The Hong Kong Polytechnic University
Department of Logistics and Maritime Studies

Matching Demand and Supply of Short Life-Cycle Products by Trading Capacity Futures

HUNG Yick Hin

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

June 2011

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

____

HUNG, Yick Hin

Abstract

The mismatching risk of the newsvendor due to demand uncertainty is a fundamental issue in inventory research. Risk pooling is a potent strategy to reduce the underlying demand uncertainty through aggregation. However, past research has been confined to investigating risk pooling among parties within a single supply chain network. We argue that the risk can be pooled and shared among different supply chains, and it can also be transferred to the public via financial derivatives, provided that suppliers have short lead-time capacities that allow retailers to replenish stocks within the season. We treat such reserved capacity (super capacity) as a commodity that can be traded as futures to retailers and speculators.

Consider a sub-industry of a certain family of short life-cycle products in which a group of suppliers have comparable production capabilities to produce goods for their "newsvendor-type" of retailers, who sell non-identical products in the market. Under the framework of a two-stage inventory model, the retailers buy physical goods and super capacity futures as inventory portfolios in the first stage. After demand realization in the second stage, the retailers make replenishment decisions, which are limited to the capacity futures on hand. However, the retailers are allowed to form coalitions to transfer the residual capacity futures among themselves. Therefore, the retailers can make bidirectional adjustments to their inventory positions. This mechanism also helps improve supply flexibility and increases the utilization of suppliers' reserved capacity.

The dissertation consists of three parts. First, we examine a case involving only two supply chains that are engaged in a co-opetition game. We compare the two scenarios in which super capacity futures can and cannot be

exchanged between the two supply chains in stage two, and prove that Pareto improvement can be obtained.

Second, we extend the model to a group of n retailers and m suppliers to form a sub-industry. We employ a biform game to analyze the risks and payoffs to the retailers as players in both the non-cooperative (first) and cooperative (second) stages. Our findings reveal that the retailers can improve their payoffs by sharing risk among different supply chains.

Third, we allow the game of trading of super capacity to include speculators from both the sub-industry and the public. We argue that to hedge against risk, the retailers can further share and transfer their risks to the speculators by means of trading super capacity as futures or as options in futures. Our results show that the whole sub-industry is better off with super capacity trading even with the presence of outside speculators.

In this thesis we also develop a time-based, value-adding capacity measurement model, which is an output-orientated input measure for super capacity trading among different supply chains involving various products.

Our study establishes that trading super capacity futures is an efficient mechanism for individual newsvendors to improve their performance in matching demand with supply by combining operational and financial hedging strategies to reduce and share the mismatching risk that is caused by demand uncertainty within a sub-industry and with the public.

Preface

This dissertation denotes a very important milestone to me. It makes my three-decade dream come true. In fact, it does not only accomplish my personal dream of pursuing a doctoral degree, but it might also shed light on a very challenging industrial problem that I have been struggling with for over thirty years in my career.

I joined the textile and garment industry in the late 1970's and attained some important results dealing with different problems in operations management and logistics, locally as well as globally. Nevertheless, matching supply with demand was always a problem that I could not solve.

The industry has experienced great pressure due to demand-supply matching problems since the late 1990's, especially after the 911 terrorist attack in the U.S.A that has caused a serious sales drop in that early winter. Over the next several years, the giant retailers believed that the problem could be solved if lead-time was shortened. As one who had played an active and assiduous role with hundreds of suppliers in Asia, I spent a great deal of effort in the early and mid 2000's working with teams to motivate and help selected firms in various supply chains to reduce lead-times that were short enough for retailers. Surprisingly, after successfully implementing the new operations in a few months to a year, the suppliers declined to provide such short lead-time delivery no matter how good the achievements were. They could not give a precise reason, but merely knew their benefits would be affected if they were continuously offering shortened lead-time delivery. Besides, due to the demand uncertainty, the industry did not know how to arrive at a formula to compensate suppliers for the short lead-time performance. Therefore, when I was applying my doctorate in 2006, I decided to seek an academic solution for the mismatching problem.

I am glad that the literature offers abundant and flourishing knowledge enabling me to thoroughly research such a genuine mystery. The abstract and sophisticated thoughts of many scholars have provided me with a fresh view, and also various advanced methodologies to detect the problem. Indeed, the strategies argued in this thesis contain several vital ingredients from my working experience. My exposure to the manifold activities of a huge group of retailers and suppliers in different regions has provided me with a good foundation to structure a sub-industry co-opetition game. My personal involvement in solving numerous dilemmas while developing short lead-time delivery in different organizations has allowed me to appreciate the value of managing short lead-time capacity. My industrial experience in the quota business in Hong Kong has given me a blueprint of capacity trading. The time-based quotation that I learned in Mexico while setting up a factory there stimulated me to conceive a capacity trading unit. Consequently, the thesis might be a synthesis of industry practice and academic theory.

I hope this dissertation may contribute to the community, no matter how small.

Yick-Hin HUNG

Acknowledgements

I am deeply grateful for the extraordinary guidance provided by Dr Leon Y. O. Li, my doctoral program chief supervisor, and Professor Edwin T. C. Cheng, my co-supervisor. Their productive advice and close support were very helpful throughout the program.

Dr Li has played a role not only as a supervisor, but also as a mentor. He was always available to discuss problems with me and answer my questions. He offered me special assistance in my first year to transform my mindset from industry to academia. His critical, detailed and patient guidance helped me overcome so many problems in my research. Moreover, his challenges motivated me to go deeper and further in the study.

Professor Cheng has also brought me to a new perspective to change my thought and to realize the nature of research at the initial stage. He was continuously enthusiastic about my work. His invaluable feedback contributed significantly to this thesis. I am very impressed by his speedy responses in giving me comments and advice on my output in an 'extremely short lead-time' manner.

I truly thank the academic and supporting staff members in my Faculty and Department who gave me a lot of knowledge and support, and had a great impact on my study. I would also like to give special thanks to all staff of Pao You-Kong Library for their professional and friendly assistance. Their excellent operations and world-class service allowed me to complete my progress of study smoothly.

In addition, I would like to express my appreciation to Tana Netting Co., Ltd. for providing me with their production and cost data. I also want to thank a lot of my friends for their moral support and encouragement, in particular, Albbie Liu, Brenda Fung, Chabi Liu, Charlie Chan, Decem Leung, Frederick Ng, Gary Lai, George Lam, Iris Wong, Kelvin Chen, Ken Wong, Mandy Wan, Maria Fung, Mun Whye Lee, Raymond Yeung and Tommy Ho, who consistently arranged big group lunches every season in the last four years to energize me in my study. Their consideration cheered my lonely research student life.

Lastly, and most importantly, I thank my wife Yan-Yu, who has accompanied me and helped me for almost thirty years to achieve my doctoral dream; words cannot convey the depth of my gratitude to her constant love, support and listening.

Table of Contents

List of Figures

List of Tables

List of Symbols

Chapter 3:

i	supplier index, $i=1$
	retailer index, $i=2$, the retailer that has *insufficient* super capacity in stage 2
	retailer index, $i=3$, the retailer that has *excess* super capacity in stage 2,
y	random variable for demand,
D_i	stochastic demand of retailer *i*,
$f(\cdot)$	probability density function of demand,
$F(\cdot)$	cumulative probability distribution function of demand,
$\overline{F}(\cdot)$	complementary cumulative probability distribution function of demand,
q_i	order quantities of retailers *i* in stage 1,
ω_i	retailer i 's wholesale price,
r_i	retailer i 's retail price,
g_i	cost of lost sales for retailer *i*,
g_i^S	cost of lost sales for the supplier for not meeting retailer *i*'s order,
v_i	net salvage value of unsold inventory for retailer *i*,
h	price of super capacity,
α	retailer *2*'s super capacity to lost sales ratio,
β	retailer *3*'s *excess* super capacity to lost sales ratio,
c_i	supplier's cost of producing retailer *i*'s orders,

s supplier's cost of setting up super capacity that is included in c_i,

l_i lost sales of retailers i,

$\pi_i(\cdot)$ retailer i's profit,

$\pi_1^T(\cdot)$ supplier's profit if trading of super capacity between retailers is allowed,

$\pi_1^N(\cdot)$ supplier's profit if trading of super capacity between retailers is not allowed.

Chapter 4:

i supplier index, $i = 1,...,m$

retailer index, $i = m+1,...,m+n$,

y random variable for demand,

D_i stochastic demand of retailer i,

$f(\cdot)$ probability density function of demand,

$F(\cdot)$ cumulative probability distribution function of demand,

$\overline{F}(\cdot)$ complementary cumulative probability distribution function of demand,

q_i^P physical inventory quantity of player i,

q_i^C super capacity quantity of player i,

q_i total inventory position on hand of player i,

ζ_i updated demand forecast of retailer i,

δ_{ij} substitution rate of the excess demand of retailer i by retailer j,

p_i inventory unit price of retailer i,

c_{ik} inventory unit cost of supplier, $i = 1,...,m$ represents the supplier and $k = m+1,...,m+n$ represents the retailer,

ω_{ik} inventory unit wholesale price of supplier, $i = 1,...,m$ represents the supplier and $k = m+1,...,m+n$ represents the retailer,

g_i goodwill penalty cost of retailer i,

h unit price of super capacity futures in stage 1,

h'	unit price of super capacity in stage 2,
$\bar{h}$	weighted average price of super capacity with different financial instruments after option exercise has been determined,
s_i	cost of supplier *i* building super capacity,
N	set of retail players,
S	set of coalition of retailers,
Q	set of strategies,
ε	confidence index,
v	a characteristic function that maps from Q to the set of maps from 2^N to $\Re$,
γ^S	super capacity residual in stage 2,
Γ^S	a collection of possible residual super capacity vector,
M^S	a reallocation matrix of γ^S,
m_{ij}^S	the volume of super capacity to be transferred from retailer *i* to retailer *j* in *S*.

Chapter 5:

i	supplier index, $i = 1,..., m$
	retailer index, $i = m+1,..., m+n$
	speculator index, $i = m+n+1,..., m+n+\eta$,
y	random variable for demand,
D_i	stochastic demand of retailer *i*,
$f(\cdot)$	probability density function of demand,
$F(\cdot)$	cumulative probability distribution function of demand,
$\overline{F}(\cdot)$	complementary cumulative probability distribution function of demand,
q_i^P	physical inventory quantity of player *i*,
q_i^C	super capacity quantity of player *i*,
q_i	total inventory position on hand of player *i*,
ζ_i	updated demand forecast of retailer *i*,

p_i	inventory unit price of retailer i,
c_{ik}	inventory unit cost of supplier, $i = 1,...,m$ represents the supplier and $k = m+1,...,m+n$ represents the retailer,
ω_{ik}	inventory unit wholesale price of supplier, $i = 1,...,m$ represents the supplier and $k = m+1,...,m+n$ represents the retailer,
g_i	goodwill penalty cost of retailer i,
v_i	salvage value of leftover inventory of retailer i,
h	unit price of super capacity futures in stage 1,
$\bar{h}$	weighted average price of super capacity with different financial instruments after option exercise has been determined,
Λ	set of strategies,
$b(\cdot)$	benefit function,
R	the Arrow-Pratt risk premium,
Π^H	total payoffs of hedgers from super capacity,
Π^T	total payoffs of all players from super capacity.

Chapter 6:

P	production function,
p	production process,
ϕ	the standard basis for the space of piecewise linear function,
l	step size between two successive knots,
$V(\cdot)$	input set,
x	resources input other than capital,
K	capital input,
$h(\cdot)$	capacity function,
$h(K^t)$	available capacity,
$h(K^{t^*})$	value-added capacity,
η	rate of damaged capacity,
q	output volume.

List of Glossary

Aggregative game – A game that the strategies of players can be aggregated in an additive way and the payoff of each player is a function of the player's own actions.

Biform game – A hybrid non-cooperative/cooperative game designed to model business interactions.

Capacity – The maximum quantity of output per unit in a given amount of time that a stock of plant and equipment is capable to complete, provided that the availability of variable factors is not restricted.

Capacity deployment – In this thesis, manufacturing capacity is deployed into the following categories:

Rated capacity – The maximum theoretical capacity of a facility assuming it is running everyday and every hour. Rated capacity is formed by unavailable and available capacities.

Unavailable capacity – Capacity that is non-saleable, or those could not utilized due to scheduled holidays and maintenance, or unscheduled material/labour shortage.

Available capacity – Capacity that can be deployed into inoperative and operative capacities.

