.

For $r = 9$, PBIA is positive but not as high as that for variable payment amount. Since the payment is always high, the main reason of positive PBIA is meeting D2's lateral transshipment requests more under full-information compared to no-information. It



(a) Policy under Full-Information (b) Policy under No-Information

Figure 4.14: Comparison of Policies under Full-Information and No-Information for $(K_2, Z_2) = (-3, -5)$, $\lambda_1 = 1.0$, $r(j)$, $c_r = 0$, $c_t = 0$, $c_p = 0$ and $c_o = 0$

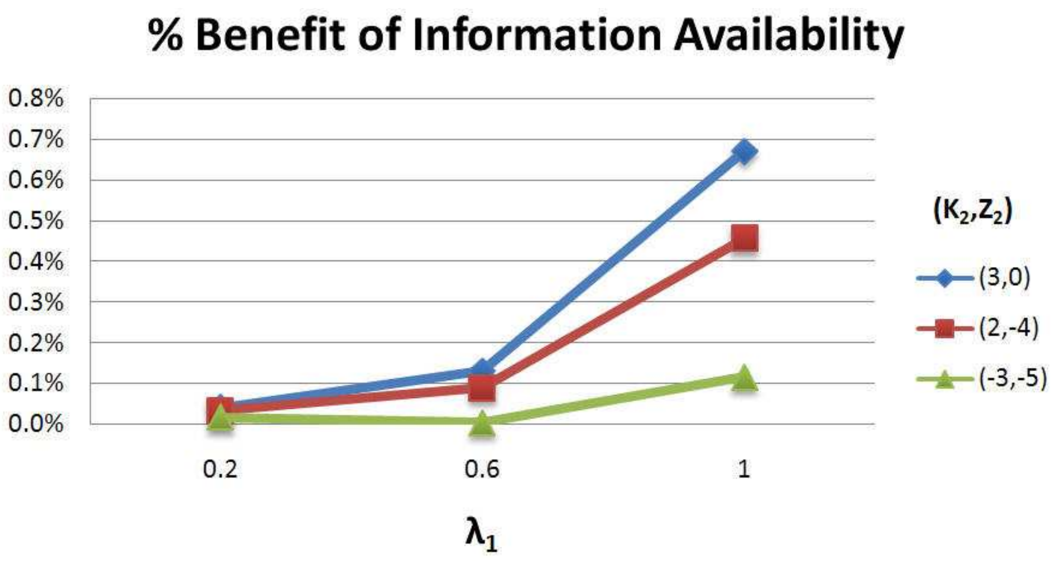


Figure 4.15: Percent Benefit Of Information Availability versus λ_1 under D2's Policies for $r = 9$, $c_r = 0$, $c_t = 0$, $c_p = 0$ and $c_o = 0$

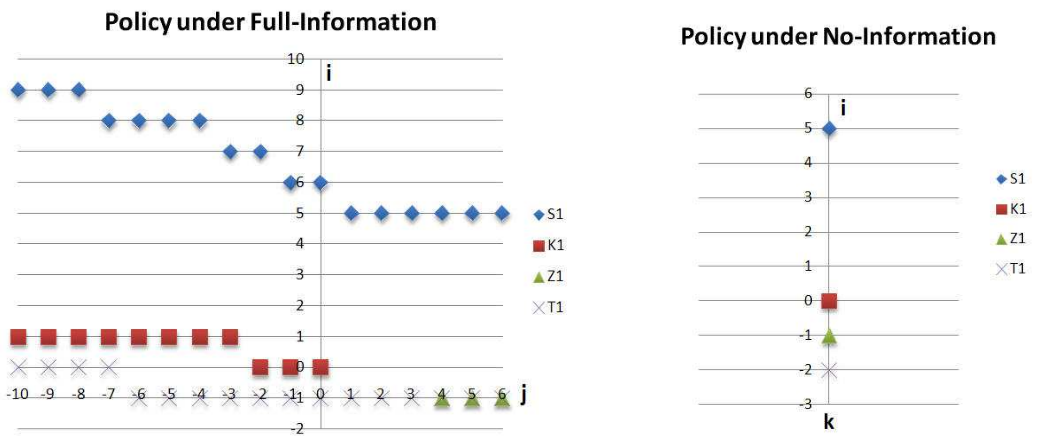
is observed that inventory level under full-information availability is higher than that under no-information availability, rationing levels are almost the same, and as a result, D1 meets requests of D2 more frequently compared to most of the instances, where inventory and rationing levels are closer.

PBIA is larger under high secondary demand when static payment amount is $r = 9$ and becomes largest under high traffic ($\lambda_1 = 1.0$). Under full-information, D1 does not perform lateral transshipment from D2 by equating transshipment and rejection levels, which means D1 does not use D2's supply at all, whereas D1 places lateral transshipment requests when $i < 0$ under no-information (see Figure 4.16). As $r = 9$, using D2's resources is not very preferable. The difference in request behavior of D1 under full-information and no-information availabilities creates a significant effect on profit, as the system spends most of the time at the states where D2 is sharing his inventory with D1. On the other hand, as payment amount is high, meeting more requests of D2 has a positive effect on profit, but high traffic has a negative affect on this tendency. Due to the dynamic adjustment ability of policy under full-information, information becomes more valuable under these conditions and full-information makes advantage over no-information. Please note that under full-information, D1 holds inventory at least equal to no-information and makes the difference between base-stock and rationing levels at least equal to no-information, i.e. D1 shares more with D2 under full-information. PBIA is decreasing with decreasing demand of D2, since the amount of revenue generated by meeting D2's requests decrease under both no-information and full-information availabilities.

For decreasing customer arrival rate of D1, PBIAs are decreasing, and at the low-traffic, they are all very close to zero. Under low arrival rate, value of information is low since D1 meets requests of D2 more frequently under both full-information and no-information availabilities. This causes D1 to use his capacity more for D2 independent of the information availability.

4.3.4.2 THE EFFECT OF ORDERING COST (c_o)

Ordering cost is a fixed cost and paid whenever a dealer places a lateral transshipment request, independent of the request being fulfilled or not. Ordering cost has a



(a) Policy under Full-Information (b) Policy under No-Information

Figure 4.16: Comparison of Policies under Full-Information and No-Information for $(K_2, Z_2) = (3, 0)$, $\lambda_1 = 1.0$, $r = 9$, $c_r = 0$, $c_t = 0$, $c_p = 0$ and $c_o = 0$

significant effect on PBIA when production costs exists, traffic intensity and payment amount are low.

PBIA sharply increases under ordering cost and the reason is the ability of D1 to visualize D2's status and therefore place lateral transshipment request whenever it is fulfilled under full-information. Under no-information, D1 does not have the information of a lateral transshipment request to be fulfilled as j is not visible, therefore due to unmet requests, excess ordering cost is paid. PBIA increases with decreasing supply of D2, because unmet lateral transshipment requests occur more frequently under no-information. In Figure 4.17, increase in PBIA under ordering cost and order in PBIA with respect to secondary supply is given.

The policies of D1 under full-information and no-information support the above explanations. D1 uses static base-stock, transshipment and rejection levels under full-information for all secondary supply and ordering cost values, which are also all equal. Specifically, $S_1(j) = 0,\ Z_1(j) = 0,\ T_1(j) = -3,\ \forall\ j$. For rationing level, D1 uses dynamic adjustment rarely and sets $K_1(j)$ to either one or zero, i.e. $K_1(j) = 0\ or\ 1,\ \forall\ j$. Under no-information, policies for all instances are almost the same, which is $S_1 = 0,\ K_1 = 0,\ Z_1 = 0,\ T_1 = -3$. For some instances, transshipment level and/or rejection level is changed by one unit. These observations on policies show that policies are more or less the same for all instances under full-information and no-information. When there is no ordering cost, the profits of corresponding (K_2, Z_2) values under full-information and no-information are almost the same, because of the similarity of the policies. In existence of ordering cost, policies under full-information and no-information for corresponding (K_2, Z_2) values are again almost the same, but not the profits. Cost due to unmet orders under no-information is the cause of both high PBIA and the increasing order of PBIA with decreasing secondary supply when ordering cost exists.

Consider extreme secondary supply cases, where D2 supplies nothing or shares all his resources. Under both these extreme situations, it is expected that the benefit of information would be low. When there is no secondary supply, D1 will not place any lateral transshipment requests under no-information and there will be no unmet orders and excess ordering cost. When secondary supply is available at all times, D1 will not

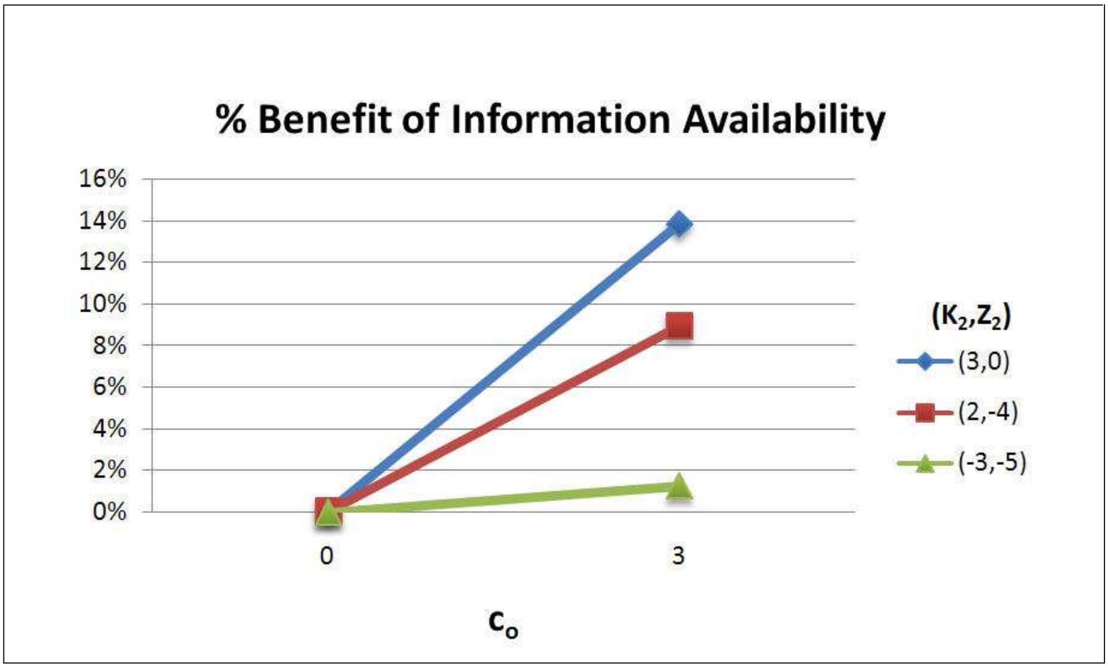


Figure 4.17: Percent Benefit of Information Availability versus c_o under D2's Policies for $\lambda_1 = 0.2$, $r = 1$, $c_r = 0$, $c_t = 0$ and $c_p = 3$

pay ordering cost for unmet orders, as there would be no unmet orders. Therefore, under no or full secondary supply availability, there will be no excess ordering cost and the benefit of information is expected to be low.

4.3.4.3 THE EFFECT OF CUSTOMER REJECTION COST (c_r)

On average, existence of customer rejection cost slightly increases PBIA, however when other costs are zero (such as c_t, c_p and c_o), it has no effect on PBIA (see Figures 4.18 and 4.19). Under variable payment and low customer arrival rate to D1, PBIA does not change with customer rejection cost, but high secondary supply causes higher PBIA (see Figure 4.18). Under $r(j)$ and low λ_1, the policies under no-information are very similar under all secondary supply levels and they are exactly the same under $c_r = 0$ and $c_r = 3$ for a given (K_2, Z_2) values ($(S_1, K_1, Z_1, T_1) = (0, -1, -1, -3)$ for high secondary supply and $(S_1, K_1, Z_1, T_1) = (0, 0, 0, -3)$ for the low and moderate secondary supply). Under full-information, $S_1(j), K_1(j), Z_1(j)$ are the same under $c_r = 0$ and $c_r = 3$ for corresponding (K_2, Z_2) values, except for very few j's, and transshipment level decreases slightly under customer rejection cost. Since there is c_r, customers are rejected more reluctantly and $T_1(j)$ levels decrease, however the decrease is not significant. Under low traffic, customer rejection occurs very rarely and the change in $T_1(j)$ does not cause a significant change in profit under full-information. For high levels of j, base-stock, rationing, transshipment and rejection levels of D1 under full-information and no-information are very similar. For low levels of j under full-information, D2 places lateral transshipment requests more frequently and D1 meets more of these by increasing base-stock level and decreasing rationing level, compared to higher levels of j. Considering that the system spends most of the time for high levels of j, the relative difference is almost the same when the profits are compared.

Figure 4.18: Percent Benefit of Information Availability versus c_r under D2's Policies for $\lambda_1 = 0.2$, $r(j)$, $c_t = 0$, $c_p = 0$ and $c_o = 0$

High secondary supply has a larger PBIA compared to moderate secondary supply and low secondary supply cases, where the latter two actually overlap. The reason is the change in the policy of D1 under no-information. D1 places fewer lateral transshipment requests by decreasing his transshipment level. Transshipment level is lower under high secondary supply ($Z_1 = -1$) compared to low and moderate secondary supplies ($Z_1 = 0$). Considering $S_1 = 0$ and $T_1 = -3$ for all related instances, this one-unit reduction in transshipment level results in a decrease in profit under no-information (profit under low secondary supply with $Z_1 = 0$ decreases almost by 4% for high secondary supply with $Z_1 = -1$). For corresponding (K_2, Z_2) values under $c_r = 0$ and $c_r = 3$, D1 does not change his policy under full-information. Especially for the states where the system spends most of the time, policies are almost the same. The profits for corresponding (K_2, Z_2) values under full-information are very close to each other, and also the profits for (K_2, Z_2) values are close. The profits under $c_r = 0$ and $c_r = 3$ for corresponding (K_2, Z_2) values under no-information are also very close to each other, but the change in the policy of D1 causes lower profits under high secondary supply, compared to moderate and low secondary supplies.

Effect of c_r on the value of information increases with production and ordering costs (see Figure 4.19). Under high traffic and low payment level, production is costly and payment is low, therefore the decision on rationing is effective. By equating base-stock and rationing levels for most of the j values, D1 holds inventory for himself only. For high j values, base-stock under full-information is lower than the base-stock under no-information, while it is higher than the base-stock level under no-information for low j values.

Under production and ordering costs, PBIA follows an increasing trend with rejection cost (see Figure 4.19). Comparing no-information and full-information policies under $c_r = 0$ and $c_r = 3$ for corresponding (K_2, Z_2) values, D1 uses the same base-stock, rationing and rejection levels under no-information, whereas base-stock and rationing levels increase and rejection level decreases when customer rejection cost exists under full-information. By decreasing rejection level and increasing base-stock and rationing levels for $c_r = 3$, D1 meets more demand and rejects fewer customer under full-information availability. Figure 4.20 compares the two policies under $c_r = 0$ and $c_r = 3$ under high secondary demand. As D1 rejects less customers under full-information when customer rejection cost exists compared to $c_r = 0$, and D1 uses the same policy under no-information independent of customer rejection cost (base-stock, rationing and rejection levels are static and does not change with customer rejection cost), PBIA increases with c_r.

Another observation in Figure 4.19 is that PBIA increases with decreasing secondary supply. Due to low payment, D1 does not change transshipment levels, where transshipment level is zero for all related six instances, i.e. $Z_1 = 0$ and $Z_1(j) = 0\ \forall\ j$. It should be noted that although transshipment levels are the same, unmet transshipment requests might be placed under no-info system causing excessive ordering cost, while D1 avoids this under full-information as D1 can observe D2's net inventory level. This is also the cause that PBIAs are ordered increasingly with decreasing secondary supply. As secondary supply decreases, the information on D2's status become more important and number of unmet request increase under no-information which causes more unnecessary ordering costs. Such type of behaviors are explained in Subsection 4.3.4.2 in detail.