Inoperative capacity – Capacity that is occupied by production line setting up and developing, and short idle time caused by training, absenteeism or breakdown.

Operative capacity – Capacity that is combined with value-added and non-value-added capacities, and might contain negative value-added capacity.

Value-added capacity – Capacity unit for capacity trading. It is used to produce output and is measured by the value-added timing of a product.

Non-value-added capacity – Capacity that is processing non-value-added work.

Negative value-added capacity – Capacity that produces damaged output.

Capacity residual – Unutilized reserved capacity of retailers or reserved capacity that is hold by speculators in stage 2.

Co-opetition game – A business game that players compete and cooperate with each other at different points along the horizon.

Confidence index – The preference of players in a biform game that is influenced by private market information, commodity futures price, etc.

Core – In a cooperation game, the set of imputations under which no coalition has a value greater than the sum of its members' payoffs.

Futures – A standard financial contract obligating buyers (sellers) to purchase (sell) an asset at a predetermined future date and price.

Hedger – A player that uses financial instrument or other means to reduce risk of an asset or profit. A hedger will make or take delivery of the futures market position unless it suffers from inaccurate forecasting.

Hindrance – Any cause that prevents the elimination of a process or activity that is non-value-added work.

Inventory risk – Total loss of leftover inventory of a retailer in the selling season.

Mismatching cost – Sum of loss due to excess inventory and excess demand that is caused by demand uncertainty.

Options on futures – An option on a futures contract gives the holder the right, but not the obligation, to buy or sell the underlying asset, e.g. super capacity, in the futures contract at the strike price on the future date or at any time prior to a specific date that depends on different kinds of contracts.

Pareto-improvement – Any improvement of a certain player in the game that will not cause any other players worse off.

Postponement – A strategy that the manufacturing process of product is delayed to permit a retailer to place a replenishment order with a supplier in the selling season.

Quick response – A strategy that shortens the production period to allow replenishment to happen in the selling season. It also describes a short lead-time production.

Risk pooling – A strategy to reduce demand variability through aggregation.

Sub-industry – A group of retailers and suppliers that sell products produced by similar facilities and capabilities within a cluster. Therefore, their capacity may become a commodity.

Speculator – A seller of option contracts and/or one merely offsets the futures positions at some point before date set for futures delivery. A speculator intends to make profit from the trade of a commodity or financial instruments.

Stage 1 – The period that is before the start of the selling season.

Stage 2 – The period that is during the selling season.

Super capacity – Reserved capacity with a short lead-time and is a commodity in a sub-industry. It is traded separately from physical inventory.

Supply chain network – An interconnection of firms that related to each other. Each firm provides value through upstream and downstream linkages to produce goods or services to the ultimate customers.

Supply risk – The sum of the lost profit due to the lost sales of a retailer in the selling season.

Transshipment – The shipment of physical goods or products between different facilities or retailers at the same level in the supply chain to meet immediate demand need.

Value-added work – i) Positive – work that creates value as perceived by a customer;

ii) Negative – work that decreases the accumulated value of the process.

Waste – Work that has been eliminated following the introduction of a change in the process.

1 Introduction

The management of inventory and supply risks of newsvendors due to demand uncertainty is a fundamental issue in inventory literature. Since it was first addressed by Arrow, Harris, and Marschak (Arrow *et al*. 1951), optimal inventory policy for uncertainty has been studied and extended by a great many researchers. Among such research, a study of interest to us is that of Jain and Silver (1995) who propose a postponement strategy to use a reserved capacity option to allow "newsvendor-type" retailers to replenish short life-cycle inventory during the selling season. This single-period two-stage model provides an opportunity for a retailer to correct its inventory position according to updated forecast, and therefore, the problem of supply-demand mismatching can be alleviated if the option of placing additional orders is available. This strategy is favourable for the retailer in terms of adjusting inventory level during the selling season to match demand closely. However, it might spoil the supplier's benefit because part of the mismatching risk is shifted along a single supply chain from downstream to upstream (Burnetas and Ritchken 2005; Donohue 2000; Wu 2005). But could the demand uncertainty be mitigated by pooling the risk among different supply chains within a sub-industry in a manner of co-opetition, and even transfer the mismatching risk to the public in order to protect the benefits of both retailer and supplier?

1.1 Background

The key characteristic of the newsvendor problem is how to determine the quantity of a single order before observing demand that will maximize profit for the entire selling period. For a two-stage problem with replenishment, the order quantity at the beginning of stage 1 and the additional order quantity at the beginning of stage 2 that maximizes expected profit is the primary problem in the literature (Lau and Lau 1999a). The problem is particularly important for items with significant demand uncertainty and large inventory and lost sales costs, such as fashions, seasonal products and new electronic devices. Frazier (1986) estimated the inventory carrying cost, shortage, and excess supply for the U. S. apparel industry was 25% of annual retail sales. In fact, for fashions items, the lost sales can only be as high as 18 – 20% of the total inventory (Hunter *et al.* 1996; Mattila *et al.* 2002). Therefore, it is worthwhile to seek a better strategy for the industry to minimize the mismatching costs that arise from demand uncertainty.

1.1.1 Mismatching cost

Mismatching losses engendered by demand uncertainty can be expressed in a profit function. Let q be the inventory quantity ordered before demand is realized, p denotes the price and ω the cost; therefore the profit is given by $p\min(D,q)-\omega q$, where $D \geq 0$ is the single-period random demand with mean $\mu = E[D]$. Newsvendor faces its own demand function with a probability density function $f(y)$ and a cumulative distribution function $F(y)$, such that the expected profit is

$$\pi(q) = E[\min(D,q) - \omega q].$$

By using the fact that $\min(D,q) = D-(D-q)^+$ and $(z)^+ = \max(z,0)$, assuming leftover inventory has salvage value v per unit and the cost of lost sales is g per unit, we can rewrite the expected profit as

$$\pi(q) = (p-\omega)\mu - L(q) \tag{1.1}$$

where $L(q) = (p+g-\omega)E(D-q)^+ + (\omega - v)E(q-D)^+ \geq 0$ is the expected loss function of the mismatching cost which is caused by the demand uncertainty. $L(q)$ is also called the expected single-period holding and shortage cost function (Porteus 2002). It is the sum of the underage cost that is incurred for each unit of lost sales and the overage cost that is incurred for each unit of leftover inventory. In the context of the newsvendor model, the mismatch risk is the sum of the lost profit due to lost sales (supply risk) and the total loss on leftover inventory (inventory risk). The mismatching cost of the newsvendor problem is therefore defined as the sum of loss due to excess inventory and excess demand. If we minimize $L(q)$, then we might obtain the maximized profit in (1.1) since its first term is a constant.

The traditional approach to minimize $L(q)$ is to find the optimal order quantity by the first order conditions since $L(q)$ is a convex function. We can obtain the classic newsvendor critical fractile solution by this approach to identify an optimal inventory level,

$$F(q^*) = \frac{p+g-\omega}{p+g-v}. \tag{1.2}$$

However, the optimum is achieved by trading off inventory leftover and lost-sales costs. It neither reduces nor avoids the risks.

1.1.2 Value of reserved capacity

As a matter of fact, $L(q)$ can be further minimized by closing the gap between D and q before finding the optimal inventory quantity. The two-stage model allows forecast update in the early season providing an opportunity for a retailer to select an adjusting inventory position by replenishment. In the apparel industry, the implementation of Quick Response (QR) is a strategy that shortens the production period to allow replenishment to happen in the season. QR is a concept first developed in 1986 by Kurt Salmon Associates, a consultancy firm in the U.S.A. to respond to competition from overseas suppliers, especially from Asia, to shorten the lead-time of supply in the apparel industry (Lowson *et al.* 1999; Hines 2004). Hunter (1990) defines QR as an operational philosophy and a set of procedures aimed at maximizing the profitability of the apparel pipeline. Wu (2005) defines QR as a response to the risk of delays in the supply chain's commitment to quantity and to the need for a shorter production lead-time. It is common now to use the term QR to describe a short lead-time production.

The literature has confirmed that QR is an effective strategy to save significant operation costs and improve the service level. Hunter *et al.* (1996) compare the performance of traditional and QR procedures for seasonal and fashion apparel. They have found that a 20-week life-cycle product can drop the lost sales from 18.3% to 4% and the leftover inventory reduces to 1.2% from 7.4% if more frequency replenishment is allowed during the selling season. Mattila *et al*. (2002) also have similar results in studying the retail performance for seasonal fashion in Finland. They have found that the lost sales are 20% for a traditional company but only 3% for QR supply chain. The sell-through rate before marking down is 65.6% in a traditional company and 91.7% with replenishment sourcing. Fisher and Raman (1996) develop an optimal production commitment model for a fashion skiwear company to respond to the early sales by QR replenishment that increases 60% in the profit. Therefore, QR suppliers can ask for a higher price than traditional suppliers because they allow the customer to reorder based on observed demand to improve service levels and inventory turns (Pinnow and King 1997). In fact, saving of in-season replenishment may justify paying from 30-50% more to a QR supplier (Gilreath *et al*. 1995). However, there is a lack of mechanism to allow efficiently splitting the extra profit between supplier and retailer that is merited by the replenishment strategy.

In the two-stage newsvendor problem, the supplier's reserved capacity for order replenishment is normally treated as a commitment made exclusively between the supplier and a particular retailer. If any reserved capacity is not used, it is wasted. This capacity is seldom decoupled from its product and treated as a commodity, which can be traded independently in the market. Nevertheless, capacity is the maximum quantity of output per unit in a given amount of time that the stock of a plant and equipment of the supplier is capable of completing (De Leeuw 1962). From a business perspective, a type of capacity that needs a shorter lead-time has a higher value than another type of capacity that requires a much longer lead-time if both types of capacity produce similar products of the same quality. This is particularly true for the case of short life-cycle products with stochastic demand. There are two main root causes for the single-period inventory problem; namely (i) it is very difficult to forecast stochastic demand accurately before the selling season, and (ii) the supplier does not respond to demand immediately. If either of these root causes can be eliminated, it will be much easier for the supplier to cope with the stochastic demand.

Unfortunately it is hard to obtain an accurate forecast of demand before the start of the selling season. Hunter and Valentino (1995) comment that the demand for fashion products is almost, by definition, impossible to forecast. Fisher and Rajaram (2000) find that forecasts in fashion retailing have an average error rate of 50% or more. However, an early season observation as small as 20% of sales can provide a much better forecast for the demand distribution over the whole season (Fisher and Raman 1996). In fact, in a multi-stage forecast, as the forecast horizon shortens, the forecast variance decreases and forecast precision increases (Sethi *et al.* 2005). Therefore, production capacity has its own value if its lead-time is short enough to allow the replenishment of goods in the post-early season after early sales information has been obtained.

On the other hand, the supplier faces several issues in offering short lead-time capacity to retailers. First, the supplier may need stronger management skills in general to overcome lead-time problems. Second, retailers may inflate their demand forecasts. Third, the supplier bears the risk that production capacity may not be fully utilized in the selling season since the impact of demand uncertainty is shifted from retailers to the supplier. If these issues are to be addressed, an incentive scheme needs to be devised to compensate the supplier for its investment, and a mechanism put in place to help the supplier to manage its risks. We argue and advocate that capacity is a product in itself, which has its own market value if its special characteristics are able to satisfy customer needs (Kotler *et al.* 2007). Thus, it may be possible to reduce the demand variability inherent in different supply chains by pooling the reserved capacity of the chains concerned by market force.

1.1.3 Risk pooling

It is well known that risk pooling is one of the most powerful tools available to address variability in the supply chain and is most effective when demands are negatively correlated (Simchi-Levi *et al.* 2008). Risk pooling suggests that demand variability is reduced if demand is aggregated across locations. Capacity is one of the risks that can be pooled, and this can benefit the supply chain (Aviv and Federgruen 1999). However, past research has been confined to investigating risk pooling among parties within a single supply chain network. We consider that the common super capacity in a sub-industry becomes a commodity in itself among a group of retailers. The retailer then uses short lead-time capacity as an alternative inventory instead of physical products.

Motivated by the need to address these issues, particularly in fashion items, we propose a new strategy based on the concepts of risk pooling and postponement to tackle the single-period inventory problem. We aim to help match supply with demand more effectively. Our research does not in itself investigate the allocation of inventory risk, capacity risk, and lost sales among different players in a single supply chain. Instead, we focus on the development of a mechanism to mitigate such risks and redistribute them among a group of supply chains as well as the public in both non-cooperative and cooperative ways.