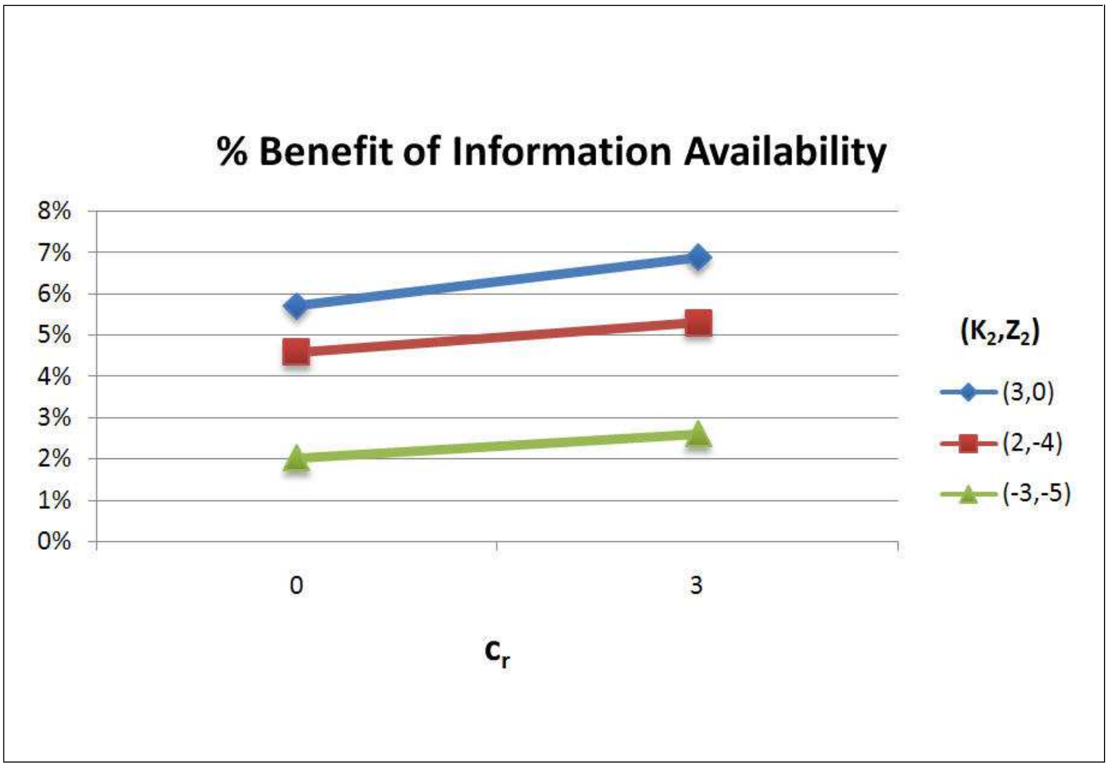
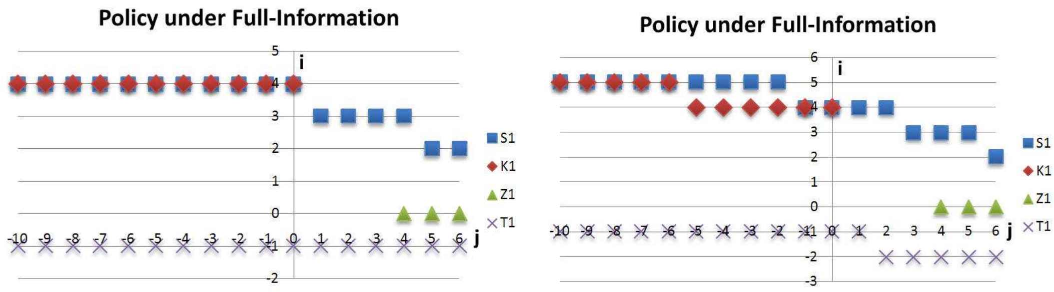


Figure 4.19: Percent Benefit of Information Availability versus c_r under D2's Policies for $\lambda_1 = 1.0$, $r = 1$, $c_t = 0$, $c_p = 3$ and $c_o = 3$

4.3.4.4 THE EFFECT OF PRODUCTION COST (c_p)

Under high traffic and low payment, secondary supply (from D2) becomes important. PBIAs are increasing with decreasing supply of D2 and highest PBIA is observed under lowest secondary supply (see Figure 4.21). The reason that PBIAs are ordered increasingly with decreasing secondary supply is due to the excess cost of unmet



(a) Policy under Full-Information for $c_r = 0$

(b) Policy under Full-Information for $c_r = 3$

Figure 4.20: Comparison of Policies under Full-Information for $c_r = 0$ versus $c_r = 3$ for $(K_2, Z_2) = (3, 0)$, $\lambda_1 = 1.0$, $r = 1$, $c_t = 0$, $c_p = 3$ and $c_o = 3$

orders under no-information, as explained in Subsection 4.3.4.2 in detail.

From Figure 4.21, it is also observed that existence of production cost increases PBIA (this behavior is also valid when ordering cost is zero). Basically, D1 adopts to production cost by decreasing base-stock (therefore producing less) and sharing less (by increasing rationing level) under both full-information and no-information, but the decrease in base-stock is sharper under full-information. Decreasing base-stock requires either using D2's resources via lateral transshipment more frequently (by redirecting more customers to D2) or rejecting customers more frequently (by increasing rejection level). Policies under full-information and no-information are analyzed in order to understand these explanations about increasing behavior of PBIA with production cost. When production cost exists, lower profit is gained per customer. This causes D1 more likely to reject and less likely to keep inventory. This can be observed in Figure 4.22 by comparing base-stock levels of (a) with (c) and (b) with (d).

When $c_p = 3$, D1 never shares his inventory with D2 under full-information availability (see Figure 4.22(c)) by equating base-stock and rationing levels. On the other

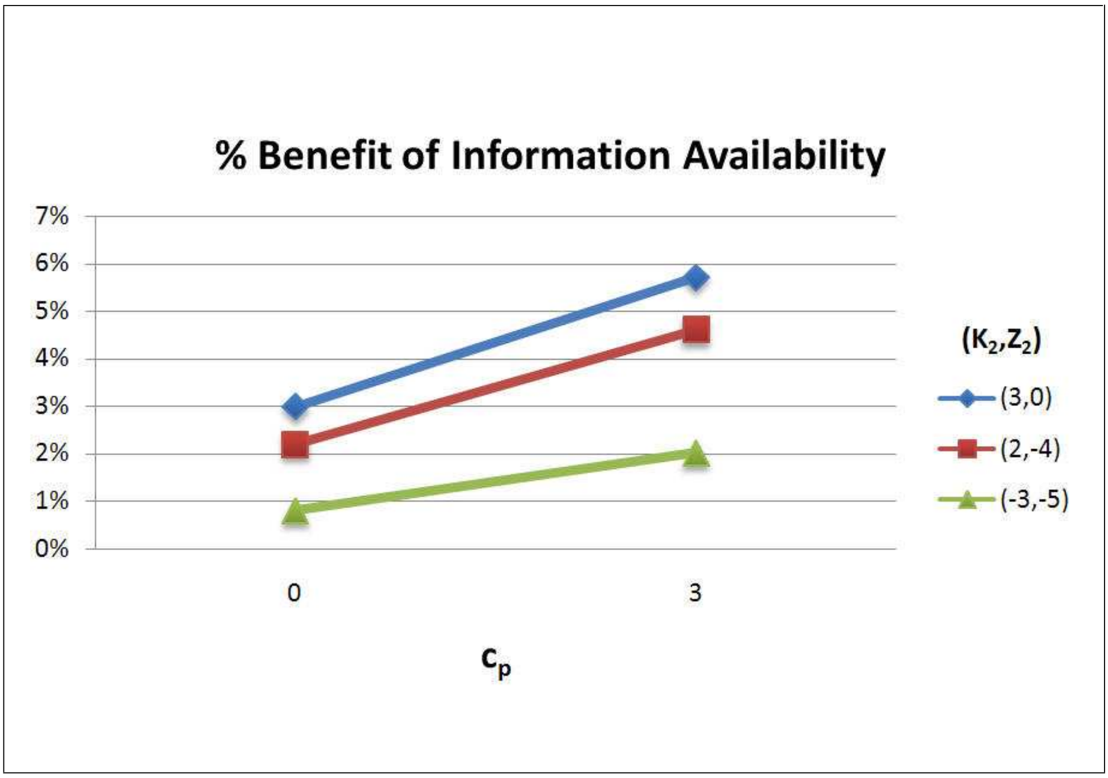
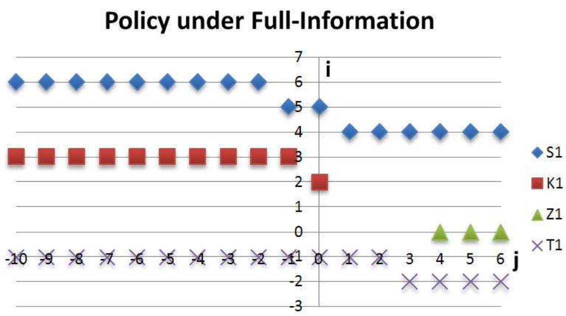


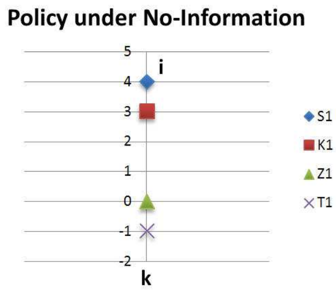
Figure 4.21: Percent Benefit of Information Availability versus c_p under D2's Policies for $\lambda_1 = 1.0$, $r = 1$, $c_r = 0$, $c_t = 0$ and $c_o = 3$

hand, under no-information, D1 meets requests of D2 at a certain level (see Figure 4.22(d)). Considering that $i = S_1$ is the most-visited state for no-information case, this makes an important effect on PBIA. For no production cost, D1 produces more compared to production cost, but again meets requests of D2 at a certain level under no-information (see Figure 4.22(b)). D1 produces more and shares more of his inventory with D2 under full-information (see Figure 4.22(a)) compared to production cost existence. Although payment is low and traffic is high, D1 finds profitable to meet lateral transshipment requests of D2 under full-information when production cost is zero.

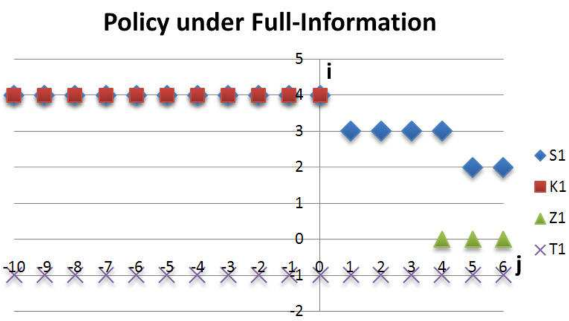
In Figure 4.22, the anticipation to decreasing base-stock by rejecting more customers is seen (but not using D2's resources more). When $c_p = 3$, D1 reject customers at the same levels under both full-information and no-information (see Figure 4.22(c) and (d)). Also for no production cost under no-information, rejection level is the same (see Figure 4.22(b)). For no production cost under full-information, D1 decreases rejection level and rejects less customers, and produces more frequently to meet the higher rate of customers (see Figure 4.22(a)).



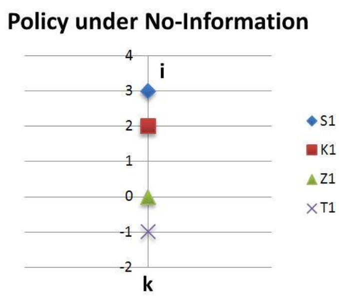
(a) Policy under Full-Information for $c_p = 0$



(b) Policy under No-Information for $c_p = 0$



(c) Policy under Full-Information for $c_p = 3$



(d) Policy under No-Information for $c_p = 3$

Figure 4.22: Comparison of Policies under Full-Information and No-Information for $(K_2, Z_2) = (3, 0)$, $\lambda_1 = 1.0$, $r = 1$, $c_r = 0$, $c_t = 0$ and $c_o = 3$

4.3.4.5 THE EFFECT OF TRANSSHIPMENT COST (c_t)

For overall average, with the increasing transshipment cost, PBIA decreases. This behavior of PBIA for transshipment cost is different compared to other costs effects (c_r, c_p and c_o). This behavior is observed almost under all parameters and Figure 4.23 shows this behavior for high traffic intensity, low payment and existence of production cost. This interesting behavior can be explained by considering extreme transshipment cost situation. When transshipment cost approaches to infinity, D1 will not request any lateral transshipment from D2 and the system converges to no-pooling system. Under a no-pooling-like system, value of benefit will be lower, and therefore PBIA will be lower. For a specific instance, the behavior of PBIA can be explained better by scrutinizing the policies under full-information and no-information cases.

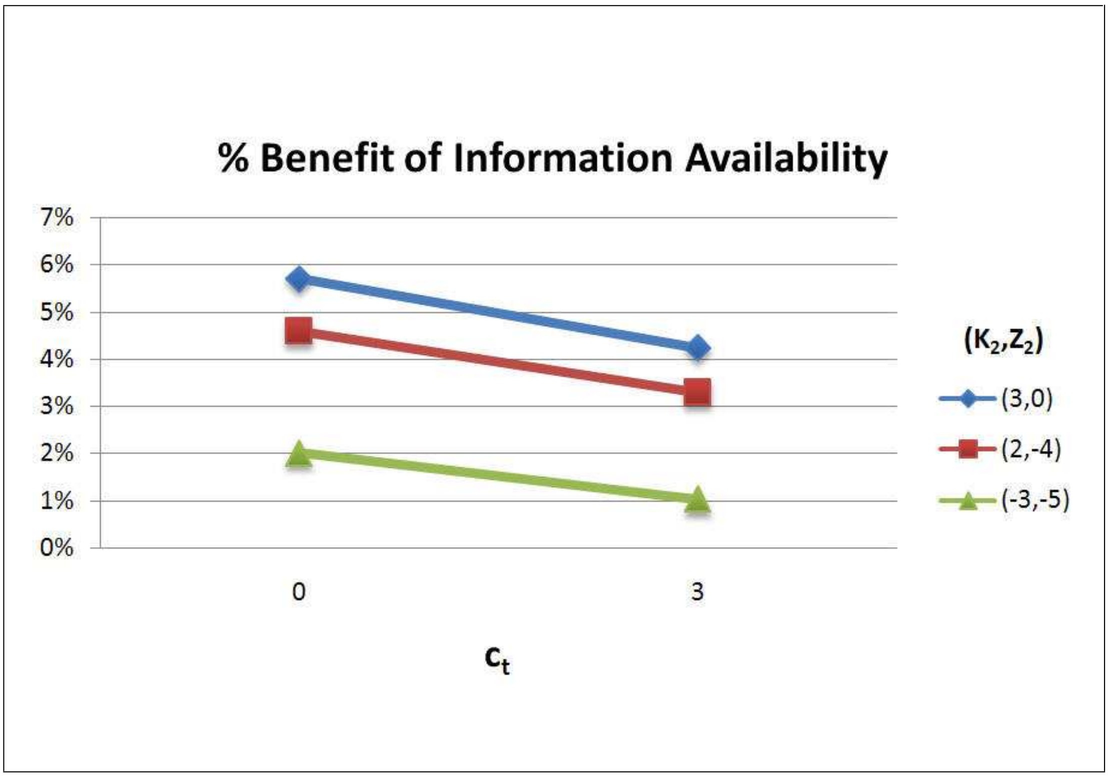


Figure 4.23: Percent Benefit of Information Availability versus c_t under D2's Policies for $\lambda_1 = 1.0$, $r = 1$, $c_r = 0$, $c_p = 3$ and $c_o = 3$

Under full-information, when transshipment cost increases, D1 does not change his rationing, transshipment and rejection levels (or changes very slightly), but increases base-stock levels. By decreasing base-stock levels for no transshipment cost, D1 pays less holding cost, and at the same time, starts using more lateral transshipments, which is profitable as there is no transshipment cost. Under no-information, a similar policy modification is made, but information availability effect on profit via dynamic policy adjustment under full-information is more fine tuned with respect to no-information adjustment, which increases PBIA. The following paragraph explains this situation under a specific set of parameters.

In Figure 4.23, observe that when secondary supply is the lowest ($(K_2, Z_2) = (3, 0)$), PBIA is the highest independent of the transshipment cost. D1 uses the same transshipment and rejection under full-information and no-information, specifically $Z_1(j) = Z_1 = 0$, $T_1(j) = T_1 = -1$, $\forall\, j$. Under full-information, D1 uses his inventory for himself only for both $c_t = 0$ and $c_t = 3$. D1 makes dynamic adjustment on base-stock levels and at the same time, inventory levels under $c_t = 0$ are lower or equal than the levels under $c_t = 3$. Figure 4.24 shows base-stock levels of D1 for low secondary supply situation under both transshipment costs. For $K_2 < j$, where lateral transshipment from D2 is possible, this base-stock level adjustment under low secondary supply situation does not only bring holding-cost advantage to D1 but also using secondary supply via lateral transshipments and as $c_t = 0$, using lateral transshipment does not deteriorate profit. Under no-information, $(S_1, K_1) = (3, 2)$ for $c_t = 0$ and $(S_1, K_1) = (4, 3)$ for $c_t = 3$. The base-stock and rationing level adjustments under no-information is not as effective as under full-information due to the dynamic adjustment availability.

The other observation from Figure 4.23 is that PBIA decreases by increasing secondary supply level. The explanation is inline with the explanations in Subsection 4.3.4.2, i.e. due to the excess ordering costs under no-information. When profits are considered, interesting observations by comparing information and secondary supply availabilities are made. Profits for full-information and no-information cases with respect to transshipment cost are given in Figure 4.25. The highest PBIA is for low secondary supply situation. From the figure, profit for low secondary supply under full-information is exceeded by high secondary supply under no-information. Therefore, it is concluded that D1 would prefer a situation where information is unavailable but secondary supply availability is higher to a situation where information is available but secondary supply availability is lower under high traffic. When the secondary supply is highest, the difference between the profits of full-information

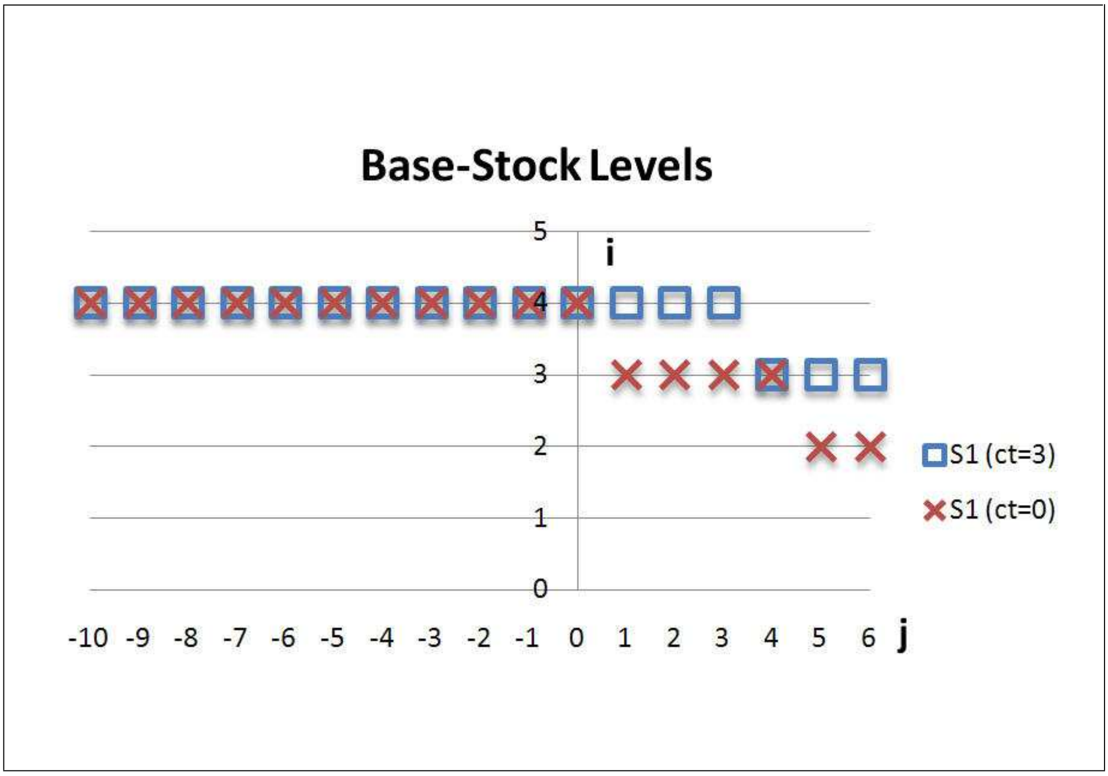


Figure 4.24: $S_1(j)$ for Full-Information under c_t for $(K_2, Z_2) = (3, 0)$, $\lambda_1 = 1.0$, $r = 1$, $c_r = 0$, $c_p = 3$ and $c_o = 3$

and no-information gets smaller, because dynamic adjustment under full-information does not bring much advantage. With decreasing secondary supply, the advantage of dynamic policy adjustment under full-information is more important and profit improves more compared to no-information.

4.3.5 ANALYSIS OF PARAMETER EFFECTS ON AVERAGE EXPECTED PROFIT UNDER INCOMPLETE INFORMATION ON D1'S STATUS

In Subsection 4.1.3, D1's inventory status and D1's customer status models to represent partial information on D1's status are given. Solution strategy used for these cases is explained in Subsection 4.3.1. Because of two reasons, the results of these cases cannot be analyzed in detail.

First reason is the results of non-linear solver solutions. To make a comparison with no-information results, average of PBIA with non-linear solutions of GAMS solvers is considered. Percent Benefit of Information Availability (PBIA) is redefined as

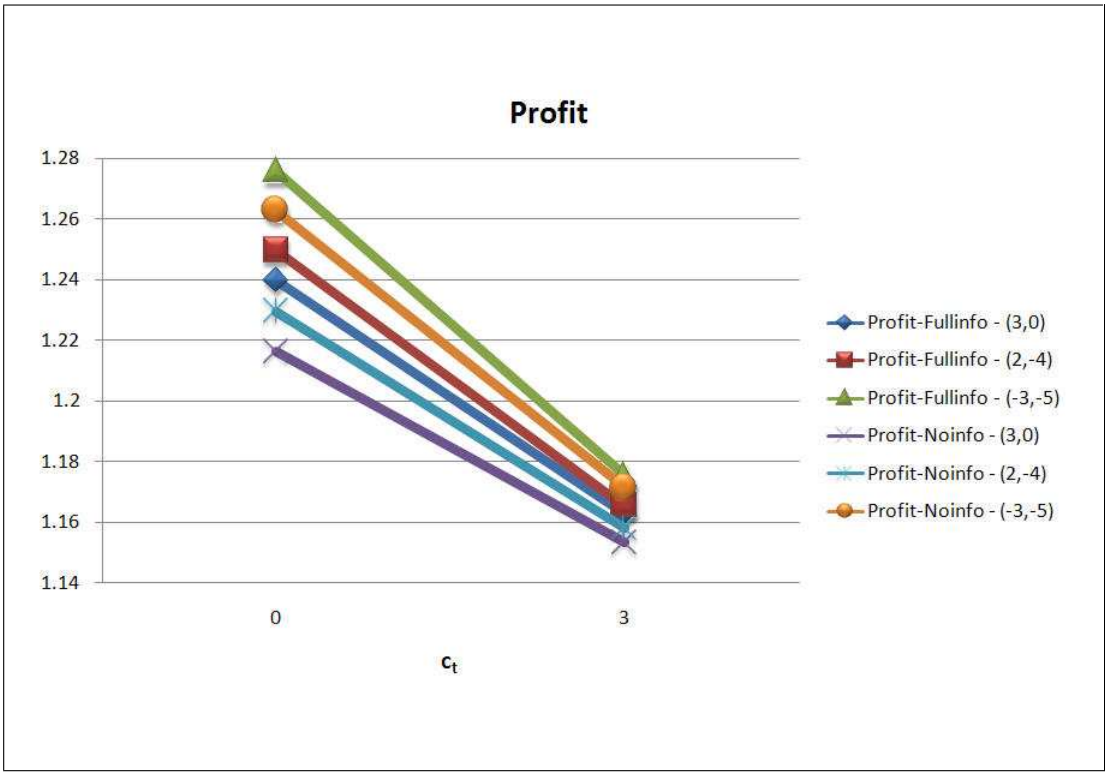


Figure 4.25: Profit versus c_t for Full-Information and No-Information under D2's Policies for $\lambda_1 = 1.0$, $r = 1$, $c_r = 0$, $c_p = 3$ and $c_o = 3$

$\frac{\overline{\Pi}_{full}-\overline{\Pi}_{partial}}{\overline{\Pi}_{partial}}\times 100\%$, where $\overline{\Pi}_{full}$ and $\overline{\Pi}_{partial}$ denote the profits of D1 for full-information availability and D1's customer (or inventory) status case, respectively. The PBIA is 5.78% excluding Baron solutions, as Baron is not available for D1's customer status and D1's inventory status models. PBIA is 0.83% and 2.56% for D1's customer status and D1's inventory status cases, respectively. With these low average PBIA values, meaningful results are not expected for quantifying the value of information on D1's status.

The other reason is the unavailability of neighborhood search algorithm. As explained in Subsection 4.3.3, due to the large number of states, it is not practically possible to code and run a neighborhood search algorithm. Average PBIA value for no-information case is reduced to 1.70%, which is a sharp reduction. It is obvious that without the neighborhood search algorithm, local optimality problem of non-linear solutions would not allow to make meaningful interpretations. Some general observations are as follows.

For D1's customer status case, 531 instances have PBIA less than or equal to 1.5% out of 576 instances. Only 14 instances have PBIA greater than 4%. For 3 instances, extreme PBIA values are observed, where the maximum is 77.7%. Those extreme PBIA values are attributed to stacking of non-linear solvers at local optimums. Excluding those 3 instances, average PBIA decreases to 0.55% from 0.83%.

For D1's inventory status case, maximum PBIA is limited to 11.6%, but average PBIA is larger than that of D1's customer status case. 275 instances have PBIA less than or equal to 1.5% and 165 instances have PBIA greater than 4% out of 576 instances. Only 16 instances have PBIA greater than 8%.

From the observations, average PBIA values are larger for D1's inventory status case compared to D1's customer status case. From the results of full-information case, the system is observed to spend more time at the states where i is positive, i.e. information on inventory is important. As D1's inventory status case makes D1 unaware of inventory status, this information unavailability largens the PBIA compared to D1's customer status case.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

In this chapter, work made in this study is summarized and conclusions are made first. Then, possible future research directions are given.

5.1 SUMMARY OF THE DISSERTATION AND CONCLUSIONS

In this study, analysis of inventory and service pooling, and information sharing in spare parts management systems is performed. There are two main parts of the study, viz. benefits of resource pooling under full information and benefits of information sharing on resource pooling.

Motivation of the study comes from increasing attractiveness of decentralized spare parts management systems. It is very costly to operate the whole spare parts service system, therefore manufacturers built a network with independent dealers to provide adequate level of service to their customer. As a result, most of the spare parts management systems are decentralized. In the decentralized network, competition exists between after-sales service providers. There are significant benefits of collaboration between the after-sales service providers, at the same time. Coopetition is a new term introduced to define collaboration efforts in a competitive environment. In this study, benefits of resource pooling and information sharing is considered, which is the benefit of coopetition, in other words.

The spare parts management system considered in this study is assumed to be a decentralized system with two independent dealers. In this setting, optimal operating policy of one of the dealers is considered given the other dealer's operating policy. In the first part of the study, collaboration among after-sales service providers of the spare parts network on inventory and service is concentrated. Optimal operating policy of a dealer considered under full information situation. The optimal policy is characterized and analyzed. Also, for comparison purposes, a centralized system setting where those two dealers operate under a single authority is considered. The comparison of optimal policies under decentralized and centralized systems shows that competition affects rationing of resources oppositely, which is inline with the previous literature (Zhao et al. [2008]). With the decreasing inventory level of the other dealer, the dealer under consideration shares more of his inventory under centralized system in order to achieve a better system profit, while under competition, the dealer shares less to prevent profit deterioration which might be caused providing poorer service to own customers. Monotonic behavior of the optimal policies under decentralized and centralized settings are analyzed. Conditions that lead monotonicity for the centralized setting are developed and shown to be the sufficient conditions for optimal monotone policy. On the other hand, necessary and sufficient conditions and sufficient conditions are presented for decentralized setting, but optimal policy cannot be proven to be monotonic. In some numerical instances, exceptions for the monotonic behavior are found, although monotonicity is satisfied in most of the cases. Special cases where monotonicity is valid could not be characterized.

Comparison of several pooling strategies is made to understand the effects on benefits of inventory and service pooling with respect to a given pooling strategy. The pooling strategies considered are dynamic-pooling (or optimal-pooling), static-pooling, full-pooling and no-pooling. Under some parameter settings, full-pooling strategy is found to be worse than no-pooling strategy. This is an important result since no-pooling strategy indicates an isolated operation of a dealer alone, which can be better than a pooling strategy if it is established myopic and inappropriate. This result contradicts with the previous results in literature. For example, Tagaras and Cohen [1992] state that complete pooling strategy is a dominating strategy. The difference comes from the definition of the strategy actually. In the literature, most of the studies are based on production system and pooling is used for transshipment of parts. The system used in this study is based on service rather than production, in which not only the parts are transshipped but also the customers may be directed to enter another queue in case of backordering.

In the second part of the study, the second dimension in collaboration among after-sales service providers of the spare parts network is concentrated, which is the information sharing dimension. A conclusion from the first part's work, which suits on information sharing is as follows. By comparing dynamic-pooling and static-pooling strategies, it is observed that the dealer under consideration has not much extra benefits by dynamically adjusting his base-stock, rationing and transshipment levels. Instead of those, using fixed values for the operating policy does not deteriorate profit much, which is practically easier to implement for a practitioner.

The work on information availability effects is not limited to the above discussion. Information incompleteness on own net inventory level of the dealer under consideration and on the other dealer's net inventory level are considered, which are the situations that companies face with in practice. For the model on the other dealer's net inventory level, no-information and full-information availabilities are considered. Under full-information, it is observed that when a variable payment amount (which depends on the other dealer's net inventory level and decreases with increasing net inventory value) is used, the behavior of rationing and transshipment levels change compared to fixed payment usage. Under variable payment amount, the dealer enjoys the other dealer's resources by lowering transshipment level at lower variable payment levels for higher net inventory levels of the other dealer and meets more requests of the other dealer's by lowering rationing level at higher inventory levels of the other dealer where variable payment is higher. Another observation about variable payment is that the benefit of information availability is higher under variable payment tariff compared to fixed payments, since the dealer can use the advantage of variable payment under full-information, which is not possible under no-information.

Information availability effects on the other dealer's net inventory level are analyzed with respect to parameter values and management insights are described with intuitions. Similar analysis is not made on the net inventory level of the dealer under consideration, because of difficulties faced at solution phase. Two levels of information incompleteness on the dealer's net inventory level are considered, viz. on inventory status and on customer status. It is found that, on average, the information on inventory status is more valuable compared to information on customer status.

5.2 FUTURE RESEARCH

There are several directions for future research opportunities. A game theoretical point of view to further analyze the optimal pooling policy is a future research direction. The model proposed in Chapter 3 is about the optimal policy of a dealer given the other dealer's known policy, and by this extension, the model can be used for an equilibrium solution assuming both dealers are optimizing their policies given the other's reaction. Alternatively, bi-level programming can be used, where one of the dealers (leader) makes policy optimization and the other makes policy optimization given the dealer's policy.

Customer waiting cost is always under question in the literature due to the problems with its quantification. Silver et al. [1998] define the costs that are parts of the waiting cost as nebulous, like uncertainty of a sales to be lost while the customer is waiting and cost of loosing goodwill due to poor service. Instead of attempting to quantify the disservice costs, a multi-objective approach to the pooling problem can be used, where an efficient frontier can be described by finding the optimal policies with respect to profit (excluding waiting cost) given a service level (or service level constraints). Then, the analysis of profit versus customer satisfaction would be useful from a managerial perspective.

Extending the problem of two dealers to multiple dealers or participants, such as including a manufacturer case is another future research opportunity. Inclusion of a manufacturer may lead to a multi-echelon structure, which might make finding the optimal policy harder. Inclusion of a manufacturer or increasing the number of dealers require a wider state-space, which brings computational burden. Consider a system composed of three dealers, where the policies of two dealers are given and the dealer under consideration optimizes his policy. Under optimal-pooling, possibly the dealer under consideration would have higher profits compared to two-dealer system, as he face with more secondary supply and demand and makes policy optimization. Under full-pooling for the same parameter values where no-pooling is better, the dealer under consideration might have lower profits compared to two-dealer system.

Single-item assumption can be challenged and multi-item structure for the case of kit structure can be considered, which may occur in practice like many different parts are required at the same time after a traffic accident and the service is completed only after having all the parts. This extension also might make finding the optimal policy harder, due to modelling complexity.

Another fruitful extension is about partial information. In Chapter 4, having incomplete information on a dealer's own status is modelled by knowing a range for net inventory amount but not the exact quantity, and it is found that such partial information not lead to significant deterioration on profit. The current model can be extended as follows: given a dealer's net inventory level, the dealer may have probabilities of observing another level. This modelling approach is similar to Smallwood and Sondik [1973] and would be more realistic.

REFERENCES

P. Alfredson and J. Verrijdt. Modeling emergency supply flexibility in a two-echelon inventory system. *Management Science*, 45(10):1416–1431, 1999.

S. Axsäter. Modelling emergency lateral transshipments in inventory systems. *Management Science*, 36(11):1329–1338, 1990.

S. Axsäter. A new decision rule for lateral transshipments in inventory systems. *Management Science*, 49(9):11681179, 2003.

A. Bensoussan, M. Çakanyıldırım, and S.P. Sethi. On the optimal control of partially observed inventory systems. *C. R. Acad. Sci. Paris*, 341:419426, 2008.

G.P. Cachon. *Competitive Supply Chain Inventory Management*. Quantitative Models for Supply Chain Management. Kluwer, 2002.

A. Clark and H. Scarf. Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4):475–490, 1960.

M. A. Cohen, N. Agrawal, and V. Agrawal. Winning in the aftermarket. *Harvard Business Review*, (May):129–138, 2006.

M.A. Cohen, C. Cull, H.L. Lee, and D. Willen. Saturn's supply-chain innovation: High value in after-sales service. *Sloan Management Review*, 41(4), 2000.

N. Çömez. Optimal transshipments and orders: A tale of two competing and cooperating retailers. Working paper, 2009.

N. Çömez, K. Stecke, and M. Çakanyıldırım. Multiple in-cycle transshipments with positive delivery times. Working paper, 2007.

M. Dada. A two-echelon inventory system with priority shipments. *Management Science*, 38(8):1140–1153, 1992.

F. de Véricourt, F. Karaesmen, and Y. Dallery. Dynamic scheduling in a make-to-stock system a partial characterization of optimal policies. *Operations Research*, 48(5):811–819, 2000.

J.C. Fransoo, M.J.F. Wouters, and T.G. de Kok. Multi-echelon multi-company inventory planning with limited information exchange. *Journal of the Operational Research Society*, 52:830–838, 2001.

A.M. Fraser. *Hidden Markov Models and Dynamical Systems*. SIAM, 2008.

J. Grahovac and A. Chakravarty. Sharing and lateral transshipment of inventory in a supply chain with expensive low-demand items. *Management Science*, 47(4): 579–594, 2001.

A.Y. Ha. Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, 43(8):1093–1103, 1997a.

A.Y. Ha. Optimal dynamic scheduling policy for a make-to-stock production system. *Operations Research*, 45:42–53, 1997b.

Y.T. Herer and A. Rashit. Lateral stock transshipments in a two-location inventory system with fixed and joint replenishment costs. *Naval Research Logistics*, 46: 525–547, 1999.

Y.T. Herer and M. Tzur. The dynamic transshipment problem. *Naval Research Logistics*, 48:386–408, 2001.

Y.T. Herer, M. Tzur, and E. Yücesan. Transshipments: An emerging inventory recourse to achieve supplychain leagility. *International Journal of Production Economics*, 80:201–212, 2002.

Y.T. Herer, M. Tzur, and E. Yücesan. The multilocation transshipment problem. *IIE Transactions*, 38:185–200, 2006.

B. Jung, B. Sun, J. Kim, and S. Ahn. Modeling lateral transshipments in multiechelon repairable-item inventory systems with finite repair channels. *Computers and Opeartions Research*, 30:1401–1417, 2003.

A. Kaplan. Stock rationing. *Management Science*, 15(5):260–267, 1969.

F.J.P. Karsten, M. Slikker, and G.J. van Houtum. Spare parts inventory pooling games. Working paper, 2009.