It is a common practice that one manufacturer supplies different products to various retailers and retailers buy goods from different suppliers. Under such circumstances, production capacity becomes a commodity in itself among a group of retailers and, accordingly, a sub-industry of those using the same capacity is formed. It is the same as if a supplier has a number of customers, and can reduce its overall risk by pooling the capacity for all its customers (Jin and Wu 2007). A sub-industry is therefore defined as a group of retailers and suppliers who sell products that are produced by similar facilities and capabilities within the cluster. It is distinguished from the supply chain network, which is defined as ‘an interconnection of organizations which relate to each other through upstream and downstream linkages between the different processes and activities that produce value in the form of products and services to the ultimate consumer'(Christopher 2005). In a sub-industry, different supply chains could run their own businesses independently without any direct interaction.

1.1.4 Trading capacity as commodity

We, therefore propose that capacity with a short lead-time is a commodity that can be sold independently as futures before it expires. Whenever this capacity is traded separately from physical inventory, we call it "super capacity". In other words, the source of replenishment of a retailer can also come from the capacity residual of other retailers by trading super capacity during the season to play a co-opetition game. Players in a co-opetition game will both compete and cooperate with each other in different points along the horizon. In the traditional view, cooperation and competition among retailers are opposed and mutually exclusive activities. However, it appears that companies can compete and cooperate with each other simultaneously in practice. In academic studies, co-opetition is not new and has been examined for more than two decades (Walley 2007). Brandenburger & Nalebuff (1996) explain that business is a cooperation when it comes to creating a pie and a competition when it comes to dividing it up because business is a game, but unlike sport, poker or chess, which must result in a single winner and losers. Business nature does not require others to fail in order to allow only one participant's success; it allows for multiple winners.

Some scholars use different names to describe the reserved capacity for responding to demand after early sales have been observed. For example, Cachon and Terwiesch (2009) and Cattani *et al.* (2008) adopt 'reactive capacity' to describe the capacity in which reacting to the demand is known with more certainty in the selling season. However, the capacity they describe is only assigned to the contracted retailer, while super capacity can be transferred freely as a commodity in a sub-industry. This is the main distinction between other definitions and our own.

Super capacity might bring many advantages if it is traded as futures. Both the supplier and the retailer can hedge against their risks by trading capacity futures before the selling season. The retailer can determine how much of the super capacity on hand should be converted into products at a very late stage, after reviewing early season or post-early season sales. This means that mismatching between supply and demand can be dramatically reduced because the late forecast, which takes into account more information about variations in demand, is more accurate (Fisher and Raman 1996). Thus a higher return on operating assets can be achieved by adding capacity and lowering physical inventories (Bradley and Arntzen 1999). However, studies examining the issues concerning the trading of reserved capacity among supply chains are very limited in the literature.

In this thesis we argue that mismatching risk can be pooled and shared among different supply chains and can also be transferred to public speculators via financial instruments, provided that suppliers' short lead-time capacities allow retailers to replenish in the season. Retailers can form a coalition to exchange the super capacity residual after realization of the demand is observed in the selling season. Thus the imbalance between aggregate demand and aggregate supply of the sub-industry is improved, and the mismatching cost of the sub-industry is also mitigated. Furthermore, the financial instruments also allow players to hedge and avoid capacity quantity and price risks by trading super capacity futures. Obviously, the commodity futures market depends not only on the existence of uncertainty about the future, but also on the existence of different preferences regarding risk and return or different utility functions that support transaction to take place between players (Gollier 2001).

The purposes of the present research are to study whether a unique best response strategy exists for the game players to pool, hedge and transfer mismatching risk to a sub-industry, and whether the whole sub-industry is better off even with outside speculators participating in the game. This study tries to partially fill the gap by examining the coordination among supply chains and the involvement of speculators in the super capacity futures game.

Moreover, there should be a capacity unit that is common to the mix of products encountered in different supply chains to allow super capacity trading to be processed. In addition, we also develop a time-based, value-adding capacity measurement model, which is an output-orientated input measure for super capacity trading among different supply chains involving various products.

1.2 Research Methodologies

We consider a group of $m \geq 1$ suppliers and $n \geq 2$ retailers to form a sub-industry with n supply chains. Each supply chain sells a different short life-cycle product to the market that has only one retailer but one or more suppliers. The m suppliers have the same facilities and similar capabilities to produce different goods for some of the n retailers with a very short lead-time. The retailers place orders and receive goods before the selling season (stage 1). We assume that all suppliers do not keep an inventory and will deliver the amounts that are requested by any retailers in the game under forced compliance (Cachon and Lariviere 2001), and that a minimum order is not requested by suppliers. Since the lead-time is short, the retailer can order additional goods after demand is realized in the beginning of the selling season (stage 2) if they have reserved super capacity on hand. Super capacity is trading before the season as futures. Players of the super capacity futures game are basically formed by three types of investors: namely the supplier i, $i = 1,...,m$; the retailer i, $i = m+1,...,m+n$; and the speculator i, $i = m+n+1,...,m+n+\eta$. We define all retailers and suppliers as hedgers. Hedgers intend to make or take delivery of the futures market position, unless they suffer from inaccurate forecasting that the residual part of the future position will be liquidated at some time prior to expiration. Speculators merely offset their positions at some point before the date set for the futures delivery. If a retailer or a supplier plays both roles, as hedger and speculator simultaneously, it is treated as two players in the game. However, if a player changes role between a hedger and a speculator along the way, in our analysis we still treat this investor as only one player at one point in this game. In case a retailer and a supplier have a common decision maker and the supplier does not offer any products to other retailers in stage 1, the condition for them to participate in the game is whether they are willing to exchange capacity residual in stage 2. We assume that all the investors do not have cash flow pressure to liquidate the capacity futures at any time. Moreover, there are zero transaction costs and no institutional restrictions on trades.

The mechanism developed in this thesis combines operational and financial hedging strategies. We use only futures but not forwards in this study because a futures contract is a standardized agreement that is much easier to trade and transfer in the market at any time until it matures. A forward contract should be held until the end of the contract term, since it is non-standardized and is developed by the two parties involved in the deal. The same volume of super capacity, however, may be traded several times before and during the selling season in order to adjust the retailer's inventory position and maximize its profit. The price paid for super capacity may in fact be a more accurate indication of a retailer's true situation and can thus provide a better basis for decisions by the sub-industry. Therefore, the mechanism suggested would discourage retailers from inflating their orders in an effort to gain a better allotment of inventory, since a retailer would face some risk if the volume of super capacity held is too great.

This study is divided into three phases that have different set up for the models. In Phase one, only one supplier and two retailers are considered. We have a group of $m \geq 1$ suppliers and $n \geq 2$ retailers but no speculators in Phase two. For Phase three, in addition to $m \geq 1$ suppliers and $n \geq 2$ retailers, $\eta \geq 1$ speculators are involved in the study.

1.2.1 Phase one – $m = 1$, $n = 2$ and $\eta = 0$

It is common in the literature to study the decentralized system of a single-period problem by comparing the centralized system as a benchmark to argue any improvement in the new strategy. In Phase one, we compare two scenarios, namely one in which super capacity is allowed to be exchanged, and the other where it is not allowed to be exchanged between two supply chains. We examine whether Pareto-improvement can be attained and the impact of the price of super capacity on both retailers and suppliers through this new mechanism of super capacity trading.

1.2.2 Phase two – $m \geq 1$, $n \geq 2$ and $\eta = 0$

In Phase two, the super capacity futures can be traded at any time before the order delivery date among the players in the sub-industry to regulate their inventory positions. We use a biform game to analyze the mechanism for n competitive retailers and identify the Nash equilibrium in the first stage and the core of the game in the second stage.

A biform game is a hybrid non-cooperative/cooperative game model designed to model business interactions as first proposed by Brandenburger and Stuart (2007). They study the players making strategic investments in the first stage and then playing a cooperative game determined by their investments in the second stage to favourably shape the competitive environment. There are only limited studies adopting this methodology in operations management; for example, Stuart (2005), Plambeck and Taylor (2005, 2007).

To define a biform game, consider a set of players (retailers) $N = \{m+1,...,m+n\}$ and a finite set Q_i of strategies, $i = m+1,...,m+n$, for each player i. We also have a subset of N, denoted by S and called a coalition. In the first stage, players make decisions among their strategies; this game can be analyzed like any other non-cooperative games. In the second stage, various coalitions S are formed to reach a common objective with the players who are in the same coalition. Competition is then modelled by a cooperative game in which the characteristic value function depends on the chosen actions. Let $Q = Q_{m+1} \times ... \times Q_{m+n}$, with typical element q, and let v be a map from Q to the set of maps from 2^N to $\Re$, where $v(q)(\phi) = 0$ for every $q \in Q$. For each player i, let ε_i be i's confidence index and it is a number in [0, 1]. An N-newsvendor biform game is then a collection $(Q_{m+1},...,Q_{m+n};v;\varepsilon_{m+1},...,\varepsilon_{m+n})$.

We determine that in the first stage the game has a unique Nash equilibrium by using the contraction mapping principle, in which the best response mapping is contracted globally to a fixed point on the whole strategy space. The contraction mapping argument is the most frequently used argument in the literature (Cachon and Netessine 2003).

Next we show that the core of the game (N,v) is non-empty in the second stage by using the result of Shapley (1967) that the balanced set $\{S_{m+1},...,S_{m+k}\}$ of coalitions of N with balancing weights $\kappa_{m+1},...,\kappa_{m+k}$ will give $\sum_{j=m+1}^{m+k}\kappa_j v(S_j) \le \pi(N) = v(N)$.

1.2.3 Phase three – $m \ge 1$, $n \ge 2$ and $\eta \ge 1$

In this phase, anyone can join the game of trading super capacity as futures and options on futures. We adopt options on futures as one of the financial instruments because the investor has an opportunity in stage 2 to decide whether or not to exercise the right after the capacity futures spot price at maturity has been observed. Hence, retailers and suppliers may hedge both quantity and price risks of the super capacity.

We consider both risk-neutral and risk-averse speculators in our study. In the model, risk-averse hedgers need to give up some of their payoffs in terms of risk premium to speculators who should earn a risk premium for bearing the commodity demand and price volatilities. The trade-off is between risk and expected return. The individual players' attitudes towards risk will affect the sharing of risk among themselves. For the risk-neutral speculators, we use the expected utility model to identify a certainty equivalent and use an auxiliary function, benefit function that is derived from a utility function to measure how much benefit an individual player would be willing to relinquish in order to reach a certain utility level. Therefore, we investigate how the capacity risk of retailers and suppliers can be transferred fully to other risk-neutral speculators.

By following this set up, we add a mean variance preference function for the risk-averse players in the game to work out an optimal risk-sharing rule that will maximize all players' payoff from the investment. We further use comparative statics to analyze behaviour of any influence by different degrees of risk aversion of individual players in the game.

1.2.4 Capacity unit

First of all we establish a proposal that a piecewise polynomial value-added production function contains only positive/negative value-added processes and non-value-added processes, but not waste. Then we use this proposition to investigate capacity deployment. We prove that the value-added capacity is fully efficient and shows no difference between different production lines to develop a time-based, value-added capacity measurement model that is an output orientated input measure to satisfy the need.

The above methodologies have been employed in our research to help us to answer the research questions in Chapter 2 of this dissertation.

1.3 Results of the Study

We propose a co-opetition game in which retailers can reduce and hedge against uncertainty by buying super capacity as futures as an alternative inventory to establish inventory portfolio with physical stock for single-period products. This novel mechanism allows retailers to make bidirectional adjustment to their inventory positions by exchanging their super capacity holdings as a commodity with their competitors in a sub-industry. In contrast, suppliers can share part of the benefit from reducing mismatching cost as well as increase capacity utilization by the mechanism. It also helps create a more stable production environment during the selling season. The following are summaries of our findings in the study.

1.3.1 Phase one

In Phase one, we study two supply chains in trading super capacity futures and show that Pareto-improvement can be attained under this mechanism. The condition to induce a supplier to build super capacity is identified. We also find that the price of super capacity has a strong effect on retailers in modifying their physical inventory levels in period 1. Retailers will hold less inventory if the price of super capacity is low compared with their profit margins.

1.3.2 Phase two

Our findings in Phase two present the situation in which the trading of super capacity as a commodity in the futures market can provide an effective mechanism to reduce demand variability for the supply chains, and shift both capacity and inventory risks from a certain supply chain to a sub-industry. We find that the game is efficient in both the non-cooperative and cooperative stages, as all the players will reach a unique equilibrium point in each stage. We also show that the inventory and capacity risks in newsvendor supply chains can be mitigated among different supply chains selling different products. Different supply chains can regulate their inventory positions by using a new tool that more accurately matches demand with supply. Moreover, the sub-industry will increase their aggregate payoff as a result of trading super capacity futures.

1.3.3 Phase three

In Phase three, the existence of the unique Nash equilibrium demonstrates that this new mechanism is an efficient means to let a single supply chain risk be shared with, or be transferred to, the other supply chains and even to the public who are in the game. Our results indicate that this new market-based risk transfer mechanism combines operational and financial hedging strategies that offer industry a new way of meeting demand more efficiently to improve profit. We also determine an optimal risk-sharing rule for the risk-averse players to share their payoff, and we find that the higher the risk aversion of a player, the lower the profit share is for that player, but the shares to other players are unaffected.

1.3.4 Capacity unit

A time-based, 'value-added capacity' measurement model for super capacity trading is established by showing the value-added capacity is fully efficient and there is no difference between different production lines. It can be easily inferred to output quantity of the retailer's order if the requested value-added input of the product is known.

This model is an innovative and apt measurement for the capacity planning of a process that produces a wide product mix and a new unit for super capacity trading that can be applied across different production facilities.

The results of this study show that the operational hedging strategies proposed in this study are valid. They therefore have satisfied the aim and the purposes and also respond to the research questions in this dissertation.

1.4 Organization of the Dissertation

The chapters of the dissertation are organized as follows:

Chapter 1 *Introduction*: the current chapter.

Chapter 2 *Literature Review*: we present a literature review on inventory policies and strategies, risk pooling and financial instrument for hedging. We look from the three different sources of additional inventory in stage 2, including supplier coordination, retailer cooperation and consumer compulsion to understand how earlier research has obtained knowledge in alleviating the mismatching problem in a single-period, two-stage environment. Then we identify research gaps for our study.

Chapter 3 *Trading super capacity between two supply chains*: we construct a model of two supply chains with two retailers and only one supplier to analyze two scenarios: one where super capacity exchange is not allowed, and the other where super capacity exchange is allowed, in order to determine whether Pareto-improvement can be attained through super capacity trading. We also study how retailers and suppliers make their decisions on capacity trading. Numerical results are demonstrated to illustrate the relationships among some important variables and parameters of the model.

Chapter 4 *Trading capacity future among n-newsvendor*: in this chapter we argue that the suppliers are induced to offer their super capacity to retailers and we also identify the characteristics of the retailers who will become players in the game. The biform analysis is adopted to study the Nash equilibrium in the non-cooperative stage and the core of players in the cooperative stage. Then we provide two numerical examples to elucidate the analysis.

Chapter 5 *Transfer mismatching risk to the public*: the optimal policy of the super capacity trading game is examined in this chapter. We also investigate how risk is transferred from hedgers to risk-natural speculators and risk-averse speculators, respectively. From there, the optimal payoff rule for players in the game is identified. We show that the whole sub-industry will gain extra profit – even speculators are in the game. We use numerical examples to explain how mismatching risks are transferred to the players in the game under different situations.

Chapter 6 *Capacity unit for trading*: we start from the literature review of capacity measurement that is divided into three streams, namely (a) engineering approach, (b) cost accounting approach, and (c) economic approach. Next we study the components of a process that consists of only value-added and non-value-added work. We also propose a new concept on waste in the production process. Then we discuss fully efficient value-added input, and define a capacity deployment according to the results in this chapter. We therefore explain how a new super capacity trading unit is innovated.

Chapter 7 *Conclusions*: we summarize our major findings in the beginning of this chapter. From that point, we discuss how our postulation fits the operational hedging strategies and we talk about the concern of the existence of short lead-time capacity. An example of capacity trading in textile and clothing that was used by some countries between the 1960s and 2004 was quoted for reference. Besides, we suggest some managerial implications in this chapter. Then the limitations of our study are indicated and some suggestions for future research are also proposed. At the end of this chapter we have concluding remarks.

Appendix A: The mathematical proofs related to Chapters 3-6.

Appendix B: *A new look at waste and value-added work in the Toyota Production System*: we provide a deeper discussion in the key concept for developing the new capacity unit in Chapter 6. We start by introducing the basic concept of Toyota Production System (TPS) and reviewing the dissemination of the concept and practice of TPS in the Western world. Then the notion of waste estimation and the new definitions of the different work components of a process are discussed to argue that waste does not exist in any process. We also derive the conditions under which productivity will improve from shortening lead-time and cycle time. A case study in which a Thai manufacturer adopts the ideas in Chapter 6 and this appendix to achieve efficiency improvement in its production process is exhibited.

In the next chapter will examine the literature on inventory policies and strategies, risk pooling and financial instruments for hedging. We also develop research questions in the last section.

2 Literature Review

The single-period two-stage inventory problem with replenishment has been studied and investigated in many different directions since the research of Bradford and Sugrue (1990), which presents a succinct model for style-goods inventory issue. However, we adopt the aspect of replenishment sources from a historical perspective to support a structural review in this field since it describes comprehensively different strategies for the problems that are related to our aim. Obviously, our literature review also covers both operation management and risk management, including commodity futures to provide a sufficient context for our study.

2.1 Introduction

This chapter presents a review of research and literature on operational and financial hedging to alleviate the supply-demand mismatching problem. There are four major sections. The next section considers the literature from inventory policy and strategy. We focus on the area showing how different strategies are formulated according to sources of replenishment to reduce mismatching risk. In Section 2.3, we then draw upon literature from the disciplines of risk pooling that relate to reduced demand variability in supply chains. Following this, in Section 2.4 we present the literature that studies financial instruments for hedging and transferring the inventory risk. In Section 2.5, we highlight the research gaps and research questions that are of particular relevance to the primary objective of this study: to mitigate mismatching cost by pooling and sharing the risk among different supply chains in a sub-industry, and even transferring it to public speculators via financial instruments. In the final section, we summarize the chapter.

2.2 Inventory Policies and Strategies

In a single-period, two-stage problem, the retailer has different ways to obtain replenishment after early sales have been realized. According to the literature, there are three different sources from which additional orders may be secured to alleviate the supply-demand mismatching problem: namely, (a) supplier coordination, (b) retailer cooperation, and (c) consumer compulsion. The following is a summary of the inventory policies and strategies that are based on the three replenishment opportunities.

2.2.1 Supplier coordination

For short life-cycle products, a simple form of postponement permits the retailer to place a second order with the supplier using an option contract and capacity commitment design. Capacity reservation for the single-period inventory problem has been studied since the 1990s. Silver and Jain (1994) initiate research on procurement strategies for reserved capacity for single- and multi-period inventory problems. Jain and Silver (1995) use a reserved capacity option to treat the single-period problem where there are uncertainties in both demand and supplier capacities. The option contract is a commonly used arrangement that provides quantity flexibility to the retailer, enabling it to respond to market changes because the supplier has earmarked either a specified or an unlimited amount of capacity for the retailer. Anupindi and Bassok (1999) present a general model of supply contracts with options for a single short life-cycle product with stochastic demand in two stages. They allow buyers to purchase options from suppliers that include an option price, in addition to firm orders to be delivered at the beginning of stages one and two. In stage two, the buyer needs to pay an exercise price for additional units, which are reserved by an option. Barnes-Schuster *et al.* (2002) provide an in-depth study of this general model by showing that an option can offer flexibility to buyers, allowing them to better react to demand changes and benefit from channel coordination. Wang and Liu (2007) develop an option contract for channel coordination and risk sharing in a retailer-led supply chain. They find that to be successful, the option price (for capacity reservation) and the exercise price (purchasing from the reserved capacity) have to be negatively correlated and that the retailer can only commit to taking a quantity that is smaller than the optimal production quantity in a centralized system.

Similarly, some researchers have made efforts to find ways of developing capacity reservation that allow retailers to buy additional goods after early sales performance has been observed. Chung and Flynn (2001) extend the classic newsvendor model to a two-stage problem where they reduce a piecewise linear convex cost using multiple replenishments in the second stage. Li and Liu (2008) design a supply contract that allows buyers to place a second order under the manufacturer’s limited reserve capacity. They discover that their design will result in higher profits if demand has a larger variability. Milner and Rosenblatt (2002) suggest a two-period contract that offers the buyer the option to adjust the second order by paying a per unit order adjustment penalty.

Each of these studies takes the buyer’s perspective in order to provide favourable quantity flexibility. Although the retailer may benefit from procuring supplies under improved conditions, the manufacturer may end up being worse off due to reductions in the volumes of the retailer’s purchases (Wu 2005). Earlier research has also found that adjusting the inventory level downstream based on demand forecast updates results in disturbance upstream (Donohue 2000). Burnetas and Ritchken (2005) find that using options is not a zero-sum game, since the retailer is better off if the uncertainty in the demand curve is low, but may be worse off if the uncertainty is sufficiently high.

2.2.2 Retailer cooperation

Cooperation among retailers is another source for replenishment during the season. Transshipment on residual inventory among retailers is a common strategy in the literature. Transshipment is similar to our study, but the key difference is that it only concerns the transfer of physical goods. Krishnan and Rao (1965) open a new stream in the analysis of transshipment of a multi-location distribution network with a number of warehouses. For single-period, two-stage problems, Rudi *et al.* (2001) examine the chances of optimal inventory orders of two newsvendors if transshipment is allowed to transfer surplus products to another retailer who is stocked out. They find that the optimal order levels under transshipment will not gain the maximum joint profits if each retailer only tries to maximize its own profits. However, they suggest using a pair of transshipment prices to overcome this problem. Hu *et al.* (2007) generalize the model of Rudi *et al.* (2001) by considering uncertain production capacity. Zou *et al.* (2010) also investigate the impact of transshipment between two retailers, but in a competitive environment. However, Zhao and Atkins (2009) realize that competitive retailers will benefit from transshipment if they have stronger differentiation, or if competition among them is weak; otherwise they will prefer consumer substitution.

Dong and Rudi (2004) study the effect of transshipment for a distribution system that contains a single supplier and n retailers. They find that transshipment makes retailers' order quantities less sensitive to the wholesale price if the supplier is a price setter. Therefore, the supplier benefits from retailers' transshipment by charging a higher wholesale price. The results of Dong and Rudi are extended by Zhang (2005), who finds that the inventory problem with transshipment is equivalent to a newsvendor problem with an adjusted demand. Anupindi *et al.* (2001) analyze a decentralized distribution system with n retailers that order inventory independently before the season and share excess goods to satisfy unmet demand after demand realization. They show that there is an equilibrium allocation mechanism in the game. Granot and Sosic (2003) extend Anupindi *et al.*'s (2001) study by considering that not all the residual inventories will be transshipped to other retailers. They find that the core allocation rules may not induce the retailers to share their entire residual inventory with others. Based on these two papers, Sosic (2006) further studies retailers that show concerns for the reactions of the other retailers to their actions, and identifies that grand coalition is a farsighted stable outcome. Slikker *et al.* (2005) and Ozen *et al.* (2008) also study cooperation games played by *n* newsvendors to improve their expected joint profit by transshipping their goods after demand realization is known. However, Huang and Sosic (2010) find that allocation of the residual profit from transshipment among players is not significantly different, no matter the transshipment price is selected before demand is known or dual allocations are adopted after demand is known.

2.2.3 Consumer compulsion

A few papers investigate the random quantity of replenishment during the selling season by consumer compulsion in particular businesses, such as catalogue and Internet mail order. The high volume of returned goods in mail order is resalable, therefore retuned goods are also a source of replenishment if they come back to retailers on time without damage. Vlachos and Dekker (2003) investigate seasonal products sold through E-commerce or mail order with a high return rate. They examine the influence of returned goods on the initial order quantity of a single-period product by assuming that the return rate is fixed and the product can only be resold once. Mostard and Teunter (2006) consider a similar newsvendor problem with resalable returns. Moreover, Mostard *et al.* (2005) study the same model but allow the product to be returned and resold infinitely in the same season. They discover that demand variability is a key factor that affects the expected profit in their models.

In the next section, we examine the demand variability moderation by risk pooling strategy. The review will extend beyond the single-period two-stage problem to a general inventory and capacity environment.

2.3 Risk Pooling

Risk pooling is an effective tool that can be used to address and reduce demand variability in supply chains (Simchi-Levi *et al.* 2008). For example, the risk-pooling effect of centralizing inventory can benefit inventory systems by reducing the need for safety stocks and consequently lowering the costs associated with inventory holding and shortage penalties (Eppen 1979). In recent years, risk-pooling strategies have received considerable attention in the academic community and in practice (Cachon and Terwiesch 2009). Schwarz (1989) studies the use of a warehouse to pool risk arising from outside-supplier lead-times in order to reduce the overall variance in the retailers' net inventory processes. Kumar *et al.* (1995) analyze static and dynamic policies for replenishing and allocating inventories among n retailers located along a fixed delivery route. Gerchak and He (2003) study the impact of demand variability on the expected overall inventory costs under risk pooling.