W.J. Kennedy, J.W. Patterson, and L.D. Fredendall. An overview of recent literature on spare parts inventories. *International Journal of Production Economics*, 76: 201–215, 2002.

P. Keskinocak and S. Savaşaneril. Collaborative procurement among competing buyers. *Naval Research Logistics*, 55:516–540, 2008.

J. Kilpi and A.P.J. Vepsalainen. Pooling of spare components between airlines. *Journal of Air Transport Management*, 10:137–146, 2004.

A. Kranenburg and G.J. van Houtum. A new partial pooling structure for spare parts networks. *European Journal of Operational Research*, 199:908921, 2009.

A. Kukreja, C.P. Schmidt, and D.M. Miller. Stocking decisions for low-usage items in a multilocation inventory system. *Management Science*, 47(10):1371–1383, 2001.

K.C. Laudon and J.P. Laudon. *Management Information Systems*. Prentice Hall, 2006. 9th edition.

H.L. Lee. A multi-echelon inventory model for repairable items with emergency lateral transshipments. *Management Science*, 33(10):1302–1316, 1987.

H.L. Lee, V. Padmanabhan, and S. Whang. The bullwhip effect in supply chains. *Sloan Management Review*, 38(3):93102, 1997.

M. Leng and M. Parlar. Game theoretic applications in supply chain management: A review. *INFOR*, 43(3):187220, 2005.

S.A. Lippman. Applying a new device in the optimization of exponential queuing systems. *Operations Research*, 23(4):687–710, 1975.

P. Milgrom and C. Shannon. Monotone comparative statics. *Econometrica*, 62(1): 157–180, 1994.

S. Mitra and A.K. Chatterjee. Leveraging information in multi-echelon inventory systems. *European Journal of Operational Research*, 152:263–280, 2004.

E.C. Moncrief, R.M. Schroder, and M.P. Reynolds. *Optimizing the MRO Inventory Asset: Production Spare Parts*. Industrial Press, 2005.

J.A. Muckstadt. A model for a multi-item, multi-echelon, multi-indenture inventory system. *Management Science*, 20(4):472–481, 1973.

S. Nahmias and W.S. Demmy. Operating characteristics of an inventory system with rationing. *Management Science*, 27(11):1236–1245, 1981.

J.A. Narus and J.C. Anderson. Rethinking distribution- adaptive channels. *Harvard Business Review*, pages 112–120, 1996.

C. Paterson, G. Kiesmüller, R. Teunter, and K. Glazebrook. Inventory models with lateral transshipments: A review. Working paper, 2009.

M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.

N. Rudi, S. Kapur, and D. F. Pyke. A two-location inventory model with transshipment and local decision making. *Management Science*, 47(12):1668–1680, 2001.

Y. Serin. *Implementable Policies for Markov Decision Processes*. University of North Carolina at Chapel Hill, 1989. PhD Dissertation.

Y. Serin and V. Kulkarni. Markov decision processes under observability constraints. *Mathematical Methods for Operations Research*, 61:311–328, 2005.

C. Sherbrooke. Metric: A multi-echelon technique for recoverable item control. *Operations Research*, 16:122–141, 1968.

E.A. Silver, D.F. Pyke, and R. Peterson. *Inventory Management and Production Planning and Scheduling*. John Wiley Sons, 1998.

R.D. Smallwood and E.J. Sondik. The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973.

E.J. Sondik. The optimal control of partially observable markov processes over the infinite horizon: Discounted costs. *Operations Research*, 26(2):282–304, 1978.

G. Tagaras and M.A. Cohen. Pooling in two-location inventory systems with non-negligible replenishment lead times. *Management Science*, 38(8):1067–1083, 1992.

D.M. Topkis. *Optimal Policies for Stocking and Issuing Inventory when Several Demand Classes Exist*. Decision Studies Group Report, 1966.

D.M. Topkis. Optimal ordering and rationing policies in a nonstationary dynamic inventory model with n demand classes. *Management Science*, 15:160–176, 1968.

G.J. van Houtum. Suggestions on parameter setting. Personal communication, 2008.

A.C.C. van Wijk, I.J.B.F. Adan, and G.J. van Houtum. Optimal lateral transshipment policy for a two location inventory problem. Working paper, 2009.

C.C. White. Procedures for the solution of a finite-horizon, partially observed, semi-markov optimization problem. *Operations Research*, 24(2):348–358, 1976.

H. Wong, D. Cattrysse, and D. van Oudheusden. Inventory pooling of repairable spare parts with non-zero lateral transshipment time and delayed lateral transshipments. *European Journal of Operational Research*, 165:207–218, 2005a.

H. Wong, B. Kranenburg, G.J. van Houtum, and D. Cattrysse. Efficient heuristics for two-echelon spare parts inventory systems with an aggregate mean waiting time constraint per local warehouse. Cardiff Business School Working Paper Series, 2005b.

H. Wong, G.J. van Houtum, D. Cattrysse, and D. van Oudheusden. Multi-item spare parts systems with lateral transshipments and waiting time constraints. *European Journal of Operational Research*, 171:10711093, 2006.

H. Wong, D. van Oudheusden, and D. Cattrysse. Cost allocation in spare parts inventory pooling. *Transportation Research Part E*, 43:370386, 2007.

K. Xu, P.T. Evers, and M.C. Fu. Estimating customer service in a two-location continuous review inventory model with emergency transhipments. *European Journal of Operational Research*, 145:569584, 2003.

X. Yan and H. Zhao. Decentralized inventory sharing with asymmetric information. Working paper (Krannert School of Management, Purdue University), 2008.

H. Zhao, V. Deshpande, and J.K. Ryan. Inventory sharing and rationing in decentralized dealer networks. *Management Science*, 51(4):531–547, 2005.

H. Zhao, V. Deshpande, and J.K. Ryan. Emergency transsipment in decentralized dealer networks: When to send and accept transshipment requests. *Naval Research Logistics*, 53:547–567, 2006.

H. Zhao, J.K. Ryan, and V. Deshpande. Optimal dynamic production and inventory transshipment policies for a two-location make-to-stock system. *Operations Research*, 56(2):400–410, 2008.

X. Zhao and M. Qiu. Information sharing in a multi-echelon inventory system. *Tsinghua Science And Technology*, 12(4):466–474, 2007.

APPENDIX A

PROOF OF PROPOSITION 3.2.1

To show that conditions **Q1-Q4** are necessary and sufficient for the optimal policy to have the monotone structure, we use the following theorem on monotonicity.

Theorem A.0.1 *(Monotonicity Theorem adopted from Milgrom and Shannon [1994]). Let A be the action space and S be the state space. Let $f : A \times S \to R$, where A is a lattice, S is a partially ordered set and $A' \subset A$. Then $argmax_{a \in A'} f(a, s)$ is monotone nondecreasing in (A', s) if and only if f is quasisupermodular in a and satisfies the single crossing property in (a, s).*

We first show that f is quasisupermodular in a. For the problem under consideration, in Section 3.1 we define three action types $a_1 \in A1 = \{accept, ltr\}$, $a_2 \in A2 = \{accept, reject\}$ and $a_3 \in A3 = \{accept, reject\}$. The action space is $A = A1$ x $A2$ x $A3$. Suppose the following ordering is defined between the actions:

$$(accept, a_2, a_3) > (ltr, a_2, a_3) \quad \forall a_2 \in A2, \forall a_3 \in A3,$$
$$(a_1, accept, a_3) > (a_1, reject, a_3) \quad \forall a_1 \in A1, \forall a_3 \in A3,$$
$$(a_1, a_2, reject) > (a_1, a_2, accept) \quad \forall a_1 \in A1, \forall a_2 \in A2.$$

Furthermore suppose the following natural ordering is defined between the states:

$$(i + 1, j) > (i, j), \ \forall j \in Z.$$
$$(i, j + 1) > (i, j), \ \forall i \in Z.$$
$$(i, j) > (i, j - 1), \ \forall i \in Z.$$

Note that both action space A and state space S are lattices, and every lattice is a partially-ordered set.

A function $f : A \to R$ is quasisupermodular if $f(x) \geq f(x \wedge y)$ implies $f(x \vee y) \geq f(y)$ (Milgrom and Shannon [1994]). In our problem, $f(a, s)$ is simply the right-hand side of Equation 3.10 evaluated under action $a = (a_1, a_2, a_3)$. Letting $x = (a, s)$ and $y = (a', s)$, the function $f(a, s)$ (or equivalently, $f((a_1, a_2, a_3), (i, j))$) should satisfy the following:

$$f(a, i, j) \geq f(a \wedge a', i, j) \Rightarrow f(a \vee a', i, j) \geq f(a', i, j) \quad \forall (i, j).$$

$$f(a_1, a_2, a_3, i, j) \geq f(\min\{a_1, a_1'\}, \min\{a_2, a_2'\}, \min\{a_3, a_3'\}, i, j) \Rightarrow \\ f(\max\{a_1, a_1'\}, \max\{a_2, a_2'\}, \max\{a_3, a_3'\}, i, j) \geq f(a_1', a_2', a_3', i, j) \; \forall (i, j). \tag{A.1}$$

Note that $f(a, i, j)$ is the sum of individual functions of $v(i, j)$ (denoted with operators in Equation 3.10) evaluated under action a, and each function is only affected by the corresponding action. In other words, if $f(a_i, a_{-i}, i, j) \geq f(a_i', a_{-i}, i, j)$ for $i, j \in Z$, for some $a_{-i} \in A_{-i}$, then $f(a_i, a_{-i}, i, j) \geq f(a_i', a_{-i}, i, j)$ for $i, j \in Z$, for all $a_{-i} \in A_{-i}$. Therefore the condition in (A.1) is trivially satisfied and f is quasisupermodular in a.

A function $f : A \times S \to R$ satisfies single-crossing property in (a, s), if for $a' > a''$ and $s' > s''$, $f(a', s'') \geq f(a'', s'')$ implies that $f(a', s') \geq f(a'', s')$ (Milgrom and Shannon [1994]). For our problem, we observe that the single crossing property is satisfied for all a and $s = (i, j)$, if $v(i, j)$ satisfies the conditions stated in Proposition 3.2.1.

Condition **Q1** is constructed as follows. Consider $(a_1, a_2, reject) > (a_1, a_2, accept)$ and $(i, j) > (i - 1, j)$. For f to satisfy the single crossing property, it should hold that

$$f(a_1, a_2, reject, i - 1, j) \geq f(a_1, a_2, accept, i - 1, j) \Rightarrow f(a_1, a_2, reject, i, j) \\ \geq f(a_1, a_2, accept, i, j).$$

Equivalently,

$$v(i - 1, j) \geq v(i, j) \Rightarrow v(i, j) \geq v(i + 1, j),$$

which is **Q1** in Proposition 3.2.1. Condition **Q1** simply implies the existence of $S_1(j)$ in that, for a given j if (upon a production completion) at state $i - 1$ accepting production is preferred to stopping production, then at state i accepting production is the preferred action. For a given j, the maximum i value at which stopping production is preferred is defined as $S_1(j)$.

Condition **Q2** is constructed as follows. Consider $(a_1, a_2, reject) > (a_1, a_2, accept)$ and $(i, j+1) > (i, j)$. For f to satisfy the single crossing property, it should hold that

$$f(a_1, a_2, reject, i, j) \geq f(a_1, a_2, accept, i, j) \Rightarrow f(a_1, a_2, reject, i, j+1)$$
$$\geq f(a_1, a_2, accept, i, j+1).$$

Equivalently,

$$v(i, j) \geq v(i+1, j) \Rightarrow v(i, j+1) \geq v(i+1, j+1),$$

which is **Q2** in Proposition 3.2.1. Condition **Q2** simply implies the monotonicity of $S_1(j)$ in that, for a given j if (upon a production completion) at state i accepting production is preferred to stopping production, then at state $j+1$ accepting production is the preferred action for the same i.

Condition **Q3** is constructed as follows. Consider $(a_1, accept, a_3) > (a_1, reject, a_3)$ and $(i+1, j) > (i, j)$. For f to satisfy the single crossing property, it should hold that

$$f(a_1, accept, a_3, i, j) \geq f(a_1, reject, a_3, i, j) \Rightarrow f(a_1, accept, a_3, i+1, j)$$
$$\geq f(a_1, reject, a_3, i+1, j).$$

Equivalently,

$$v(i-1, j) + r \geq v(i, j-1) \Rightarrow v(i, j) + r \geq v(i+1, j-1),$$

which is **Q3** in Proposition 3.2.1. Condition **Q3** simply implies the existence of $K_1(j)$ in that, for a given j if (upon a request arrival from D2) at state i transshipping a part is preferred to not transshipping it, then at state $i+1$ transshipping is the preferred action. For a given j, the maximum i value at which not transshipping is preferred is defined as $K_1(j)$.

Condition **Q4** is constructed as follows. Consider $(a_1, accept, a_3) > (a_1, reject, a_3)$ and $(i, j+1) > (i, j)$. For f to satisfy the single crossing property, it should hold that

$$f(a_1, accept, a_3, i, j) \geq f(a_1, reject, a_3, i, j) \Rightarrow f(a_1, accept, a_3, i, j+1)$$
$$\geq f(a_1, reject, a_3, i, j+1).$$

Equivalently,

$v(i-1, j) + r \geq v(i, j-1) \Rightarrow v(i-1, j+1) + r \geq v(i, j)$,

which is **Q4** in Proposition 3.2.1. Condition **Q4** simply implies the monotonicity of $K_1(j)$ in that, for a given j if (upon a request arrival from D2) at state i transshipping a part is preferred to not transshipping it, then at state $j+1$ transshipping is the preferred action for the same i.□

APPENDIX B

PROOF OF THEOREM 3.3.1

In this part, an outline of the proof of Lemma 1 is given first to help the reader to follow the proof. Then, the proof continues.

OUTLINE OF THE PROOF:

Outline of the proof is given in Figure B.1. Each step of the proof is explained below.

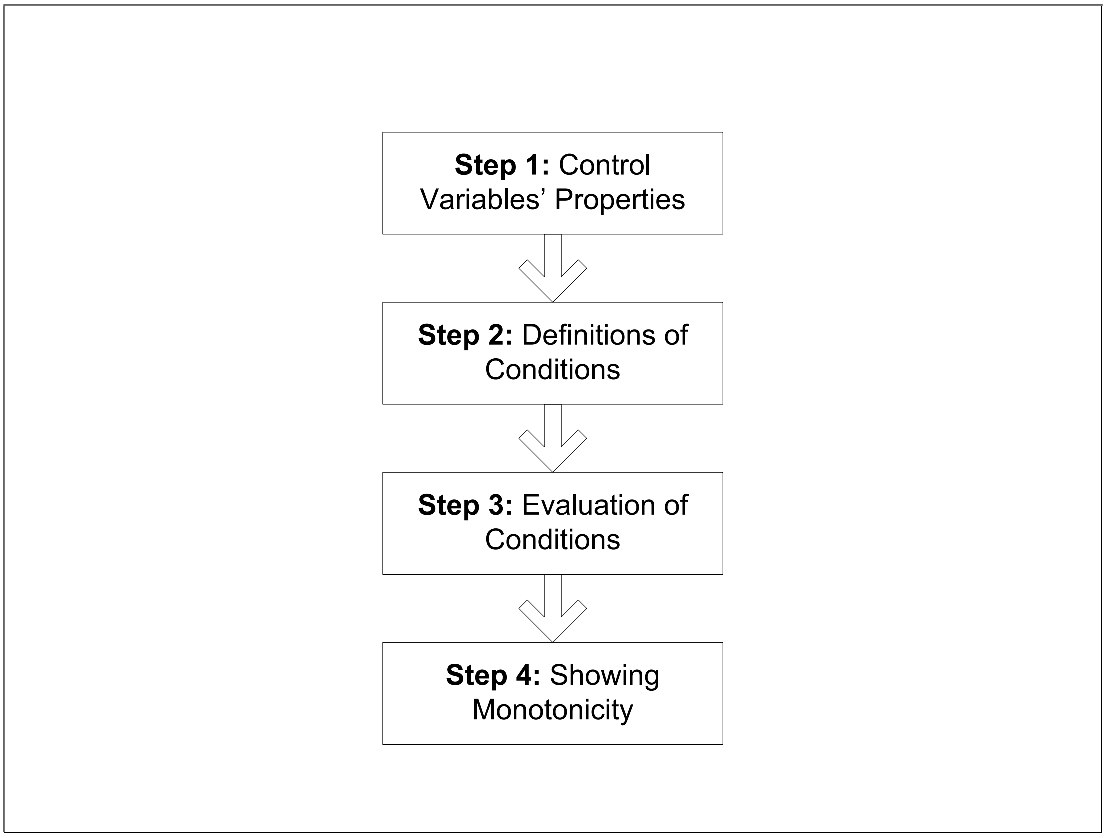


Figure B.1: Outline of the Proof of Monotonicity for Simplified Model

Step 1 Control Variables' Properties

As explained in Section 3.3, the central authority has several available actions upon occurrence of different events. Control Variables of the central authority indicate some states for which the actions change above and below those states for given events. Let S_1 be the state such that for $i < S_1$, manufacturing order is given to increase status at D1, and for $i \geq S_1$, no manufacturing order is given at D1. We can express this as:

$$v(i+1, j) > v(i, j) \quad for\ i < S_1\ and$$

$$v(i+1, j) \leq v(i, j) \quad for\ i \geq S_1$$

Therefore, S_1 can be defined as:

$$S_1(j) = min\{i \geq 0 | v(i+1, j) - v(i, j) \leq 0\} \quad \text{(B.1)}$$

Figure B.2 depicts plausible behavior of $v(i+1, j) - v(i, j)$ and the definition of S_1 based on that behavior.

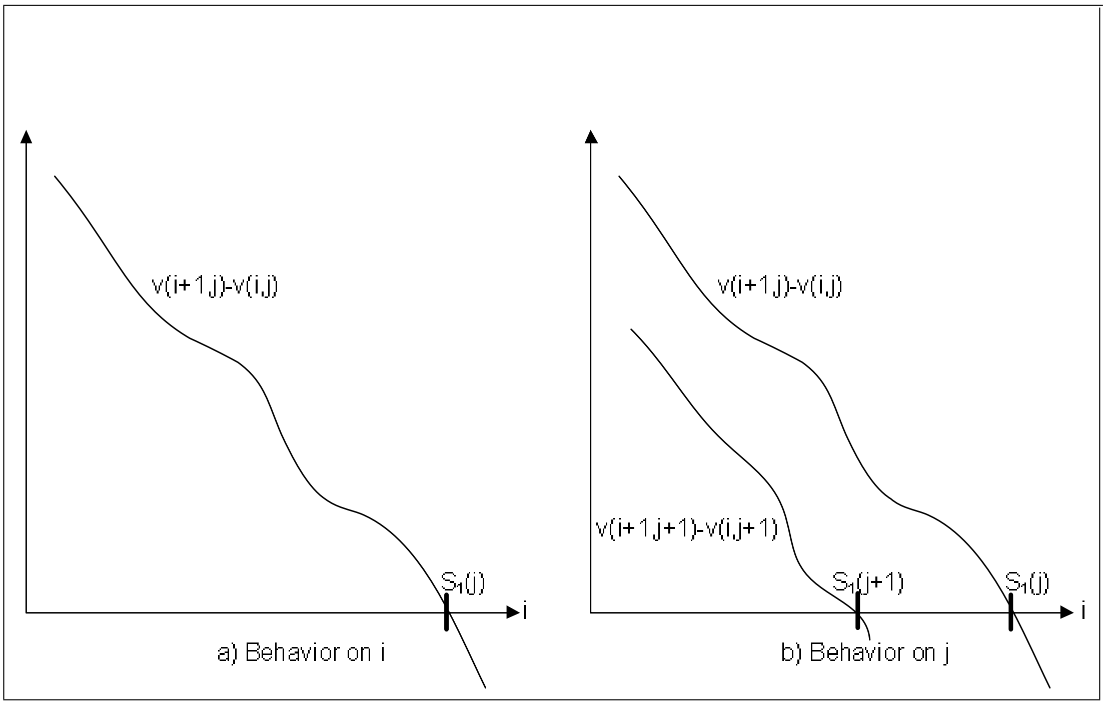


Figure B.2: Behavior of $v(i+1, j) - v(i, j)$ and S_1

In this proof, optimal function is shown to satisfy $v(i+1, j) - v(i, j)$ *non − increasing in i* and therefore from the definition of S_1 in Equation B.1, S_1

exists. Please note that $v(i+1, j) - v(i, j)$ non-increasing in i is a sufficient condition for the existence of S_1 (see Figure B.2-a). This results in condition 1 as follows:

$$v(i, j) - v(i-1, j) \geq v(i+1, j) - v(i, j) \qquad (C1')$$

Furthermore, S_1 is shown to be non-increasing in j. A sufficient condition, which is condition 3 actually, can be expressed as follows (see Figure B.2-b):

$$v(i+1, j) - v(i, j) \geq +v(i+1, j+1) - v(i, j+1) \qquad (C3')$$

This is equivalent to saying that the optimality function is submodular (or sub-additive).

In a similar manner, K_1 and Z_1 can be defined as follows. Let K_1 be the state such that for $K_1 < i \leq S_1$, lateral transshipment from D1 to D2 is made upon a customer arrival to D2, and for $K_1 \geq i$, D2 serves his customer by himself (i.e. without lateral transshipment). This means $v(i, j-1) > v(i-1, j) \quad for\ i \leq K_1\ and\ v(i, j-1) \leq v(i-1, j) \quad for\ K_1 < i < S_1$. Therefore, K_1 can be defined as:

$$K_1(j) = max\{i | v(i, j-1) - v(i-1, j) > 0\} \qquad (B.2)$$

Figure B.3 depicts plausible behavior of $v(i, j-1) - v(i-1, j)$ and the definition of K_1 based on that behavior.

In this proof, optimal function is shown to satisfy $v(i, j-1) - v(i-1, j)\ non-increasing\ in\ i$ and therefore from the definition of K_1 in Equation B.2, K_1 exists. Please note that $v(i, j-1) - v(i-1, j)$ non-increasing in i is a sufficient condition for the existence of K_1 (see Figure B.3-a). This results in condition 4 as follows:

$$v(i, j-1) - v(i-1, j) \geq v(i+1, j-1) - v(i, j) \qquad (C4')$$

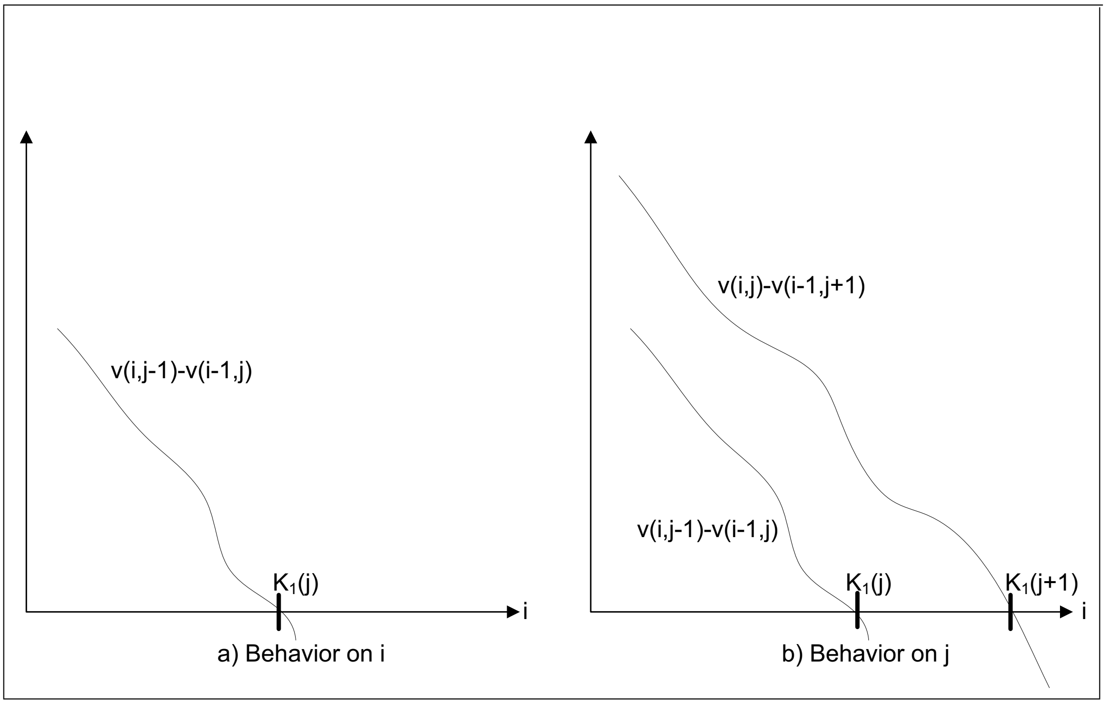


Figure B.3: Behavior of $v(i, j-1) - v(i-1, j)$ and K_1

Furthermore, K_1 is shown to be non-decreasing in j. A sufficient condition, which is condition 5 actually, can be expressed as follows (see Figure B.3-b):

$$v(i, j) - v(i-1, j+1) \geq v(i, j-1) - v(i-1, j) \qquad (C5')$$

This is equivalent to saying that the optimality function is submodular (or subadditive).

Let Z_1 be the state such that for $i \leq Z_1$, lateral transshipment from D2 to D1 is made upon a customer arrival to D1, and for $Z_1 < i \leq S_1$, D1 serves his customer by himself (i.e. without lateral transshipment). This means $v(i, j-1) > v(i-1, j) \quad for\ i \leq Z_1\ and\ v(i, j-1) \leq v(i-1, j) \quad for\ Z_1 < i \leq S_1$. Therefore, Z_1 can be defined as:

$$Z_1(j) = max\{i | v(i, j-1) - v(i-1, j) > 0\} \qquad (B.3)$$

Figure B.4 depicts plausible behavior of $v(i, j-1) - v(i-1, j)$ and the definition of Z_1 based on that behavior. It can be seen in the figure that Z_1 is the largest integer satisfying $v(i, j-1) - v(i-1, j) > 0$.

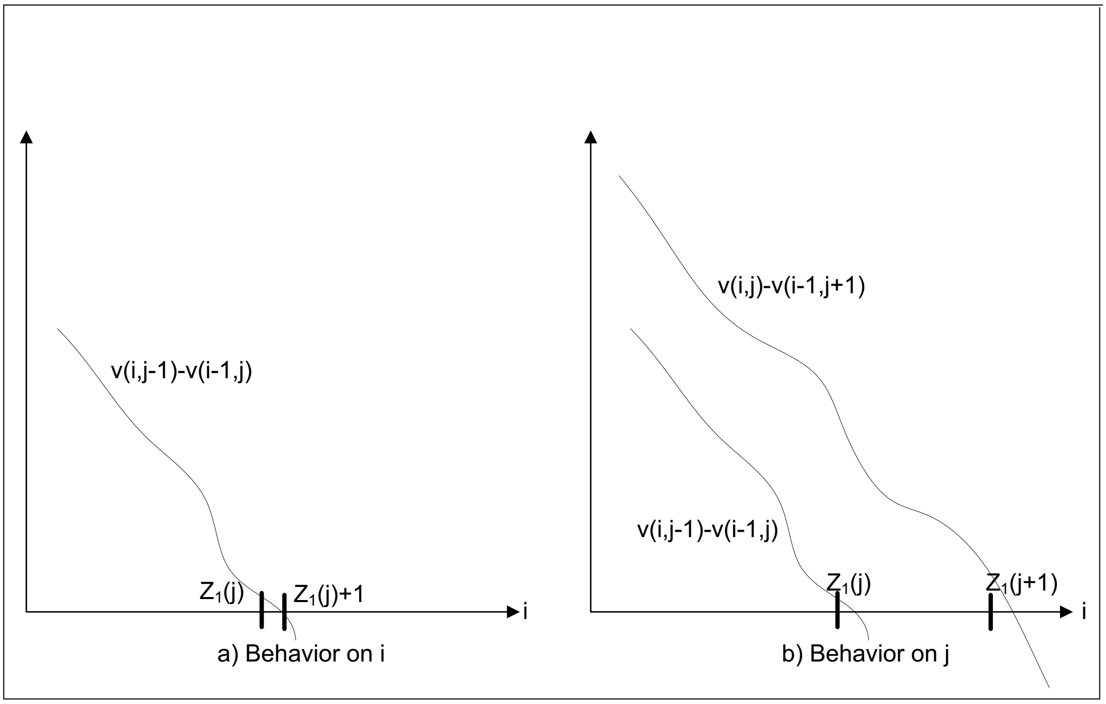


Figure B.4: Behavior of $v(i, j-1) - v(i-1, j)$ and Z_1

In this proof, optimal function is shown to satisfy $v(i, j-1) - v(i-1, j)$ *non* − *increasing in i* and therefore from the definition of Z_1 in Equation B.3, Z_1 exists. Please note that $v(i, j-1) - v(i-1, j)$ non-increasing in i is a sufficient condition for the existence of Z_1 (see Figure B.3-a). This results in condition 4 as follows:

$$v(i, j-1) - v(i-1, j) \geq v(i+1, j-1) - v(i, j) \qquad (C4')$$

Furthermore, Z_1 is shown to be non-decreasing in j. A sufficient condition, which is condition 5 actually, can be expressed as follows (see Figure B.4-b):

$$v(i, j) - v(i-1, j+1) \geq v(i, j-1) - v(i-1, j) \qquad (C5')$$

This is equivalent to saying that the optimality function is submodular (or sub-additive).

In a similar manner, S_2, K_2, Z_2 can be defined.

Step 2 Definitions of Conditions

Behavior of Control Variables on state is the basic question while analyzing monotonicity. For example, S_1 is the smallest status of D1 where $v(i+1, j) - v(i, j) \leq 0$. The behavior of S_1 (consequently the function $v(i+1, j) - v(i, j)$) on i is important to see the relation between state and D1's decision. Evaluating the difference on i of $v(i+1, j) - v(i, j)$ for observing the behavior of S_1 on i helps to see the behavior, and it is expected that the difference is nondecreasing on i, i.e. $v(i, j) - v(i-1, j) \geq v(i+1, j) - v(i, j)$. This is Condition 1, actually. It is also required to see the behavior of S_1 on j, which results in Condition 3 with a similar approach used above, but the difference the function $v(i+1, j) - v(i, j)$ is evaluated on j. Table B.1 summarizes the relation between Control Variables and the conditions.

Step 3 Evaluation of Conditions:

In Section 3.2, detailed explanation of the relations between conditions and control variables are referred to here. The relations are explained for centralized system setting below, but they also hold for decentralized system setting.

Conditions are as follows:

C1': $v(i, j) - v(i-1, j) \geq v(i+1, j) - v(i, j), \forall i, j$
C2': $v(i, j) - v(i, j-1) \geq v(i, j+1) - v(i, j), \forall i, j$
C3': $v(i+1, j) - v(i, j) \geq +v(i+1, j+1) - v(i, j+1), \forall i, j$
C4': $v(i, j-1) - v(i-1, j) \geq v(i+1, j-1) - v(i, j), \forall i, j$
C5': $v(i, j) - v(i-1, j+1) \geq v(i, j-1) - v(i-1, j), \forall i, j$

C1' and C2' are concavity conditions. C3', C4' and C5' are sub/supermodularity conditions and they imply C1' and C2'. When C3' at $(i, j-1)$ is added to C4' at (i, j), C1' is obtained; and when C3' at $(i-1, j)$ is added to C5' at (i, j), C2' is obtained. Therefore, it is only enough to consider V under C3', C4' and C5'.

Step 4 Showing Monotonicity:

Monotonicity is shown below in Lemma B.0.2.

Lemma B.0.2 *If $v \in V$, then $Tv \in V$, where $Tv(i, j) = -c_h i^+ - c_l i^- + \lambda_1 T_1 v(i, j) + \lambda_2 T_2 v(i, j) + \mu_1 T_3 v(i, j) + \mu_2 T_4 v(i, j)$.*

Proof (Lemma B.0.2). In the proof we first show that if $v \in V$ then $T_k v \in V$ for $k = 1, 2, 3, 4$. Then, we show that $Tv \in V$. Note that **C1'-C2'** are implied by **C3'-**

Table B.1: Relation between Control Variables and Conditions

Condition	Expression	Related Control Variable(s)	Indicates
C1'	$2v(i, j) \geq v(i+1, j) + v(i-1, j)$	S_1	existence
C2'	$2v(i, j) \geq v(i, j+1) + v(i, j-1)$	S_2	existence
C3'	$v(i+1, j) + v(i, j+1) \geq v(i, j) + v(i+1, j+1)$	S_1	monotone ↓
		S_2	monotone ↓
C4'	$v(i, j) + v(i, j-1) \geq v(i-1, j) + v(i+1, j-1)$	K_1	existence
		K_2	monotone ↑
		Z_1	existence
		Z_2	monotone ↑
C5'	$v(i, j) + v(i-1, j) \geq v(i, j-1) + v(i-1, j+1)$	K_1	monotone ↑
		K_2	existence
		Z_1	monotone ↑
		Z_2	existence

C5'. In the proof we only show that Tv satisfies **C3'-C5'**. The proof follows similar lines with Ha [1997a] and Zhao et al. [2008].

FIRST PART: Show $T_1 v \in V$.

$$T_1 v(i, j) = \max\{v(i-1, j) + R, v(i, j-1) + R\}$$

defining w to be a function on $\{0, 1\} \times Z^2$ such that

$$\begin{aligned} w(u, i, j) &= u[v(i-1, j) + R] + (1-u)[v(i, j-1) + R] \\ &= \begin{cases} v(i, j-1) + R & \text{if } u = 0 \\ v(i-1, j) + R & \text{if } u = 1 \end{cases} \end{aligned}$$

thus $T_1 v(i, j) = \max_{u \in \{0,1\}} w(u, i, j)$.

(i) $T_1 v$ satisfies **C3'**: We want to prove that $T_1 v(i+1, j) + T_1 v(i, j+1) \geq T_1 v(i, j) + T_1 v(i+1, j+1)$. Let $u_1,\ u_2 \in \{0, 1\}$ be such that $T_1 v(i, j) = w(u_1, i, j),\ T_1 v(i+1, j+1) = w(u_2, i+1, j+1)$.

Case 1: $u_1 = 0, u_2 = 0$

$$\begin{aligned} T_1 v(i+1, j) + T_1 v(i, j+1) &\geq w(0, i+1, j) + w(0, i, j+1) \\ &= v(i+1, j-1) + v(i, j) + 2R \\ &\geq v(i, j-1) + v(i+1, j) + 2R \\ &= w(0, i, j) + w(0, i+1, j+1) \\ &= T_1 v(i, j) + T_1 v(i+1, j+1) \end{aligned}$$

where the first inequality follows from the definition of $T_1 v(i, j)$, the second inequality follows from C3', and the others (i.e., equalities) follow from the definition of $w(u, i, j)$.