Thomas and Tyworth (2007) investigate pooling the lead-time risk by splitting orders and simultaneously offering significant opportunities to reduce inventory-system costs. Cachon and Harker (2002) demonstrate that manufacturers can benefit from capacity pooling by outsourcing. Gerchak *et al.* (1988) investigate the inventory effects of using commonality in assemble-to-order systems and find that it is beneficial to employ commonality whenever possible. Kurata *et al.* (2007) study how a two-stage supply chain system consisting of a supplier and two manufacturers can reduce order variability by applying a form of risk pooling and bundling. However, they only look at a single supply network.

In fact, it is hard to obtain an optimal result if capacity and inventory are addressed separately, because this ignores the interaction between capacity and inventory within a manufacturing system (Bradley and Arntzen 1999). The substitution between capacity and inventory for short life-cycle products with stochastic demand has also been studied by Angelus and Porteus (2002).

Having discussed how operations can be structured to mitigate the mismatching risk, we will now turn to hedge inventory risk by different financial instruments in the next section.

2.4 Hedging by Financial Instrument

There are a few other papers considering integration of financial and operational risk management tools to hedge inventory risk for short life cycle products. Nevertheless, the studies are either only to combine risk pooling with financial hedging or about postponement with financial hedging. Ding *et al.* (2007) investigate the risk pooling effect of global firms that have production capacity in different countries to obtain a joint optimal capacity and financial option decision. Gaur and Seshadri (2005) use different underlying financial assets to hedge inventory risk. They use the price information of the financial asset to determine both optimal inventory level and hedge. Caldentey and Haugh (2006) extend their work to study the problem of continuous hedging of profit risk. Chod *et al.* (2010) examine the relationship between capacity flexibility and the value of financial hedging in the process of minimizing the risk of stochastic demand. The use of mean-variance and risk-aversion parameters to measure a retailer's profit risk in these papers is also a common technique to be adopted in the study of risk-averse agents in newsvendor problems as shown by Lau (1980), Lau and Lau (1999b), Agrawal and Seshadri (2000), and Gan *et al.* (2005). The common finding in the literature is that the risk-averse newsvendor normally orders less than the risk-neutral newsvendor. However, Wu *et al.* (2009) show that a risk-averse newsvendor may order more than a risk-neutral newsvendor if the lost sales cost is considered with mean-variance trade-off.

2.5 Research Gaps and Questions

It seems that the literature has not yet considered shifting or reducing the inventory and capacity risks among a group of supply chains using futures trading as a pooling mechanism. In the literature, transshipment study is close to the concept of our exchange capacity. The joint profit can be improved by transshipping goods among a group of newsvendors to adjust their inventory positions after demand realization is known. However, most papers limited their investigations to a single product, for example, Rudi *et al.* (2001), Dong and Rudi (2004), and Sosic (2006). Our study differs from their studies because capacity pooling can not only come across different products to benefit a certain sub-industry, but we can also trade capacity as a commodity in the futures market as an underlying asset for both suppliers and retailers in order to hedge the demand risk. Therefore, retailers can avoid holding too much inventory by sharing reserved capacity. Suppliers can also benefit from this mechanism by increasing the utilization of their reserved capacity.

Moreover, there is also a gap in the literature to study how risk-averse players transfer their risk beyond the sub-industry. This study tries to fill the gap partially by also examining the involvement of speculators in the super capacity futures game.

Therefore we want to study in this thesis a novel coordination strategy whereby the mismatch between supply and demand is mitigated, and risks of inventory and capacity are transferred from a single supply chain to a sub-industry via the trading of super capacity with competitors. Hence we propose the following research questions in order to address the research gaps:

1. Is the pooling mechanism of trading super capacity futures driving to Pareto-improvement?
2. How does the price of super capacity in stage 1 affect ordering quantity of physical inventory under demand uncertainty environment?
3. Are there any entry barriers for a retailer to be a hedger in the super capacity trading game?
4. Will the expected additional profit of hedgers who participate in the game of exchanging super capacity residual be maximized?
5. Does a unique Nash equilibrium exist in the game, with and without speculators?
6. Will risk-averse hedgers transfer their risk beyond the sub-industry?
7. Will the benefits of the entire sub-industry deteriorate if speculators are allowed to involve themselves in the capacity trading?

Moreover, we need to develop a capacity unit that is a common measurement among all the products and facilities for super capacity trading to be processed. Our last research question, therefore, is:

8. What is the capacity unit model that can serve super capacity trading among different supply chains involving various products?

2.6 Conclusions

This chapter has reviewed different strategies to alleviate the supply-demand mismatching problem. The option contact is a common tool to be used along a supply chain between players. The retailers can also cooperate amongst themselves by transshipping of inventory, but this policy is limited to the retailers who sell identical and/or substituted products. However, risk pooling is an effective strategy to reduce demand variability in supply chains. The mismatching risk can also be hedged and transferred via financial instruments. We have developed eight research questions to try to fill up some of the gaps in the study of mitigating mismatching cost by pooling and sharing the risk among supply chains and transferring it to the public via financial instruments.

The next chapter will propose a model with two newsvendor-type of supply chains to examine a co-opetition game between them in order to reduce and hedge against capacity risks and inventory risks by trading “super capacity” futures.

3 Trading Super Capacity Between Two Supply Chains

In this chapter we propose a novel mechanism for pooling the reserved capacity (super capacity) of different supply chains so that they can more effectively match their single-period inventory supplies with their demands. In the current set up, however, only two retailers buy or sell unutilized super capacity independently as a commodity in a sub-industry before and during the selling season, which helps improve supply flexibility and increases the utilization of suppliers' reserved capacity.

3.1 Introduction

We consider and compare two scenarios, namely one where capacity can be exchanged between two supply chains, and another where this is not allowed. We examine the impact of the price of super capacity on both retailers and suppliers and whether Pareto-improvement can be achieved through this new mechanism of super capacity trading. Our aim is to study the effects of capacity pooling among different supply chains on the mitigation of single-period inventory and capacity risks.

The rest of the chapter is organized as follows: in the next section we construct a model of two supply chains with two retailers and only one supplier. In Section 3.3 we analyze both super capacity exchange as it is and is not allowed, to determine whether Pareto-improvement can be attained. In Section 3.4, we also study how suppliers make their decisions on capacity trading. In Section 3.5 we provide numerical results to illustrate the relationships among some important variables and parameters of the model.

Finally, we conclude the chapter in Section 3.6.

3.2 The Model for Two Supply Chains

We examine two supply chains in an industry that sells non-identical short life-cycle products. Each supply chain has one retailer with a common supplier, with a planning horizon of two stages and correlated demands. The two retailers, each facing stochastic demand, buy goods from the same supplier that operates two separate production lines. These two production lines use the same facility and are able to produce goods for either of the retailers. However, each line usually serves only one particular retailer before the selling season. Even though the supplier needs a very short lead-time to fill the retailers' orders, the retailers have to place orders both before the start of the season (stage 1) and during the season (stage 2) since the supplier's capacity is limited. However, the retailers may buy super capacity in stage 1 for stage 2 orders.

Let the index i of the supplier be $i=1$ and the two retailers are $i=2,3$. Thus, our two retailers i place orders based on their own forecasts for q_i units at wholesale prices ω_i per unit with the supplier before the start of the selling season. To supply retailers i, the supplier's cost is c_i per unit. The timings of placing an order and starting production depend on the volume of the retailer's forecast. The larger the volume, the earlier production should start to overcome the risk of capacity limitation. At the same time, each retailer may also buy super capacity for stage 2 at a unit price h in stage 1.

The supplier offers super capacity from both production lines at the same price because the capacity is identical. Retailers i sell the goods to consumers at unit prices r_i. The supplier adopts the "make-to-order" policy, so it need not hold any inventory. If a retailer i cannot satisfy the demand eventually, it will incur a goodwill penalty g_i and the supplier will also suffer a similar penalty g_i^S for failing to meet retailer i's demand. Moreover, the supplier will bear the cost of unutilized super capacity at s per unit. To ensure meaningful analysis, we assume $r_i > \omega_i > c_i \geq s$.

We assume that the order quantities are determined only by the retailers as mentioned in Section 1.2. Therefore, the retailers must each choose an order quantity before the start of the selling season and may also buy a certain amount of super capacity to reduce the risk of holding too much inventory, even though its demand information is uncertain at this stage. They will place a second order in stage 2 after early sales performance has been observed so as to adjust their inventory levels according to the available market information. To simplify the analysis, we follow the approach of Smith *et al.* (2002) by assuming that the second orders are placed at the end of the selling season when realization of demand has been fully observed. The second order can therefore be treated as a backlog order and the retailers' lost sales are assumed to be due to unsatisfied demand because of insufficient inventory ordered in stage 1. Thus, we use α to represent the ratio of super capacity to lost sales for retailer *2* and β to represent the ratio of *excess* super capacity to lost sales for retailer *3*. We consider the case where only one of the retailers (i.e., retailer *2*) always has stage 1 inventory less than demand because if both retailers have surplus inventory in stage 1, then super capacity trading will not take place. For retailer *i*, the net salvage value of unsold inventory at the end of the season is v_i, where $v_i < c_i$. The production costs and the wholesale prices of the second order remain the same as in stage 1; i.e., c_i, and ω_i, respectively, for retailer *i*. These values do not change in stage 2 because the supplier uses the same production lines and methods to make the products as in stage 1. However, retailers' payments for the super capacity for stage 2 orders are non-refundable.

Let $D_i \geq 0$ be the random demand during the selling season with mean $\mu_i = E[D_i]$. Retailer i faces its own demand function with a probability density function $f_i(y)$ and a cumulative distribution function $F_i(y)$. Let $\overline{F}_i(y) = 1 - F_i(y)$, where $F_i(0) = 0$. We assume that all the distribution functions are continuous, invertible, twice differentiable, and independent of the wholesale price offered by the supplier.

3.3 Two Scenarios

We consider two different scenarios in this chapter. In both scenarios, the retailers place one order and buy super capacity in stage 1. Then they determine the additional inventory they will need at the end of stage 2. We do not consider cases where both retailers have insufficient super capacity and where both retailers have excess super capacity on hand, because no exchange of super capacity will occur under such circumstances. Therefore, we suppose that one of the retailers, $i=2$ has insufficient super capacity, while the other $i=3$ has excess super capacity in stage 2. The super capacity that retailer *2* holds is equal to α of its lost sales. The *excess* super capacity that retailer *3* holds is equal to β of its lost sales. We assume that the profit functions are concave. The decision problems of retailers *2* and *3* are $\max_{q_2 \geq 0} E[\pi_2(q_2,\alpha,h)]$ and $\max_{q_3 \geq 0} E[\pi_3(q_3,\beta,h)]$, respectively.

3.3.1 Scenario 1 - Super capacity exchange is not allowed

In the first scenario, we consider the case where super capacity cannot be exchanged between the two supply chains. We first consider retailer *2,* which has insufficient super capacity. The expected profit for it is

$$\begin{aligned}\pi_2(q_2,\alpha,h) = -\omega_2 q_2 + r_2\int_0^{q_2} y f_2(y)dy + \alpha(r_2-\omega_2-h)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy \\ - g_2(1-\alpha)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy.\end{aligned} \tag{3.1}$$

The first term in Eq. (3.1) is the cost of goods bought from the supplier before the season. The second term is the sales revenue from selling the product in the inventory built up before the season. The third term is the profit at the end of stage 2 from the quantity sold by exercising super capacity on hand. The last term is the cost of lost sales. The problem for retailer *2* is to determine the optimal ordering quantity in stage 1 to maximize its expected profit (3.1).

Similarly, the expected profit for retailer *3* that has surplus super capacity is:

$$\pi_3(q_3,\beta,h) = -\omega_3 q_3 + r_3\int_0^{q_3} y f_3(y)dy + (r_3 - \omega_3 - h)\int_{q_3}^{\infty}(y - q_3)f_3(y)dy$$
$$- \beta h\int_{q_3}^{\infty}(y - q_3)f_3(y)dy + v_3\int_0^{q_3}(q_3 - y)f_3(y)dy\,. \tag{3.2}$$

It should be noted that the expected leftover inventory is added in (3.2). The quantity of super capacity is greater than the lost sales in this case because retailer *3* has excess super capacity after fulfilling its own backlog order.