Case 2: $u_1 = 0, u_2 = 1$

$$\begin{aligned} T_1 v(i+1, j) + T_1 v(i, j+1) &\geq w(1, i+1, j) + w(0, i, j+1) \\ &= v(i, j) + v(i, j) + 2R \\ &\geq v(i, j-1) + v(i, j+1) + 2R \\ &= w(0, i, j) + w(1, i+1, j+1) \\ &= T_1 v(i, j) + T_1 v(i+1, j+1) \end{aligned}$$

where the second inequality follows from C2'.

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_1v(i+1,j) + T_1v(i,j+1) &\geq w(1,i+1,j) + w(0,i,j+1) \\ &= v(i,j) + v(i,j) + 2R \\ &\geq v(i-1,j) + v(i+1,j) + 2R \\ &= w(1,i,j) + w(0,i+1,j+1) \\ &= T_1v(i,j) + T_1v(i+1,j+1) \end{aligned}$$

where the first inequality follows from the definition of $T_1v(i,j)$, the second inequality follows from C1'.

Case 4: $u_1 = 1, u_2 = 1$

$$\begin{aligned} T_1v(i+1,j) + T_1v(i,j+1) &= w(1,i+1,j) + w(1,i,j+1) \\ &= v(i,j) + v(i-1,j+1) + 2R \\ &\geq v(i-1,j) + v(i,j+1) + 2R \\ &= w(1,i,j) + w(1,i+1,j+1) \\ &= T_1v(i,j) + T_1v(i+1,j+1) \end{aligned}$$

where the first equality follows from the definition of $T_1v(i,j)$, the inequality follows from C3'.

(ii) T_1v satisfies **C4'**:

We want to prove that $T_1v(i,j) + T_1v(i,j-1) \geq T_1v(i-1,j) + T_1v(i+1,j-1)$. Let u_1 and u_2 be such that $T_1v(i-1,j) = w(u_1,i-1,j),\ T_1v(i+1,j-1) = w(u_2,i+1,j-1)$

Case 1: $u_1 = 0, u_2 = 0$

$$\begin{aligned} T_1v(i,j) + T_1v(i,j-1) &\geq w(0,i,j) + w(0,i,j-1) \\ &= v(i,j-1) + v(i,j-2) + 2R \\ &\geq v(i-1,j-1) + v(i+1,j-2) + 2R \\ &= w(0,i-1,j) + w(0,i+1,j-1) \\ &= T_1v(i-1,j) + T_1v(i+1,j-1) \end{aligned}$$

where the first inequality follows from the definition of $T_1v(i,j)$, the second inequality follows from C4'.

Case 2: $u_1 = 0, u_2 = 1$

$$\begin{aligned} T_1v(i,j) + T_1v(i,j-1) &\geq w(0,i,j) + w(1,i,j-1) \\ &= v(i,j-1) + v(i-1,j-1) + 2R \\ &= v(i-1,j-1) + v(i,j-1) + 2R \\ &= w(0,i-1,j) + w(1,i+1,j-1) \\ &= T_1v(i-1,j) + T_1v(i+1,j-1) \end{aligned}$$

where the first inequality follows from the definition of $T_1v(i,j)$.

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_1v(i,j) + T_1v(i,j-1) &\geq w(1,i,j) + w(0,i,j-1) \\ &= v(i-1,j) + v(i,j-2) + 2R \\ &\geq v(i-2,j) + v(i+1,j-2) + 2R \\ &= w(1,i-1,j) + w(0,i+1,j-1) \\ &= T_1v(i-1,j) + T_1v(i+1,j-1) \end{aligned}$$

where the first equality follows from the definition of $T_1v(i,j)$, the second inequality follows from C4'.

Case 4:$u_1 = 1, u_2 = 1$

$$\begin{aligned} T_1v(i,j) + T_1v(i,j-1) &\geq w(1,i,j) + w(1,i,j-1) \\ &= v(i-1,j) + v(i-1,j-1) + 2R \\ &\geq v(i-2,j) + v(i,j-1) + 2R \\ &= w(1,i-1,j) + w(1,i+1,j-1) \\ &= T_1v(i-1,j) + T_1v(i+1,j-1) \end{aligned}$$

where the first inequality follows from the definition of $T_1v(i,j)$, the second inequality follows from C4'.

(iii) T_1v satisfies **C5'**:

We want to prove that $T_1v(i,j) + T_1v(i-1,j) \geq T_1v(i,j-1) + T_1v(i-1,j+1)$. Let u_1 and u_2 be such that $T_1v(i,j-1) = w(u_1,i,j-1)$, $T_1v(i-1,j+1) = w(u_2,i-1,j+1)$

Case 1:$u_1 = 0, u_2 = 0$

$$\begin{aligned}
T_1v(i,j) + T_1v(i-1,j) &\geq w(0,i,j) + w(0,i-1,j) \\
&= v(i,j-1) + v(i-1,j-1) + 2R \\
&\geq v(i,j-2) + v(i-1,j) + 2R \\
&= w(0,i,j-1) + w(0,i-1,j+1) \\
&= T_1v(i,j-1) + T_1v(i-1,j+1)
\end{aligned}$$

where the first inequality follows from the definition of $T_1v(i,j)$, the second inequality follows from C5'.

<u>Case 2:$u_1 = 0, u_2 = 1$</u>

$$\begin{aligned}
T_1v(i,j) + T_1v(i-1,j) &\geq w(0,i,j) + w(1,i-1,j) \\
&= v(i,j-1) + v(i-2,j) + 2R \\
&\geq v(i,j-2) + v(i-2,j+1) + 2R \\
&= w(0,i,j-1) + w(1,i-1,j+1) \\
&= T_1v(i,j-1) + T_1v(i-1,j+1)
\end{aligned}$$

where the first inequality follows from the definition of $T_2v(i,j)$, the second inequality follows from C5'.

<u>Case 3: $u_1 = 1, u_2 = 0$</u>

$$\begin{aligned}
T_1v(i,j) + T_1v(i-1,j) &\geq w(1,i,j) + w(0,i-1,j) \\
&= v(i-1,j) + v(i-1,j-1) + 2R \\
&= v(i-1,j-1) + v(i-1,j) + 2R \\
&= w(1,i,j-1) + w(0,i-1,j+1) \\
&= T_1v(i,j-1) + T_1v(i-1,j+1)
\end{aligned}$$

where the first inequality follows from the definition of $T_1v(i,j)$.

<u>Case 4:$u_1 = 1, u_2 = 1$</u>

$$\begin{aligned}T_1v(i,j)+T_1v(i-1,j) &\geq w(1,i,j)+w(1,i-1,j)\\ &= v(i-1,j)+v(i-2,j)+2R\\ &\geq v(i-1,j-1)+v(i-2,j+1)+2R\\ &= w(1,i,j-1)+w(1,i-1,j+1)\\ &= T_1v(i,j-1)+T_1v(i-1,j+1)\end{aligned}$$

where the second inequality follows from C5'.

SECOND PART: Show $T_2v(i,j) \in V$

$$T_2v(i,j) = \max\{v(i-1,j)+R, v(i,j-1)+R\}$$

defining w to be a function on $\{0,1\} \times Z^2$ such that

$$\begin{aligned}w(u,i,j) &= u[v(i-1,j)+R]+(1-u)[v(i,j-1)+R]\\ &= \begin{cases} v(i,j-1)+R & \text{if } u=0\\ v(i-1,j)+R & \text{if } u=1\end{cases}\end{aligned}$$

thus

$$T_2v(i,j) = \max_{u\in\{0,1\}} w(u,i,j)$$

(i) T_2v satisfies **C3'**:

We want to prove that $T_2v(i+1,j)+T_2v(i,j+1) \geq T_2v(i,j)+T_2v(i+1,j+1)$. Let u_1 and u_2 be such that $T_2v(i,j) = w(u_1,i,j)$, $T_2v(i+1,j+1) = w(u_2,i+1,j+1)$

Case 1: $u_1 = 0, u_2 = 0$

$$\begin{aligned}T_2v(i+1,j)+T_2v(i,j+1) &= w(0,i+1,j)+w(0,i,j+1)\\ &= v(i+1,j-1)+v(i,j)+2R\\ &\geq v(i,j-1)+v(i+1,j)+2R\\ &= w(0,i,j)+w(0,i+1,j+1)\\ &= T_2v(i,j)+T_2v(i+1,j+1)\end{aligned}$$

where the first equality follows from the definition of $T_2v(i, j)$, the second inequality follows from C3', and the others follow from the definition of $w(u, i, j)$.

Case 2: $u_1 = 0, u_2 = 1$

$$\begin{aligned} T_2v(i+1, j) + T_2v(i, j+1) &\geq w(1, i+1, j) + w(0, i, j+1) \\ &= v(i, j) + v(i, j) + 2R \\ &\geq v(i, j-1) + v(i, j+1) + 2R \\ &= w(0, i, j) + w(1, i+1, j+1) \\ &= T_2v(i, j) + T_2v(i+1, j+1) \end{aligned}$$

where the second inequality follows from C2'.

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_2v(i+1, j) + T_2v(i, j+1) &\geq w(1, i+1, j) + w(0, i, j+1) \\ &= v(i, j) + v(i, j) + 2R \\ &\geq v(i-1, j) + v(i+1, j) + 2R \\ &= w(1, i, j) + w(0, i+1, j+1) \\ &= T_2v(i, j) + T_2v(i+1, j+1) \end{aligned}$$

where the second inequality follows from C1'.

Case 4: $u_1 = 1, u_2 = 1$

$$\begin{aligned} T_2v(i+1, j) + T_2v(i, j+1) &\geq w(1, i+1, j) + w(1, i, j+1) \\ &= v(i, j) + v(i-1, j+1) + 2R \\ &\geq v(i-1, j) + v(i, j+1) + 2R \\ &= w(1, i, j) + w(1, i+1, j+1) \\ &= T_2v(i, j) + T_2v(i+1, j+1) \end{aligned}$$

where the second inequality follows from C3'.

(ii) T_2v satisfies **C4'**:

We want to prove that $T_2v(i, j) + T_2v(i, j-1) \geq T_2v(i-1, j) + T_2v(i+1, j-1)$. Let u_1

and u_2 be such that $T_2v(i-1, j) = w(u_1, i-1, j)$, $T_2v(i+1, j-1) = w(u_2, i+1, j-1)$

Case 1: $u_1 = 0, u_2 = 0$

$$\begin{aligned} T_2v(i, j) + T_2v(i, j-1) &\geq w(0, i, j) + w(0, i, j-1) \\ &= v(i, j-1) + v(i, j-2) + 2R \\ &\geq v(i-1, j-1) + v(i+1, j-2) + 2R \\ &= w(0, i-1, j) + w(0, i+1, j-1) \\ &= T_2v(i-1, j) + T_2v(i+1, j-1) \end{aligned}$$

where the second inequality follows from C4'.

Case 2: $u_1 = 0, u_2 = 1$

$$\begin{aligned} T_2v(i, j) + T_2v(i, j-1) &\geq w(0, i, j) + w(1, i, j-1) \\ &= v(i, j-1) + v(i-1, j-1) + 2R \\ &= w(0, i-1, j) + w(1, i+1, j-1) \\ &= T_2v(i-1, j) + T_2v(i+1, j-1) \end{aligned}$$

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_2v(i, j) + T_2v(i, j-1) &\geq w(0, i, j) + w(1, i, j-1) \\ &= v(i, j-1) + v(i-1, j-1) + 2R \\ &\geq v(i-2, j) + v(i+1, j-2) + 2R \\ &= w(1, i-1, j) + w(0, i+1, j-1) \\ &= T_2v(i-1, j) + T_2v(i+1, j-1) \end{aligned}$$

where the second inequality follows from C4' and C5'.

Case 4: $u_1 = 1, u_2 = 1$

$$
\begin{aligned}
T_2v(i,j) + T_2v(i,j-1) &\geq w(1,i,j) + w(1,i,j-1) \\
&= v(i-1,j) + v(i-1,j-1) + 2R \\
&\geq v(i-2,j) + v(i,j-1) + 2R \\
&= w(1,i-1,j) + w(1,i+1,j-1) \\
&= T_2v(i-1,j) + T_2v(i+1,j-1)
\end{aligned}
$$

where the second inequality follows from C4'.

(iii) T_2v satisfies **C5'**:

We want to prove that $T_2v(i,j) + T_2v(i-1,j) \geq T_2v(i,j-1) + T_2v(i-1,j+1)$. Let u_1 and u_2 be such that $T_2v(i,j-1) = w(u_1,i,j-1)$, $T_2v(i-1,j+1) = w(u_2,i-1,j+1)$

Case 1: $u_1 = 0, u_2 = 0$

$$
\begin{aligned}
T_2v(i,j) + T_2v(i-1,j) &\geq w(0,i,j) + w(0,i-1,j) \\
&= v(i,j-1) + v(i-1,j-1) + 2R \\
&\geq v(i,j-2) + v(i-1,j) + 2R \\
&= w(0,i,j-1) + w(0,i-1,j+1) \\
&= T_2v(i,j-1) + T_2v(i-1,j+1)
\end{aligned}
$$

where the second inequality follows from C5'.

Case 2: $u_1 = 0, u_2 = 1$

$$
\begin{aligned}
T_2v(i,j) + T_2v(i-1,j) &\geq w(0,i,j) + w(1,i-1,j) \\
&= v(i,j-1) + v(i-2,j) + 2R \\
&\geq v(i,j-2) + v(i-2,j+1) + 2R \\
&= w(0,i,j-1) + w(1,i-1,j+1) \\
&= T_2v(i,j-1) + T_2v(i-1,j+1)
\end{aligned}
$$

where the second inequality follows from C5' (writing C5' for (i-1,j) and (i,j-1), then summing two inequalities).

Case 3: $u_1 = 1, u_2 = 0$

$$
\begin{aligned}
T_2v(i,j) + T_2v(i-1,j) &\geq w(1,i,j) + w(0,i-1,j) \\
&= v(i-1,j) + v(i-1,j-1) + 2R \\
&= v(i-1,j-1) + v(i-1,j) + 2R \\
&= w(1,i,j-1) + w(0,i-1,j+1) \\
&= T_2v(i,j-1) + T_2v(i-1,j+1)
\end{aligned}
$$

Case 4: $u_1 = 1, u_2 = 1$

$$
\begin{aligned}
T_2v(i,j) + T_2v(i-1,j) &\geq w(1,i,j) + w(1,i-1,j) \\
&= v(i-1,j) + v(i-2,j) + 2R \\
&\geq v(i-1,j-1) + v(i-2,j+1) + 2R \\
&= w(1,i,j-1) + w(1,i-1,j+1) \\
&= T_2v(i,j-1) + T_2v(i-1,j+1)
\end{aligned}
$$

where the second inequality follows from C5'.

THIRD PART: Show $T_3v(i,j) \in V$

$$T_3v(i,j) = \max\{v(i+1,j), v(i,j)\}$$

defining w to be a function on $\{0,1\} \times Z^2$ such that

$$
\begin{aligned}
w(u,i,j) &= u[v(i+1,j)] + (1-u)[v(i,j)] \\
&= \begin{cases} v(i,j) & \text{if } u = 0 \\ v(i+1,j) & \text{if } u = 1 \end{cases}
\end{aligned}
$$

thus

$$T_3v(i,j) = \max_{u \in \{0,1\}} w(u,i,j)$$

(i) T_3v satisfies **C3'**:

We want to prove that $T_3v(i+1,j) + T_3v(i,j+1) \geq T_3v(i,j) + T_3v(i+1,j+1)$. Let u_1 and u_2 be such that $T_3v(i,j) = w(u_1,i,j)$, $T_3v(i+1,j+1) = w(u_2,i+1,j+1)$

Case 1: $u_1 = 0, u_2 = 0$

$$\begin{aligned} T_3v(i+1,j) + T_3v(i,j+1) &\geq w(0,i+1,j) + w(0,i,j+1) \\ &= v(i+1,j) + v(i,j+1) \\ &\geq v(i,j) + v(i+1,j+1) \\ &= w(0,i,j) + w(0,i+1,j+1) \\ &= T_3v(i,j) + T_3v(i+1,j+1) \end{aligned}$$

where the first inequality follows from the definition of $T_3v(i,j)$, the second inequality follows from C3', and the others follow from the definition of $w(u,i,j)$.

Case 2: $u_1 = 0, u_2 = 1$

$$\begin{aligned} T_3v(i+1,j) + T_3v(i,j+1) &\geq w(1,i+1,j) + w(0,i,j+1) \\ &= v(i+2,j) + v(i,j+1) \\ &\geq v(i,j) + v(i+2,j+1) \\ &= w(0,i,j) + w(1,i+1,j+1) \\ &= T_3v(i,j) + T_3v(i+1,j+1) \end{aligned}$$

where the second inequality follows from C3' (writing C3' for (i,j) and (i+1,j), then summing two inequalities).

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_3v(i+1,j) + T_3v(i,j+1) &\geq w(0,i+1,j) + w(1,i,j+1) \\ &= v(i+1,j) + v(i+1,j+1) \\ &= w(1,i,j) + w(0,i+1,j+1) \\ &= T_3v(i,j) + T_3v(i+1,j+1) \end{aligned}$$

Case 4: $u_1 = 1, u_2 = 1$

$$\begin{aligned}
T_3v(i+1,j) + T_3v(i,j+1) &\geq w(1,i+1,j) + w(1,i,j+1) \\
&= v(i+2,j) + v(i+1,j+1) \\
&\geq v(i+1,j) + v(i+2,j+1) \\
&= w(1,i,j) + w(1,i+1,j+1) \\
&= T_3v(i,j) + T_3v(i+1,j+1)
\end{aligned}$$

where the second inequality follows from C3'.