Lemma 3.1. *Even if the supplier operates under forced compliance, it can manipulate the stage 1 quantity by* h*, but not the size of the super capacity.*

Proof. Since the expected profit function (3.1) is concave, the optimal quantity q_2^* satisfies

$$q_2^* = \overline{F}_2^{-1}\frac{\omega_2}{r_2 + g_2 - \alpha(r_2 + g_2 - \omega_2 - h)}\,. \tag{3.3}$$

For the expected profit function (3.2), the optimal quantity q_3^* satisfies

$$q_3^* = \overline{F}_3^{-1}\frac{\omega_3 - v_3}{\omega_3 - v_3 + h(1+\beta)}\,. \tag{3.4}$$

Therefore, both q_2^* and q_3^* increase as h increases and vice versa. However, q_2^* and q_3^* are moving in opposite directions when α and β are moving in the same direction. Since the supplier does not know the size of the demand before the selling season, it is hard for it to use super capacity to manage stage 1 quantity. The supplier can, therefore, influence the optimal quantity in stage 1 by adjusting the unit price of the available super capacity, but not the volume of super capacity. □

From Lemma 3.1, we see that retailers will try to keep the physical inventory as small as possible if the price of super capacity is small. However, if h increases, retailers may increase the physical inventory in stage 1 and reduce the holding of super capacity.

3.3.2 Scenario 2 - Super capacity exchange is allowed

In this scenario, the retailer that has insufficient super capacity is allowed to buy super capacity from the other retailer that holds excess super capacity in stage 2. We assume that additional orders will be placed with the supplier to utilize the full amount of the traded super capacity in stage 2. Since only two retailers are involved here, they will have to exchange super capacity with each other. For simplicity, it is assumed that the two retailers adopt the same unit price as in stage 1 for the trading of super capacity between them and that no transaction costs will be incurred. We consider two situations as follows.

3.3.2.1 The case where the retailer buys super capacity

Retailer *2* can only buy the exact quantity of extra super capacity from retailer *3* that sells the same quantity in excess. We denote the quantity of super capacity bought as α' of retailer *2*'s lost sales. The quantity of super capacity sold is equal to β' of retailer *3*'s lost sales. Obviously, the condition $\alpha + \alpha' \leq 1$ must hold. Hence, the new expected profit function for retailer *2* is

$$\pi_2'(q_2,\alpha,h) = -\omega_2 q_2 + r_2\int_0^{q_2} y f_2(y)dy + (\alpha+\alpha')(r_2-\omega_2-h)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy$$
$$-g_2(1-\alpha-\alpha')\int_{q_2}^{\infty}(y-q_2)f_2(y)dy. \tag{3.5}$$

3.3.2.2 The case where the retailer sells super capacity

Retailer *3* that has excess super capacity sells as much as possible of the unutilized super capacity that it holds to minimize cost. We denote the super capacity units that retailer *3* transfers to retailer *2* as β' of its lost sales, where $\beta' \leq \beta$. Hence, the new expected profit function for retailer *3* is

$$\pi_3'(q_3,\beta,h) = -\omega_3 q_3 + r_3\int_0^{q_3} y f_3(y)dy + (r_3-\omega_3-h)\int_{q_3}^{\infty}(y-q_3)f_3(y)dy$$
$$-h(\beta-\beta')\int_{q_3}^{\infty}(y-q_3)f_3(y)dy + v_3\int_0^{q_3}(q_3-y)f_3(y)dy. \tag{3.6}$$

Proposition 3.1. *Regardless of whether super capacity is insufficient or in excess, retailers will reduce their optimal inventory quantities in stage 1 if they are allowed to trade super capacity in stage 2 between themselves.*

Proof. From (3.5) and (3.6), we obtain a new optimal q_2^0 that satisfies

$$q_2^0 = \overline{F}_2^{-1}\frac{\omega_2}{r_2+g_2-(\alpha+\alpha')(r_2+g_2-\omega_2-h)} \tag{3.7}$$

and a new q_3^0 that satisfies

$$q_3^0 = \overline{F}_3^{-1}\frac{\omega_3-v_3}{\omega_3-v_3+h(1+\beta-\beta')}. \tag{3.8}$$

Comparing (3.3) to (3.7), and (3.4) to (3.8), we see that that $q_2^0 < q_2^*$ and $q_3^0 < q_3^*$. □

Proposition 3.1 shows that both retailers will minimize risk by keeping less inventory in stage 1 when they are given the option of trading the alternative inventory in stage 2. The retailers therefore use short lead-time capacity as an alternative inventory instead of the physical product to adjust their inventory positions, as well as avoiding losses resulting from mismatching between supply and demand. In other words, retailers can hedge against inventory risk by trading super capacity with their competitors in the season after early sales have been observed. Similarly, from the supplier's viewpoint, the shifting of some order volume from stage 1 to stage 2 can help ease the peak production pressure before the selling season. In addition, the supplier will also increase its income by selling super capacity to retailers.

Lemma 3.2. *Retailers that buy or sell super capacity at spot markets are better off than if there is no trading of super capacity in stage 2.*

Proof. a) Comparing (3.1) and (3.5), we obtain $\pi_2(q_2,\alpha,h) < \pi_2'(q_2,\alpha,h)$, where the difference is $\alpha'(r_2-\omega_2-h+g_2)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy$.

b) Comparing (3.2) and (3.6), we obtain $\pi_3(q_3,\beta,h)<\pi_3'(q_3,\beta,h)$, where the difference is $h\beta'\int_{q_3}^{\infty}(y-q_3)f_3(y)dy$. □

A retailer can improve its profit if it has a chance to adjust its inventory position during the selling season. If market demand is higher than expected, it can use the super capacity it has reserved and can also buy extra super capacity from the other retailer in order to generate additional profits by selling more products in stage 2. On the other hand, if demand is less than anticipated, it can save some costs by selling the surplus super capacity that it holds. Obviously, the profit of the retailer that buys super capacity will be higher than that of the retailer that sells super capacity if h is smaller than the profit margin.

3.4 The Supplier's Profit

We now examine the status of the supplier under different scenarios, in particular whether it will be better off under the proposed mechanism.

Lemma 3.3. *The supplier will reduce the amount of unutilized super capacity and will improve its expected profit if super capacity is traded among the retailers in stage 2.*

Proof. If trading of super capacity is not allowed, the supplier has the following expected profit function

$$\begin{aligned}\pi_1^N(\alpha,\beta,h)&=\omega_2q_2+\omega_3q_3-c_2q_2-c_3q_3+\alpha(\omega_2+h-c_2)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy\\&\quad+(\omega_3+h-c_3)\int_{q_3}^{\infty}(y-q_3)f_3(y)dy+\beta(h-s)\int_{q_3}^{\infty}(y-q_3)f_3(y)dy\\&\quad-g_2^S(1-\alpha)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy\\&=q_2(\omega_2-c_2)+q_3(\omega_3-c_3)-g_2^S\int_{q}^{\infty}(y-q_2)f_2(y)dy\\&\quad+\pi_1^N(\alpha,h)+\pi_1^N(\beta,h),\end{aligned} \tag{3.9}$$

where $\pi_1^N(\alpha,h) = \alpha(\omega_2 + h - c_2 + g_2^S)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy$ is the supplier's profit in stage 2 from an order of retailer *2*, including the super capacity income, and $\pi_1^N(\beta,h) = (\omega_3 + h - c_3 + \beta[h-s])\int_{q_3}^{\infty}(y-q_3)f_3(y)dy$ is the supplier's profit in stage 2 from an order of retailer *3*, including the net super capacity income. However, if trading of super capacity does occur between the retailers, then the supplier has the following expected profit function

$$
\begin{aligned}
\pi_1^T(\alpha,\beta,h) &= \omega_2 q_2 + \omega_3 q_3 - c_2 q_2 - c_3 q_3 + \alpha(\omega_2 + h - c_2)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy \\
&\quad + \alpha'(\omega_2 - c_2)\int_{q_2}^{\infty}(y-q_2)f_2(y)dy + (\omega_3 + h - c_3)\int_{q_3}^{\infty}(y-q_3)f_3(y)dy \\
&\quad + \beta h\int_{q_3}^{\infty}(y-q_3)f_3(y)dy - s(\beta-\beta')\int_{q_3}^{\infty}(y-q_3)f_3(y)dy \\
&\quad - g_2^S(1-\alpha-\alpha')\int_{q_2}^{\infty}(y-q_2)f_2(y)dy \\
&= q_2(\omega_2 - c_2) + q_3(\omega_3 - c_3) - g_2^S\int_{q}^{\infty}(y-q_2)f_2(y)dy \\
&\quad + \pi_1^T(\alpha,h) + \pi_1^T(\beta,h),
\end{aligned}
\tag{3.10}
$$

where

$$\pi_1^T(\alpha,h) = \{\alpha(\omega_2 + h - c_2 + g_2^S) + \alpha'(\omega_2 - c_2 + g_2^S)\}\int_{q_2}^{\infty}(y-q_2)f_2(y)dy$$

and $$\pi_1^T(\beta,h) = \{(\omega_3 + h - c_3) + \beta(h-s) + \beta' s\}\int_{q_3}^{\infty}(y-q_3)f_3(y)dy.$$

From (3.10), since $\beta \geq \beta'$, the unutilized super capacity is reduced to $(\beta-\beta')\int_{q_3}^{\infty}(y-q_3)f_3(y)dy$. So, from (3.9) and (3.10), we see that $\pi_1^T(\alpha,\beta,h) > \pi_1^N(\alpha,\beta,h)$. □

It is evident that the trading of super capacity will increase the utilization of super capacity in stage 2. Therefore, the supplier will reduce the costs due to unutilized capacity, and crucially, will also reduce the risk of not fully utilizing capacity. However, to be sure that the supplier can earn a higher profit in this way, we need to compare this scenario with one where there is trading of super capacity under the single-period newsvendor ordering framework.

Proposition 3.2. *The supplier will be induced to build super capacity if the unit price of super capacity satisfies:*

$$h \geq \gamma(\omega_3 - c_3 + s),\ where\ \ 0 \leq \gamma < 1.$$

Proof. See Appendix A. □

Proposition 3.2 shows that the higher the profit margin is, and/or the higher the cost of unutilized super capacity is, the higher is the value of h. However, the supplier will sell its super capacity if it has some value since the discount factor γ can be close or equal to zero. In fact, γ is the ratio of the unutilized super capacity to the total super capacity that the supplier has sold. If the supplier predicts that the utilization of super capacity will be low in the season, then it will ask for a higher h and vice versa. This means that the supplier considers two factors in determining h: namely, (i) the profit margin and the cost of unutilized super capacity, and (ii) the predicted percentage of unutilized super capacity. If the supplier is willing to sell super capacity, it will be interesting to test whether this game is Pareto-improving with respect to the non-trading policy.

Proposition 3.3. *If super capacity is allowed to be traded, then Pareto-improvement will be attained.*

Proof. Lemmas 3.2 and 3.3 show that both the retailers and the supplier will make more profits if super capacity is allowed to be traded between the retailers. Therefore, Pareto-improvement will be attained because all the players in the game are better off, and none is worse off. □

Whenever retailers are allowed to trade super capacity during the season, the utilization of super capacity is improved. Retailer *2* will reduce lost sales by selling more products to the market. It will benefit more than retailer *3* if its profit margin exceeds h. Retailer *3*, however, is only compensated by the cost of unutilized super capacity. In this way they can adjust their own inventory positions so as to be able to react more effectively and expeditiously to market demands. Nevertheless, the supplier's profit will increase from producing more goods. We have proved that all the parties will be better off by engaging in super capacity trading.

From the above analysis, we realize that trading super capacity can pool capacity among a group of supply chains to hedge against both inventory and capacity risks. This innovative concept also allows competitors to help one another to reduce the mismatch between supply and stochastic demand, thus reducing the cost of inventory leftover and lost sales for both the retailer and supplier in the sub-industry.

3.5 Numerical Results

In order to gain better insights into the theoretical results presented in Sections 3.3 and 3.4, we numerically studied the effects of a variety of the model's parameters and variables on the supplier and retailers. For each numerical example, we calculated the optimal solution and the corresponding total expected profit, if necessary. In order to be realistic, we talked to a couple of seasonal garment retailers to understand their cost structures. We learned that the wholesale price is approximately double the factory cost, and retail price is approximately double the wholesale price. This pricing configuration along the clothing supply chain is in line with the general practice of the clothing industry (Magretta 1998). Therefore, we set the values along the supply chain at $r_2 = 1$, $r_3 = 1$, $\omega_2 = 0.5$, $\omega_3 = 0.5$, $c_2 = 0.25$ and $c_3 = 0.25$, and let the demand follow a normal distribution with mean 10,000 and standard deviation 5,000. Other data were as follows: $s = 0.02$, $g_2 = 0.02$, $v_3 = 0.05$, $l_2 = 4{,}000$, and $l_3 = 600$, where l_2 and l_3

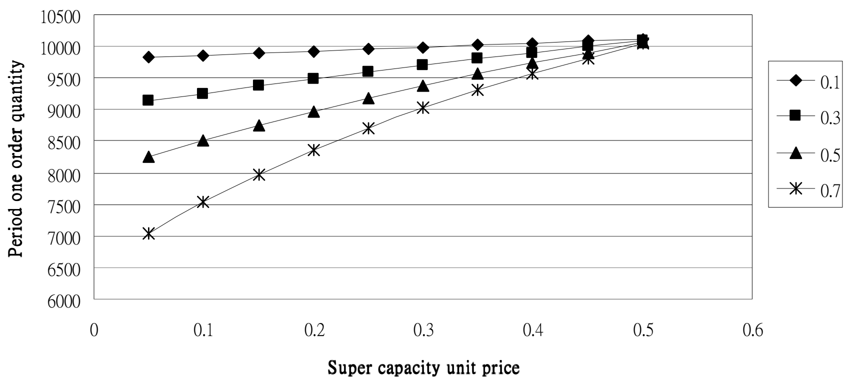

Figure 3.1: Order quantity of retailer 2 vs super capacity unit price in scenario 1

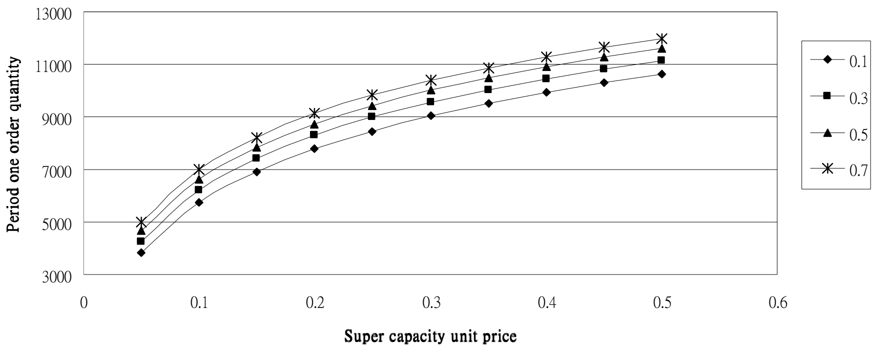

Figure 3.2: Order quantity of retailer 3 vs super capacity unit price in scenario 1

are the quantities of the lost sales of retailers *2* and *3*, respectively. These values were close to the market information we gathered from practitioners.

Figures 3.1 and 3.2 show the impacts of h, the unit price of super capacity, on the quantity of pre-season orders in Scenario 1. We tested the model by fixing both α and β at 0.1, 0.3, 0.5, and 0.7 to have a sufficient range of different situations. The numerical results were as predicted by Lemma 3.1, that both q_2^* and q_3^* increase as h increases, and vice versa under different values of α and β. We observed that α has a stronger impact than β. When h is at a low price level, retailer *2* buys fewer products in stage 1 if it predicts that α will be high. The difference in stage 1 order quantity was about 30% in our examples when α was between 0.1 and 0.7. However, when h was as high as the profit margin, there was no difference for different values of α. On the other hand, the impact of β on retailer *3* was only minimal, as h dominates the decisions. Retailer *3* buys less super capacity and keeps more inventory if h is high.

In Scenario 2, when super capacity exchange is allowed between retailers, the behaviour of retailers changes when deciding the pre-season order quantities. Figure 3.3 shows that the net order quantities are reduced in stage 1 if trading of super capacity takes place. We see that an increase in super capacity unit price closes the gap in order quantities in the two scenarios for different values of α for retailer *2*. When α' = 0.1 and h was as low as

5% of the retail price, retailer *2*'s reduced order quantities increased from 328 units to 833 units if α rose from 0.1 to 0.7, and if super capacity could be traded during the season. β has a stronger impact than the unit price of super capacity for retailer *3*, even if both retailers reduce their stage 1 order quantities by the same amount when super capacity exchange is allowed. Figure 3.4 shows that the reduced order quantities are quite stable when h ranges between 25% and 45% of the retail price. Both Figures 3.3 and 3.4 demonstrate the characteristics of Proposition 3.1 presented in Section 3.3.2.2. These results reflect that h is a main factor in determining stage 1 inventory level if hedging against risk by super capacity is available. We can see that the capacity trading mechanism provides a reasonable means for retailers to choose both inventory and hedge positions before the selling season.

In Lemma 3.2, we prove that retailers that buy or sell super capacity at spot markets are better off than when there is no trading of super capacity in stage 2. From (3.1) and (3.5), we know that retailer *2* increases its profit by

$$\Delta\pi_2 = \alpha' l_2 (r_2 - \omega_2 - h + g_2), \tag{3.11}$$

where l_2 is the quantity of the lost sales of retailer *2*. Figure 3.5 shows that α' has a stronger impact than h on profit improvement when h is small. For example, if $h = 0.05$, then the net increased profit was \$188 when $\alpha' = 0.1$, but it jumped to \$1,316 when $\alpha' = 0.7$. However, the amount of increased profit becomes increasingly closer among different α' when h rises. The retailer that exceeds its super capacity also increases its profits.

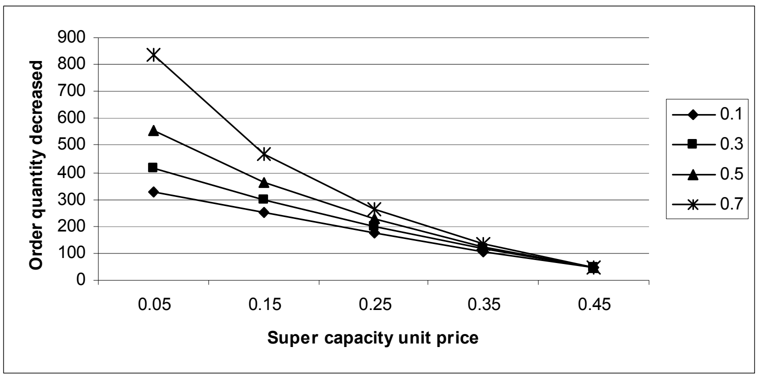

Figure 3.3: Retailer 2 decreases order quantity when super capacity exchanged is allowed

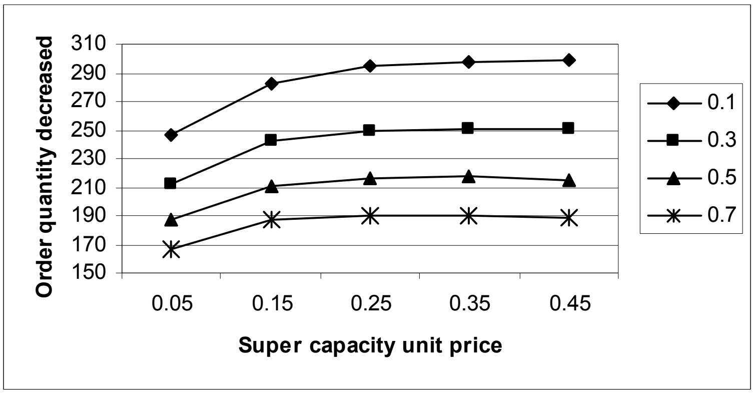

Figure 3.4: Retailer 3 decreases order quantity when super capacity exchange is allowed

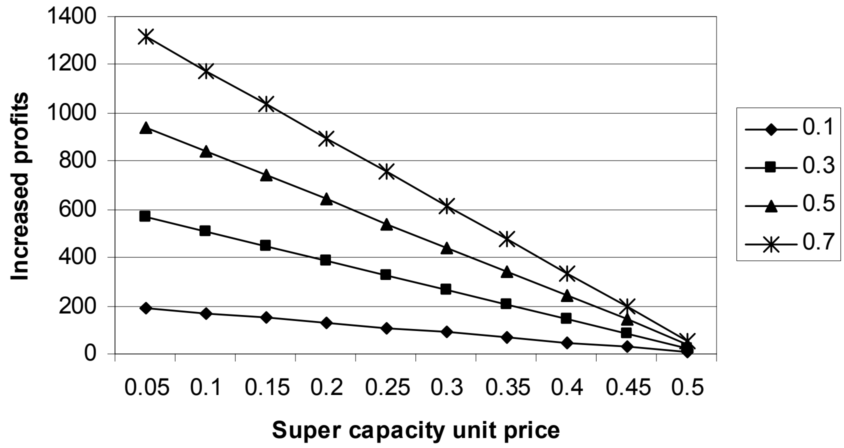

Figure 3.5: Increased profit of retailer 2 under different α' values

However, its profit is proportional to both h and β' as shown in Figure 3.6. The reason is that retailer *3* can only increase its profit by selling its excess super capacity, but not its profit margin. If we look at the supplier, its profit will mainly increase by producing more products. The supplier also slightly reduces its super capacity setup cost. Table 3.1 shows that α' has a much bigger effect than β' on profit improvement when trading of super capacity is allowed. However, the effect due to β' is very small. For example, when β' increased from 5% to 65%, profit only decreased by 2% when α' = 0.35. However, profit increased very significantly—by 1,377%—when α' increased from 5% to 65% at β' = 0.25. Obviously, retailer *2* may gain more benefit from the profit margin earning in this capacity trading, but retailer *3* can only compensate the loss in excess capacity.

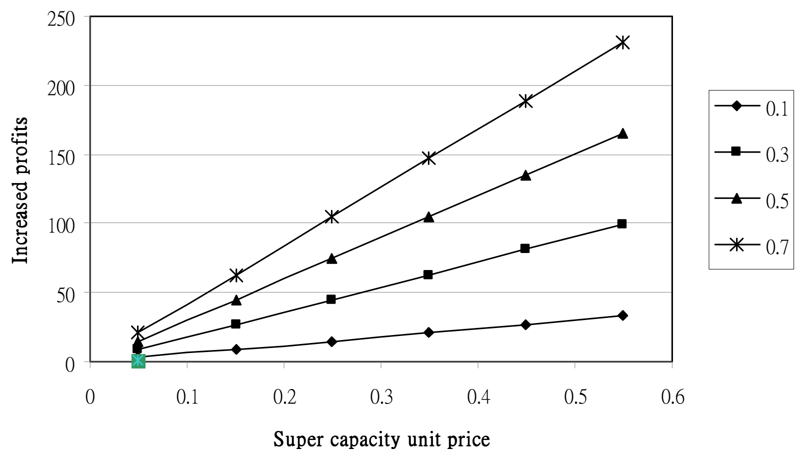

Figure 3.6: Increased profit of retailer 3 under different α' values

From the numerical examples, we obtain a better understanding of the behaviour of h under different situations and see that the results agree with the theoretical analysis presented in Sections 3.3 and 3.4. We observe that both the retailer and the supplier increase their profits if super capacity trading is allowed, attaining Pareto-improvement. Since the retailer is not willing to hold any super capacity if its price is equal or more than the profit margin, h has a limited price range. However, it is difficult for retailers to know whether they will have insufficient or excess inventory before the selling season, so the exchange of super capacity after demand is realized offers them an opportunity to adjust their inventory positions under an uncertain environment.

Table 3.1: Results of the numerical examples on the increased profit of the supplier under various of α' and β'

	$\alpha' = 0.05$	0.20	0.35	0.50	0.65
$\beta' = 0.05$	50	200	350	500	650
0.25	47	197	347	497	647
0.45	45	195	345	495	645
0.65	43	193	343	493	643

3.6 Conclusions

In this chapter we have studied a co-opetition game in which retailers can reduce and hedge against uncertainty by buying super capacity as futures as an alternative inventory for single-period products. We showed that the proposed super capacity trading mechanism can induce retailers to reduce their inventories before the selling season, since they can alter their inventory positions during the season. We also found that the price of super capacity has a strong effect on retailers in modifying their inventory levels in stage 1. Retailers will hold less inventory if the price of super capacity is low compared with their profit margins. We provided a theoretical framework to show that Pareto-improvement can be attained under this mechanism.