(ii) T_3v satisfies **C4'**:

We want to prove that $T_3v(i,j) + T_3v(i,j-1) \geq T_3v(i-1,j) + T_3v(i+1,j-1)$. Let u_1 and u_2 be such that $T_3v(i-1,j) = w(u_1,i-1,j)$, $T_3v(i+1,j-1) = w(u_2,i+1,j-1)$

Case 1: $u_1 = 0, u_2 = 0$

$$\begin{aligned}
T_3v(i,j) + T_3v(i,j-1) &\geq w(0,i,j) + w(0,i,j-1) \\
&= v(i,j) + v(i,j-1) \\
&\geq v(i-1,j) + v(i+1,j-1) \\
&= w(0,i-1,j) + w(0,i+1,j-1) \\
&= T_3v(i-1,j) + T_3v(i+1,j-1)
\end{aligned}$$

where the second inequality follows from C4'.

Case 2: $u_1 = 0, u_2 = 1$

$$\begin{aligned}
T_3v(i,j) + T_3v(i,j-1) &\geq w(1,i,j) + w(0,i,j-1) \\
&= v(i+1,j) + v(i,j-1) \\
&\geq v(i-1,j) + v(i+2,j-1) \\
&= w(0,i-1,j) + w(1,i+1,j-1) \\
&= T_3v(i-1,j) + T_3v(i+1,j-1)
\end{aligned}$$

where the second inequality follows from C4' (writing C4' for (i,j) and (i+1,j), then summing two inequalities).

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_3v(i,j) + T_3v(i,j-1) &\geq w(0,i,j) + w(1,i,j-1) \\ &= v(i,j) + v(i+1,j-1) \\ &= w(1,i-1,j) + w(0,i+1,j-1) \\ &= T_3v(i-1,j) + T_3v(i+1,j-1) \end{aligned}$$

where the inequality follows from the definition of $T_3v(i, j)$, and the others follow from the definition of $w(u, i, j)$.

<u>Case 4: $u_1 = 1, u_2 = 1$</u>

$$\begin{aligned} T_3v(i,j) + T_3v(i,j-1) &\geq w(1,i,j) + w(1,i,j-1) \\ &= v(i+1,j) + v(i+1,j-1) \\ &\geq v(i,j) + v(i+2,j-1) \\ &= w(1,i-1,j) + w(1,i+1,j-1) \\ &= T_3v(i-1,j) + T_3v(i+1,j-1) \end{aligned}$$

where the second inequality follows from C4'.

(iii) T_3v satisfies **C5'**:

We want to prove that $T_3v(i, j) + T_3v(i-1, j) \geq T_3v(i, j-1) + T_3v(i-1, j+1)$. Let u_1 and u_2 be such that $T_3v(i, j-1) = w(u_1, i, j-1)$, $T_3v(i-1, j+1) = w(u_2, i-1, j+1)$

<u>Case 1: $u_1 = 0, u_2 = 0$</u>

$$\begin{aligned} T_3v(i,j) + T_3v(i-1,j) &\geq w(0,i,j) + w(0,i-1,j) \\ &= v(i,j) + v(i-1,j) \\ &\geq v(i,j-1) + v(i-1,j+1) \\ &= w(0,i,j-1) + w(0,i-1,j+1) \\ &= T_3v(i,j-1) + T_3v(i-1,j+1) \end{aligned}$$

where the second inequality follows from C5'.

<u>Case 2: $u_1 = 0, u_2 = 1$</u>

$$\begin{aligned} T_3v(i,j) + T_3v(i-1,j) &\geq w(0,i,j) + w(1,i-1,j) \\ &= v(i,j) + v(i,j) \\ &\geq v(i,j-1) + v(i,j+1) \\ &= w(0,i,j-1) + w(1,i-1,j+1) \\ &= T_3v(i,j-1) + T_3v(i-1,j+1) \end{aligned}$$

where the second inequality follows from C2'.

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_3v(i,j) + T_3v(i-1,j) &\geq w(0,i,j) + w(1,i-1,j) \\ &= v(i,j) + v(i,j) \\ &\geq v(i+1,j-1) + v(i-1,j+1) \\ &= w(1,i,j-1) + w(0,i-1,j+1) \\ &= T_3v(i,j-1) + T_3v(i-1,j+1) \end{aligned}$$

where the second inequality follows from C4' and C5' (writing C4' for (i,j+1) and C5' for (i+1,j), then summing two inequalities).

Case 4: $u_1 = 1, u_2 = 1$

$$\begin{aligned} T_3v(i,j) + T_3v(i-1,j) &\geq w(1,i,j) + w(1,i-1,j) \\ &= v(i+1,j) + v(i,j) \\ &\geq v(i+1,j-1) + v(i,j+1) \\ &= w(1,i,j-1) + w(1,i-1,j+1) \\ &= T_3v(i,j-1) + T_3v(i-1,j+1) \end{aligned}$$

where the second inequality follows from C5'.

FOURTH PART : Show $T_4v(i,j) \in V$

$$T_4v(i,j) = \max\{v(i,j+1), v(i,j)\}$$

defining w to be a function on $\{0,1\} \times Z^2$ such that

$$w(u,i,j) = u[v(i,j+1)] + (1-u)[v(i,j)]$$
$$= \begin{cases} v(i,j) & \text{if } u = 0 \\ v(i,j+1) & \text{if } u = 1 \end{cases}$$

thus

$$T_4 v(i,j) = \max_{u \in \{0,1\}} w(u,i,j)$$

(i) $T_4 v$ satisfies **C3'**:

We want to prove that $T_4v(i+1,j) + T_4v(i,j+1) \geq T_4v(i,j) + T_4v(i+1,j+1)$. Let u_1 and u_2 be such that $T_4v(i,j) = w(u_1,i,j)$, $T_4v(i+1,j+1) = w(u_2,i+1,j+1)$

<u>Case 1: $u_1 = 0, u_2 = 0$</u>

$$\begin{aligned} T_4v(i+1,j) + T_4v(i,j+1) &\geq w(0,i+1,j) + w(0,i,j+1) \\ &= v(i+1,j) + v(i,j+1) \\ &\geq v(i,j) + v(i+1,j+1) \\ &= w(0,i,j) + w(0,i+1,j+1) \\ &= T_4v(i,j) + T_4v(i+1,j+1) \end{aligned}$$

where the second inequality follows from C3'.

<u>Case 2: $u_1 = 0, u_2 = 1$</u>

$$\begin{aligned} T_4v(i+1,j) + T_4v(i,j+1) &\geq w(1,i+1,j) + w(0,i,j+1) \\ &= v(i+1,j+1) + v(i,j+1) \\ &\geq v(i,j) + v(i+1,j+2) \\ &= w(0,i,j) + w(1,i+1,j+1) \\ &= T_4v(i,j) + T_4v(i+1,j+1) \end{aligned}$$

where the second inequality follows from C2' and C3' (writing C2' for (i,j+1) and C3' for (i,j+1), then summing two inequalities).

Case 3: $u_1 = 1, u_2 = 0$

$$\begin{aligned} T_4 v(i+1, j) + T_4 v(i, j+1) &\geq w(1, i+1, j) + w(0, i, j+1) \\ &= v(i+1, j+1) + v(i, j+1) \\ &= v(i, j+1) + v(i+1, j+1) \\ &= w(1, i, j) + w(0, i+1, j+1) \\ &= T_4 v(i, j) + T_4 v(i+1, j+1) \end{aligned}$$

Case 4: $u_1 = 1, u_2 = 1$

$$\begin{aligned} T_4 v(i+1, j) + T_4 v(i, j+1) &\geq w(1, i+1, j) + w(1, i, j+1) \\ &= v(i+1, j+1) + v(i, j+2) \\ &\geq v(i, j+1) + v(i+1, j+2) \\ &= w(1, i, j) + w(1, i+1, j+1) \\ &= T_4 v(i, j) + T_4 v(i+1, j+1) \end{aligned}$$

where the second inequality follows from C3'.

(ii) $T_4 v$ satisfies **C4'**:

We want to prove that $T_4 v(i, j) + T_4 v(i, j-1) \geq T_4 v(i-1, j) + T_4 v(i+1, j-1)$. Let u_1 and u_2 be such that $T_4 v(i-1, j) = w(u_1, i-1, j)$, $T_4 v(i+1, j-1) = w(u_2, i+1, j-1)$

Case 1: $u_1 = 0, u_2 = 0$

$$\begin{aligned} T_4 v(i, j) + T_4 v(i, j-1) &\geq w(0, i, j) + w(0, i, j-1) \\ &= v(i, j) + v(i, j-1) \\ &\geq v(i-1, j) + v(i+1, j-1) \\ &= w(0, i-1, j) + w(0, i+1, j-1) \\ &= T_4 v(i-1, j) + T_4 v(i+1, j-1) \end{aligned}$$

where the second inequality follows from C4'.

Case 2: $u_1 = 0, u_2 = 1$

$$
\begin{aligned}
T_4v(i,j) + T_4v(i,j-1) &\geq w(0,i,j) + w(1,i,j-1) \\
&= v(i,j) + v(i,j) \\
&\geq v(i-1,j) + v(i+1,j) \\
&= w(0,i-1,j) + w(1,i+1,j-1) \\
&= T_4v(i-1,j) + T_4v(i+1,j-1)
\end{aligned}
$$

where the second inequality follows from C1'.

Case 3: $u_1 = 1, u_2 = 0$

$$
\begin{aligned}
T_4v(i,j) + T_4v(i,j-1) &\geq w(0,i,j) + w(1,i,j-1) \\
&= v(i,j) + v(i,j) \\
&\geq v(i-1,j+1) + v(i+1,j-1) \\
&= w(1,i-1,j) + w(0,i+1,j-1) \\
&= T_4v(i-1,j) + T_4v(i+1,j-1)
\end{aligned}
$$

where the second inequality follows from C4' and C5' (writing C4' for (i,j) and C5' for (i,j), then summing two inequalities).

Case 4: $u_1 = 1, u_2 = 1$

$$
\begin{aligned}
T_4v(i,j) + T_4v(i,j-1) &\geq w(1,i,j) + w(1,i,j-1) \\
&= v(i,j+1) + v(i,j) \\
&\geq v(i-1,j+1) + v(i+1,j) \\
&= w(1,i-1,j) + w(1,i+1,j-1) \\
&= T_4v(i-1,j) + T_4v(i+1,j-1)
\end{aligned}
$$

where the second inequality follows from C4'.

(iii) T_4v satisfies **C5'**:

We want to prove that $T_4v(i,j) + T_4v(i-1,j) \geq T_4v(i,j-1) + T_4v(i-1,j+1)$. Let u_1 and u_2 be such that $T_4v(i,j-1) = w(u_1,i,j-1)$, $T_4v(i-1,j+1) = w(u_2,i-1,j+1)$

Case 1: $u_1 = 0, u_2 = 0$

$$
\begin{aligned}
T_4v(i,j) + T_4v(i-1,j) &\geq w(0,i,j) + w(0,i-1,j) \\
&= v(i,j) + v(i-1,j) \\
&\geq v(i,j-1) + v(i-1,j+1) \\
&= w(0,i,j-1) + w(0,i-1,j+1) \\
&= T_4v(i,j-1) + T_4v(i-1,j+1)
\end{aligned}
$$

where the second inequality follows from C5'.

Case 2: $u_1 = 0, u_2 = 1$

$$
\begin{aligned}
T_4v(i,j) + T_4v(i-1,j) &\geq w(0,i,j) + w(1,i-1,j) \\
&= v(i,j) + v(i-1,j+1) \\
&\geq v(i,j-1) + v(i-1,j+2) \\
&= w(0,i,j-1) + w(1,i-1,j+1) \\
&= T_4v(i,j-1) + T_4v(i-1,j+1)
\end{aligned}
$$

where the second inequality follows from C2' and C5' (writing C2' for (i,j) and C5' for (i,j+1), then summing two inequalities).

Case 3: $u_1 = 1, u_2 = 0$

$$
\begin{aligned}
T_4v(i,j) + T_4v(i-1,j) &\geq w(0,i,j) + w(1,i-1,j) \\
&= v(i,j) + v(i-1,j+1) \\
&= v(i,j) + v(i-1,j+1) \\
&= w(1,i,j-1) + w(0,i-1,j+1) \\
&= T_4v(i,j-1) + T_4v(i-1,j+1)
\end{aligned}
$$

Case 4: $u_1 = 1, u_2 = 1$

$$
\begin{aligned}
T_4v(i,j) + T_4v(i-1,j) &\geq w(1,i,j) + w(1,i-1,j) \\
&= v(i,j+1) + v(i-1,j+1) \\
&\geq v(i,j) + v(i-1,j+2) \\
&= w(1,i,j-1) + w(1,i-1,j+1) \\
&= T_4v(i,j-1) + T_4v(i-1,j+1)
\end{aligned}
$$

where the second inequality follows from C5'.

LAST PART: Show $Tv(i,j) \in V$

$$Tv(i,j) = [-c_h(i^+ + j^+) - c_l(i^- + j^-)] + \Big[\lambda_1 T_1 v(i,j) + \lambda_2 T_2 v(i,j) + \mu_1 T_3 v(i,j) + \mu_2 T_4 v(i,j)\Big] \tag{B.4}$$

In Lemma B.0.2, the term $\lambda_1 T_1 v(i,j) + \lambda_2 T_2 v(i,j) + \mu_1 T_3 v(i,j) + \mu_2 T_4 v(i,j)$ in Equation (B.4) satisfies the conditions **C3'-C5'**. To show that $Tv(i,j) \in V$ we show that $f(i,j) = -c_h(i^+ + j^+) - c_l(i^- + j^-) \in V$, i.e., satisfies the conditions **C3'-C5'**.

(i) Tv satisfies **C3'**:

We want to prove that $f(i+1,j) + f(i,j+1) \geq f(i,j) + f(i+1,j+1)$.

$$\begin{aligned}
&f(i+1,j) + f(i,j+1) - f(i,j) - f(i+1,j+1)\\
&= \Big(-c_h((i+1)^+ + j^+) - c_l((i+1)^- + j^-)\Big) + \Big(-c_h(i^+ + (j+1)^+) - c_l(i^- + (j+1)^-)\Big)\\
&-\Big(-c_h(i^+ + j^+) - c_l(i^- + j^-)\Big)\\
&-\Big(c_h((i+1)^+ + (j+1)^+) + c_l((i+1)^- + (j+1)^-)\Big)\\
&= 0
\end{aligned}$$

(ii) Tv satisfies **C4'**:

We want to prove that $f(i,j) + f(i,j-1) \geq f(i-1,j) + f(i+1,j-1)$.

$$\begin{aligned}
&f(i,j) + f(i,j-1) - f(i-1,j) - f(i+1,j-1)\\
&= \Big(-c_h(i^+ + j^+) - c_l(i^- + j^-)\Big)\\
&+\Big(-c_h(i^+ + (j-1)^+) - c_l(i^- + (j-1)^-)\Big)\\
&-\Big(-c_h((i-1)^+ + j^+) - c_l((i-1)^- + j^-)\Big)\\
&-\Big(-c_h((i+1)^+ + (j-1)^+) - c_l((i+1)^- + (j-1)^-)\Big)\\
&= \Big(-c_h(2i^+ - (i-1)^+ - (i+1)^+) - c_l(2i^- - (i-1)^- - (i+1)^-)\Big)\\
&\geq 0
\end{aligned}$$

<u>Case 1: $i = 0$</u>

$$
\begin{aligned}
&f(i,j)+f(i,j-1)-f(i-1,j)-f(i+1,j-1)\\
&=c_h+c_l\\
&\geq 0
\end{aligned}
$$

The inequality follows from positivity of parameters c_h and c_l.

<u>Case 2: $i \neq 0$</u>

$$
\begin{aligned}
&f(i,j)+f(i,j-1)-f(i-1,j)-f(i+1,j-1)\\
&=\Big(c_h(((i+1)^+-i^+-(i^+-(i-1)^+)+c_l(((i+1)^+-i^+-(i^+-(i-1)^+)\Big)\\
&=0
\end{aligned}
$$

(iii) Tv satisfies **C5'**: We want to prove that $f(i,j)+f(i-1,j)\geq f(i,j-1)+f(i-1,j+1)$.

$$
\begin{aligned}
&f(i,j)+f(i-1,j)-f(i,j-1)-f(i-1,j+1)\\
&=\Big(-c_h(i^++j^+)-c_l(i^-+j^-)\Big)\\
&+\Big(-c_h((i-1)^++j^+)-c_l((i-1)^-+j^-)\Big)\\
&-\Big(-c_h(i^++(j-1)^+)-c_l(i^-+(j-1)^-)\Big)\\
&-\Big(-c_h((i-1)^++(j+1)^+)-c_l((i-1)^-+(j+1)^-)\Big)\\
&=\Big(c_h(((j+1)^+-j^+-(j^+-(j-1)^+)+c_l((((j+1)^+-j^+-(j^+-(j-1)^+)\Big)\\
&\geq 0
\end{aligned}
$$

<u>Case 1: $j=0$</u>

$$
\begin{aligned}
&f(i,j)+f(i-1,j)-f(i,j-1)-f(i-1,j+1)\\
&=c_h+c_l\\
&>0
\end{aligned}
$$

where the inequality follows from positivity of parameters c_h and c_l.

<u>Case 2: $j \neq 0$</u>

$$f(i, j) + f(i-1, j) - f(i, j-1) - f(i-1, j+1)$$
$$\geq 0$$

This completes the proof of Lemma B.0.2.□

Proof (Theorem 3.3.1). We use Lemma B.0.2 in the proof.

Existence and Monotonicity of $S_1(j)$ ($S_2(i)$). Condition **C1'** implies that, for a given j, if stopping production is the optimal action under i, then it is the optimal action under $i+1$. Therefore, it is possible to define $S_1(j)$ as the control-limit such that for $i \geq S_1(j)$ stopping production is the optimal action. Condition **C3'** implies that the minimum i value at which stopping production is preferred over continuing production is lower under $j+1$ compared to j. Therefore $S_1(j)$ is decreasing in j. Existence and monotonicity of $S_2(i)$ follows from conditions **C2'** and **C3'**.

Existence and Monotonicity of $K_1(j)$ ($K_2(i)$). Condition **C4'** implies that, for a given j, if rejecting the transshipment request of D2 is the optimal action (over accepting the request) under $i+1$, then rejecting the request is optimal action under i. Therefore, it is possible to define $K_1(j)$ as the control-limit such that for $i \leq K_1(j)$ rejecting the request of D2 is the optimal action. Condition **C5'** implies that the maximum i value at which rejecting the request is preferred over accepting the request is lower under $j+1$ compared to j. Therefore $K_1(j)$ is decreasing in j. Existence and monotonicity of $Z_1(j)$ follows similar lines to $K_1(j)$, while existence and monotonicity of $K_2(j)$ and $Z_2(j)$ follows from conditions **C5'** and **C4'**, respectively.□

APPENDIX C

PROOF OF PROPOSITION 3.2.3

The expression in (3.21) gives the total discounted expected profit under a given policy. Please refer to the definitions of $\mathbf{P}_{from}$ and $\mathbf{P}_{to}$ in Subsection 3.4.1. For a given parameter setting, consider a policy. If $\mathbf{P}_{from} > \mathbf{P}_{to}$ ($\mathbf{P}_{from} < \mathbf{P}_{to}$) under this policy, then profit is linear decreasing (increasing) in r.

Now, consider the profits under all possible policies including the optimal policy. The optimal profit is the maximum of all profits under a given r, i.e., optimal profit function is the upper envelope of profit functions and therefore it is convex in r. The optimal profit function is maximized at an extreme point, therefore under the given parameter setting either $r = 0$ or $r = 10$ is preferred under the optimal policy.□

APPENDIX D

INITIAL NUMERICAL STUDY SETTING AND RESULTS FOR INCOMPLETE INFORMATION MODEL

An initial numerical study is made for the partial-information model given in Section 4.3 in order to understand the system behavior. Table D.1 shows the setting of the numerical study. Full-factorial experiment is made for changing parameter values. In the experimental setting, the following parameters are fixed: the production rates at D1 and D2, the revenue obtained per customer, policy of D2 (i.e. (S_2, K_2, Z_2, T_2)), waiting cost at D1, unit production cost and fixed ordering cost, which are $\mu_1 = 1$, $\mu_2 = 1$, $S_2 = 4$, $K_2 = 0$, $Z_2 = -1$, $T_2 = -4$, $R = 10$, $c_l = 2$, $c_p = 0$ and $c_o = 0$, respectively. The number of cases solved are 432, which is simply the multiplication of the numbers of different parameter values.

Table D.1: Initial Parameter Setting for the Numerical Study for Partial-Information Model

Parameters	r	c_h	c_r	c_t	c_d	λ_1	λ_2
Values	1,5,9	0.5,2	0,2	0,1	0,2	0.5,1.0,2.0	0.5,1.0,2.0

Three different information availability levels are considered, viz. no-information, partial-information and full-information. No-information corresponds to a single state for D2's status, which is assumed to be $\widetilde{j} = 0$. Three intervals are assumed on D2's status for partial-information, which are $S_2 \geq j > K_2$, $K_2 \geq j > Z_2$ and $Z_2 \geq j \geq T_2$, which correspond to $\widetilde{j} = 2$, $\widetilde{j} = 1$ and $\widetilde{j} = 0$, respectively. Full-information denotes availability of all values of j to D1.

Solutions are obtained on a Pentium 3.0 Ghz. processor computer with 2 GB RAM using GAMS Optimization Package Version 23.0. Problem instances are solved for full-information case as explained in Section 3.4, i.e. with a single run for each instance using GAMS with CplexPar as the solver. As the models for no-information and partial-information cases are non-linear, multiple non-linear solvers are used interacting with each other. The non-linear solvers used are Baron, Snopt and MinosD. All those three solvers are used with default options, and modified options as well. Baron does not make use of initial point, therefore 2 runs are made for each instance using two different options. Snopt and MinosD are used with no initial solution, initial solution derived from full-information solution and initial solution provided from each other. Figure 4.7 explain the strategy used in non-linear solution. As a result, 40 runs are made for each instance using Snopt and MinosD. Including Baron and considering both information availability levels, 84 runs are made for each instance for non-linear solution. Best solution in terms of profit is selected among those results as the non-linear solution. Total of 36,720 runs are made for each instance, considering all information availability levels.

For comparison, the percentage deviation in profit under partial-information $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{partial}}{\overline{\Pi}_{full}}\times 100\%$ and the percentage deviation in profit under no-information $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{no}}{\overline{\Pi}_{full}}\times 100\%$ will be used, where $\overline{\Pi}_{full}$, $\overline{\Pi}_{partial}$ and $\overline{\Pi}_{no}$ denote the profits for full-information, partial-information and no-information cases. Addition to percentages, the profit deviation ratio $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{partial}}{\overline{\Pi}_{full}-\overline{\Pi}_{no}}$ is considered. It is always expected the profit deviation ratio to be positive and less than or equal to one, since full-information will perform better than or equal to partial-information and partial-information will perform better than or equal to no-information.

Starting with the averages, for the deviation percentages $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{partial}}{\overline{\Pi}_{full}}\times 100\%$ and $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{no}}{\overline{\Pi}_{full}}\times 100\%$, and the deviation ratio $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{partial}}{\overline{\Pi}_{full}-\overline{\Pi}_{no}}$, average values are 0.21%, 1.15% and 0.19, respectively. Table D.2 shows the deviation percentages and ratio defined above for the numerical study. It is clear from the table and the averages that the information effect is almost negligible under this problem setting.

No-information percentages are higher than partial-information percentages, as ex-

Table D.2: Percent Profit Deviation and Profit Deviation Ratio Results for the Initial Numerical Study for Partial-Information and No-Information

$\frac{\overline{\Pi}_{full}-\overline{\Pi}_{partial}}{\overline{\Pi}_{full}} \times 100\%$		$\frac{\overline{\Pi}_{full}-\overline{\Pi}_{no}}{\overline{\Pi}_{full}} \times 100\%$		$\frac{\overline{\Pi}_{full}-\overline{\Pi}_{partial}}{\overline{\Pi}_{full}-\overline{\Pi}_{no}}$	
Range	# of inst.	Range	# of inst.	Range	# of inst.
= 0.00%	91	= 0.00%	2	= 0.00	91
(0.00%, 0.05%]	112	(0.00%, 0.05%]	14	(0.00, 0.10]	131
(0.05%, 0.10%]	53	(0.05%, 0.10%]	11	(0.10, 0.50]	161
(0.10%, 0.50%]	125	(0.10%, 0.50%]	113	(0.50, 1.00]	45
(0.50%, 1.50%]	44	(0.50%, 1.50%]	163	= 1.00	4
(1.50%, 3.00%]	7	(1.50%, 3.00%]	116		
3.00% <	0	3.00% <	13		

pected. For the instances where percentage is zero, the decision is the same with the full-information solution and there is no randomization in decision.

The policies are scrutinized for 13 instances where $\frac{\overline{\Pi}_{full}-\overline{\Pi}_{no}}{\overline{\Pi}_{full}} \times 100\%$ is greater than 3.00% for the no-information case. In those instances, randomized decisions are more-likely expected. For production decision, randomization is not made in any instance out of 13 instances. This indicates that production decision is not affected by information on D2's status. Only 2 instances have randomized decision for lateral transshipment requests of D2. Randomization is mostly used for decision about customer arrival. In 7 instances, randomization is made between $a_1 = Accept$ and $a_1 = ltr$ or $a_1 = Reject$ and $a_1 = ltr$. From this analysis, it is concluded that in the current problem setting, information is mostly valuable for decisions made at customer arrival, but there is little effect of information on profit.

APPENDIX E

NOTATIONS USED IN THE STUDY

Notations used in the study are given below.

STATES

I : set of status of D1 ($I = \{i : i \in \mathbb{Z}\}$, where negative values stand for queue status and positive values stand for inventory status)

J : set of status of D2 ($J = \{j : j = -\infty, \ldots, S_2 - 1, S_2\}$ for models in Chapter 3 and $J = \{j : j = T_2, T_2 + 1, \ldots, S_2 - 1, S_2\}$ for models in Chapter 4, where negative values stand for queue status and positive values stand for inventory status)

S : set of all states for full information system ($S = \{s : s = (i, j) | i \in I, j \in J\}$)

$\{S_1, S_2, \ldots, S_K\}$: a given partition of the state space S

K : set of all states for partial information system for partial information on D2's status given in Subsection 4.1.2 ($K = \{k : k = (i, \widetilde{j}) | i \in I, \widetilde{j} \in \{0\}$ or $\{0, 1\}$ or $\{0, 1, 2\}\}$)

G : set of all states for partial information system for partial information on D1's status given in Subsection 4.1.3 ($G = \{g : g = (\widetilde{i}, j) | \widetilde{i} \in \{-\infty, \ldots, -1, 0, \mathbf{1}\}, j \in \mathbb{Z}\}$ for D1's inventory status case and $G = \{g : g = (\widetilde{i}, j) | \widetilde{i} \in \{\mathbf{-1}, 0, 1, \ldots, \infty\}, j \in \mathbb{Z}\}$ for D1's customer status case)

$\{X_t, t \geq 0\}$: internal process (or original process) over continuous time t for D1

$\{Y_t, t \geq 0\}$: observation process over continuous time t for D1

ACTIONS

A : set of actions ($A = \{a : a = (a_1, a_2, a_3)\}$ for models in Section 3.1 and in Chapter 4; and $A = \{a : a = (a_1, a_2, a_3, a_4)\}$ for model in Section 3.3)

For Decentralized Model given in Section 3.1:

a_1 : actions available at customer arrival ($a_1 = \{Accept, ltr\}$)

a_2 : actions available when D2 makes a lateral transshipment request ($a_2 = \{Accept, Reject\}$)

a_3 : actions available for production completion ($a_3 = \{Accept, Reject\}$)

For Centralized Model given in Section 3.1:

a_1 : actions available at customer arrival to D1 ($a_1 = \{Accept, ltr, \}$

a_2 : actions available at customer arrival to D2 ($a_1 = \{Accept, ltr, \}$

a_3 : actions available for production completion at D1 ($a_3 = \{Accept, Reject\}$)

a_4 : actions available for production completion at D2 ($a_3 = \{Accept, Reject\}$)

For models in Chapter 4:

a_1 : actions available at customer arrival ($a_1 = \{Accept, ltr, Reject\}$)

a_2 : actions available when D2 makes a lateral transshipment request ($a_2 = \{Accept, Reject\}$)

a_3 : actions available for production completion ($a_3 = \{Accept, Reject\}$)

PARAMETERS

α : discount rate

μ_1 : manufacturing rate of D1's production line

μ_2 : manufacturing rate of D2's production line

λ_1 : arrival rate of customers at D1

λ_2 : arrival rate of customers at D2

β : uniform rate which is defined as $\beta = \lambda_1 + \lambda_2 + \mu_1 + \mu_2$

τ : expected time to be in a state , $\frac{1}{\beta}$

c_l : waiting cost per customer per unit time

c_h : holding cost per item per unit time

c_t : lateral transshipment cost per item

c_r : customer rejection cost per customer demand

c_d : dealer rejection cost per lateral transshipment request

c_o : fixed ordering cost per lateral transshipment request

c_p : production cost per item per unit time

R : Net revenue per customer (also called as **revenue amount**)

r : Net revenue per lateral transshipment (also called as **payment amount**)

$r(j)$: Variable net revenue per lateral transshipment as a function of j(also called as **variable payment amount**)

π : a policy

$a^{\pi}(i, j) = (a_1, a_2, a_3)$: actions for a given state under policy π

π^* : policy which maximizes expected discounted profit

$\delta(s)$: initial probability distribution of state s

(S_2, K_2, Z_2) : D2's threshold policy for models in Chapter 3; where S_2, K_2, and Z_2 are base-stock, rationing and transshipment levels, respectively

(S_2, K_2, Z_2, T_2) : D2's threshold policy for models in Chapter 4; where S_2, K_2, Z_2 and T_2 are base-stock, rationing, transshipment and customer rejection levels, respectively

VARIABLES

$\tau U(s, a)$: total discounted joint probability of state s under action a

$X(s, a)$: the discounted fraction of visits to state s under action a

$\alpha(a|k)$: conditional probability of choosing action a given that the observation process is at state k

$\alpha(a|g)$: conditional probability of choosing action a given that the observation process is at state g

FUNCTIONS AND OPERATORS

i^+ : positive value function, where $i^+ = max\{0, i\}$

i^- : negative value function, where $i^- = max\{0, -i\}$

$I_{[z]}$: indicator function, where $I_{[z]} = \begin{cases} 1 & \text{if } z \text{ is true} \\ 0 & \text{if } z \text{ is false} \end{cases}$

$N_1(t)$: the number of customer arrivals by time t

$N_2(t)$: the number of accepted lateral transshipment requests by time t

$N_3(t)$: the number of customers served by lateral transshipment by time t

$\O_1 v(i, j)$: one period return operator for a customer arrival to D1

$\O_2 v(i, j)$: one period return operator for a customer arrival to D2

$\O_3 v(i, j)$: one period return operator for a production completion for D1

$\O_4 v(i, j)$: one period return operator for a production completion for D2

$T_1 v(i, j)$: one period return operator for a customer arrival to D1

$T_2 v(i, j)$: one period return operator for a customer arrival to D2

$T_3 v(i, j)$: one period return operator for a production completion for D1

$T_4 v(i, j)$: one period return operator for a production completion for D2

$H(i, j)$: holding cost

$L(i, j)$: customer waiting cost

R1 : expected discounted revenue by own customers

R2 : expected discounted revenue by sending parts

RLT : expected discounted revenue by lateral transshipment

$\rho(s, a)$: one-step profit at state s under action a

$p(s|s', a)$: transition probability from state s' to state s under action a

$v^{\pi}(i, j)$: expected discounted profit starting with initial state (i, j) under policy π

$v^* \equiv v^{\pi^*}$: optimal profit function

SL : service level, or the total discounted expected number of customers waiting in D1's queue

$\mathbf{P}_{from}$: total discounted fraction of time part transshipment from D2 is made

$\mathbf{P}_{to}$: total discounted fraction of time transshipment requests of D2 is met

$\overline{\Pi}_{DP}$: expected discounted profit for dynamic-pooling strategy

$\overline{\Pi}_{NP}$: expected discounted profit for no-pooling strategy

$\overline{\Pi}_{FP}$: expected discounted profit for full-pooling strategy

$\overline{\Pi}_{SP}$: expected discounted profit for static-pooling strategy

$\overline{\Pi}_{full}$: average expected profit under full-information

$\overline{\Pi}_{partial}$: average expected profit under partial-information

$\overline{\Pi}_{no}$: average expected profit under no-information

SYNONYMS

Pooling system, sharing system.

After-sales service provider, dealer, retailer, service center.

Status (of a dealer), net inventory level, inventory/queue status, inventory/queue level.

ABBREVIATIONS

D1 : Dealer 1 (policy optimizing dealer or the dealer under consideration)

D2 : Dealer 2 (dealer with a given known policy or the other dealer)

PBIA: Percent Benefit of Information Availability

VITA

Benhür Satır is from Republic of Turkey. He received his B.S., M.S. in Industrial Engineering and M.S. in Economics degrees from Middle East Technical University, Turkey, in 1999, 2003 and 2004, respectively. After receiving B.S. degree, he worked for a poultry integration from July 1999 to April 2000, in the position of Director of Production Planning and Inventory Control Department. After this work experience, he commenced his academic career and between September and December 2000, he worked as Expert at Industrial Engineering Department of Çankaya University, and since January 2001, he is Instructor at the same department. Benhür Satır has gained a wide educational experience for the last ten years. He has taught sixteen different undergraduate courses and advised nineteen different senior student projects. His research interests are combinatorial optimization, scheduling, multi-criteria optimization, mathematical modeling applications, inventory control systems, supply chain optimization and management, management information systems, microeconomics, pricing, technology policies, regional development and agricultural economics.

Benhür Satır has close relations with the industry, especially with SMEs in Organized Industrial Zones in Ankara Region (OSTİM, İVEDİK and SİNCAN). He is actively involved in İŞİM Clustering Study in Work and Construction Machines Industry (`http://www.isim.org.tr/`) for three years, since the beginning. He is the contact person in Çankaya University of two European Union Funded Transfer of Innovation Projects related with İŞİM Clustering Study. He prepared and conducted professional training programs ("Clustering and Collaboration", OSTİM, 2010 and "Cost Optimization in Steel Industry", ERDEMİR, 2006). He is currently a consultant in a collaboration project of OSTİM companies on a new product design and production.

Benhür Satır received Honor Award from President of Turkish Republic, Süleyman Demirel, in 1994 and Special Award from Governor of Isparta, Ertuğrul Dokuzoğlu, in 1995 for his successes in National Physics Olympics and University Entrance Exam, respectively.

Benhür Satır is a referee of Information World Journal (Bilgi Dünyası , ISSN: 1302-3217). He is a member of Operations Research Association of Turkey (Yöneylem Araştırma Derneği), Applied Probability Society, Work and Construction Machines Clustering Association (İş ve İnşaat Makinaları Kümelenme Derneği). He holds the title “Brand Attorney”.

Benhür Satır is married with Ferda Satır and has a year-old daughter, Ayşe Râna Satır. For a detailed Vita, please contact with him using the following e-mail address: benhursatir@yahoo.com.