However, we confine our study to a setting with only one pair of supply chains in our model. We have not worked out the characteristics of super capacity and their impacts if multiple supply chains are involved in trading of super capacity. As a natural extension of our work, in the next chapter we analyze risk sharing in a sub-industry when there are multiple suppliers and retailers. We also study the increases in profit within a sub-industry as a result of the trading of super capacity futures and the improvement in performance from inventory matching of products with stochastic demand.

4 Trading Capacity Futures Among *N* Newsvendors

In this chapter we consider $n \geq 2$ competitive retailers (newsvendors) who sell non-identical products. Similarly to the last chapter, retailers buy physical goods and capacity futures as inventory portfolios in the first stage to determine their inventory positions in the selling season. After realization of demand observed in the second stage, retailers make a replenishment decision limited to the capacity futures on hand. However, retailers are allowed to form coalitions to transfer the capacity futures residual among themselves. Therefore, retailers can make bidirectional adjustments to their inventory positions.

4.1 Introduction

We employ a biform game in this chapter to analyze the risks and payoffs to retailers as players in both the non-cooperative (first) and cooperative (second) stages. We find the Nash equilibrium in the first stage and a core of the players who trade capacity futures residual in the second stage. Our findings reveal that retailers can share risk among different supply chains involving different products to mitigate inventory risk and improve their payoffs. However, the game discriminates against retailers who have lower profit margins, inventory costs, and lost sales penalties.

The remainder of this chapter is organized as follows: Section 4.2 describes the set up of our models. We argue that the suppliers are induced to offer their super capacity to retailers, and we also identify the characteristics of the retailers who will become players in the game. Section 4.3 uses biform analysis to study the Nash equilibrium in the non-cooperative stage and the core of players in the cooperative stage. Section 4.4 provides two numerical examples to illustrate the analysis. Section 4.5 concludes the chapter.

4.2 The Model for $m \geq 1$ Suppliers and $n \geq 2$ Retailers

There is a group of n retailers and m suppliers that form a sub-industry with n supply chains that sell different short life-cycle products to the market. Each supply chain is dominated by one retailer that buys goods from one or more of the suppliers. The suppliers all have the same facilities and similar capabilities to produce goods with a very short lead-time. Each supply chain faces a general two-stage and correlated demand that has a probability density function $f_i(y)$, $i = m+1,...,m+n$ and also a cumulative distribution $F_i(y)$ $i = m+1,...,m+n$. Similarly to Chapter 3, let $\overline{F_i}(y) = 1 - F_i(y)$ and $F_i(0) = 0$. We assume that all the distribution functions are continuous, invertible, and double differentiable. Since the suppliers each have their own limited capacity before the start of the selling season (stage 1), the retailers have to determine the order quantity, $q_i^p = (q_{m+1}^p,...,q_{m+n}^p)$, for inventory, and must consider buying super capacity $q_i^c = (q_{m+1}^c,...,q_{m+n}^c)$ as reserve capacity. During the season (stage 2), after receiving the forecast update ζ, the retailers will adjust their super capacity on hand by trading the unbalanced volume of capacity in the market to secure a total inventory of $q_i = q_i^p + q_i^c$ for the whole season. Let $D_i \geq 0$, $i = m+1,...,m+n$ be the initial demand during the selling season and $\mu_i = E[D_i]$, where $[\cdot]$ is the expected value operator. If there is excess demand $\left(y_i - q_i | \zeta_i\right)^+$ in firm i, a customer will visit another firm j in order to satisfy demand. Thus, the actual demand a firm faces is $D'_j = y_j\left(\zeta_j\right) + \delta_{ij}\left(y_i - q_i | \zeta_i\right)^+$, where $\delta_{ij} \in [0,1]$ is the substitution rate of the excess demand of firm i that will be met by the inventory of firm j. We assume no second substitution and any unmet demand of D'_j is lost. Hence $\sum_{j=m+1}^{m+n} \delta_{ij} \leq 1$. We also assume that all the suppliers do not keep inventories and will deliver the orders by the retailers under forced compliance as the previous set up for two-retailer model.

The supply chains sell the goods to consumers at unit prices $p_i = (p_{m+1},...,p_{m+n})$. The costs of the goods for suppliers are c_{ik} per unit, where $i = m+1,...,m+n$ represents the retailer and $k = 1,...,m$ represents the supplier. The wholesale prices are ω_{ik} per unit in both stages. If a retailer does not satisfy demand, that retailer will incur a goodwill penalty cost g_i per unit. We assume that $p_i > \omega_{ik} > c_{ik}$.

Furthermore, we treat the super capacity of the suppliers in stage 2 as a commodity and assume that all the capacity is of the same quality as in Chapter 3. Super capacity can be traded independently as futures any time before it expires. In stage 1, the suppliers will determine and prepare a volume of super capacity for stage 2 and hedge their super capacity by selling it in the futures market at a price of h per unit. We assume that the price of the super capacity is transparent at any time in the market. Since a retailer may trade super capacity a couple of times in the entire season, we define the weighted average of the super capacity price as $\bar{h}$ per unit. The cost of building super capacity is $s_k = (s_1,...,s_m)$ per unit. The retailers will also buy super capacity futures as hedging to substitute inventory before the selling season. After demand realization, the retailers can place replenishment orders, provided they have the same volume of super capacity on hand; otherwise, they have to buy the inadequate volume of capacity in the market for the replenishment orders. Any participant in this futures market that does not hedge is defined as a speculator. In this chapter we place a restriction on all the players in that they cannot be speculators that merely offset their position at some point before the date set for the futures' delivery. We assume that all the participants intend to make or take delivery of the futures market position, unless they suffer from inaccurate forecasting in which the residual part of the future position will be liquidated at some time prior to expiration. Therefore, the retailers do not necessarily build the optimal inventory in stage

1, but hold a certain volume of super capacity to mix with the physical goods as an inventory portfolio before the season. After early sales have been observed, the retailers decide the replenishment order quantities and then exchange super capacity in the market within a coalition to adjust their inventory positions and reduce the mismatching costs. This design also provides a mechanism for the suppliers to pool their capacity, which reduces waste in the facilities and increases the utilization of each factory involved in super capacity trading.

4.2.1 Supplier's risk and induction

The suppliers may not be keen to develop their short lead-time capacity unless they can benefit from using quick response (Iyer and Bergen 1997). For example, although a retailer may benefit by procuring supplies under improved conditions, the supplier may suffer due to a reduction in the retailer's volume of purchases (Wu 2005). Earlier research has found that adjusting the inventory level downstream by using updated forecasting causes disturbance upstream (Donohue 2000). Hence, suppliers must be provided with an incentive that compensates them for the challenges caused by short lead-time, and given a mechanism for managing their risks.

Our analysis in this chapter focuses on the effects and benefits of capacity pooling by trading super capacity futures. Trading of super capacity is the same as other form of risk pooling for reducing demand variance—such as location pooling, product pooling, and lead-time pooling—and can reduce the uncertainty faced by supply chains, or hedge against the risks of uncertainty in supply chains, particularly when demand is approximated by the normal distribution (Berman *et al.* 2011). In fact, the benefits of capacity pooling always increase with increasing variability of individual random demands under a mean-preserving transformation (Gerchak and He 2003).[1] Thus,

[1] In some situations, risk pooling might lead to inventory anomaly, especially when right skewness of the demand distribution exists (see, e.g., Gerchak and Mossman 1992; Yang and Schrage 2009).

capacity pooling allows the utilization of super capacity to increase in stage 2, enabling the suppliers in the sub-industry to maintain a more stable and continuous production environment.

The suppliers will be induced to join the system of trading super capacity because they can benefit more in this system than when they sell their own reserved capacity as an option to only one retailer. This is important to the suppliers because the lower the utilization of the reserved capacity, the more issues it causes in management, such as labour-related issues caused by excess manpower. Moreover, the suppliers' benefits can be hedged by selling their super capacity as commodity futures in order to protect their revenues from their capacities. Like other commodities, the price and demand for super capacity in stage 2 are uncertain. However, the suppliers can decide on the volume of super capacity independently in stage 1 because the futures market already offers a market price that guarantees their income from building super capacity. It is assumed that supplier k chooses l inputs $x_k = (x_{1k},\ldots,x_{lk})$ to produce ψ units of super capacity by paying $\tau_k = (\tau_{1k},\ldots,\tau_{lk})$ for the inputs x_k. Hence, $s_k(\tau_k,\psi) = \{\min_{xk}[\tau_k\psi : \psi = f(x)]\}$ is the cost function of the supplier in a standard cost minimization problem. If supplier k sells all the super capacity at price h' in stage 2, the capacity profit function is

$$\pi_k = h'\psi - s_k(\tau_k,\psi). \tag{4.1}$$

The revenue is uncertain before the season because the market price of super capacity in stage 2 is unknown. The supplier can keep the same profit level if it sells ψ' units of super capacity as futures at price h in stage 1 and then buys back from the market in stage 2. Therefore, the profit function becomes

$$\pi_k = h'\psi - s_k(\tau_k,\psi) + h\psi' - h'\psi' \tag{4.2}$$

In this case the variance of revenue under price uncertainty is:

$$\begin{aligned} Var(\pi_k) &= Var[h'\psi - s_k(\tau_k,\psi) + h\psi' - h'\psi'] \\ &= (\psi - \psi')^2 Var(h'). \end{aligned} \tag{4.3}$$

If $\psi' = \psi$ is a fully hedged super capacity in stage 2, then $Var(\pi_k) = 0$.

Therefore, the futures market provides a powerful tool for the supplier to manage price uncertainty. On the other hand, considering the first-order conditions of (4.2), we have

$$\frac{\partial \pi}{\partial \psi} = h' - s_k = 0 \text{, hence } h' = s_k ;$$

and $$\frac{\partial \pi}{\partial \psi'} = h - h' = 0 \text{, hence } h = h' .$$

The marginal cost of super capacity is equal to the capacity price. Therefore, both price uncertainty and demand uncertainty do not affect the decision of choosing an output level of super capacity in stage 1. The function of a futures contract can therefore protect the supplier from price uncertainty in the super capacity in the future as they hedge their risk asset. Moreover, the futures price in stage 1 gives the suppliers an important indication of how the market foresees the aggregate demand in the season, helping them determine the output level of super capacity before the season. Obviously, the futures contract cannot help the supplier to make extra profit from the super capacity price arising in stage 2.

4.2.2 The game players

The futures price will be an entry barrier and will determine which retailers are players in the game. Under the principle of profit maximization, the retailers do not pay for the super capacity futures that are either equal to or higher rates than their inventory cost or profit margin; if they do, they will gain nothing or lose in the trade. Therefore, super capacity is a good substitute for their inventories and will reduce their inventory risks if the price is below these limits. However, every retailer would have its own entry level, based on the product prices and their costs. As a result, only a limited number of retailers are entitled to join the game. We develop two sets to define the players in Lemma 4.1 and Lemma 4.2. Please note that a notation $a \wedge b = \min\{a,b\}$, $a,b \in \Re$ is used in the Lemmas.

Lemma 4.1. *A retailer* i *may buy super capacity for hedging in stage 1 if* $i \in A$, $i = m+1,...,m+n$, *where*

$$A = \left\{ H_i^{\alpha} \in A : H_i^{\alpha} = \frac{h}{(p_i - \omega_i) \wedge \omega_i}, \forall H_i^{\alpha} \in (0,1),\ h \geq 0 \right\}.$$

Proof. See Appendix A. □

By the same token, if the price of super capacity in stage 2 is less than the penalty cost of lost sales of a retailer, then such a retailer would buy super capacity in order to reduce loss in the season.

Lemma 4.2. *A retailer* i *may buy super capacity to balance the lost sales in stage 2 if* $i \in B, i = m+1,...,m+n$ *where*

$$B = \left\{ H_i^{\beta} \in B : H_i^{\beta} = \frac{h'}{(p_i - \omega_i) \wedge g_i}, \forall H_i^{\beta} \in [0,1],\ h' \geq 0 \right\},$$ *in which* g_i *is the penalty cost due to lost sales and* h' *is the unit price of super capacity in stage 2.*

Proof. In stage 2, the worst case for the super capacity holder is to surrender the excess quantities unconditionally, i.e., $0 \in H_i^{\beta}$. However, following the same argument in Lemma 4.1, the highest price that the retailer will pay for the super capacity will not be more than the lost sales penalty or profit margin. Hence, $H_i^{\beta} \in [0,1]$ □

In fact, Lemmas 4.1 and 4.2 define the potential retailers that will participate in the game of trading super capacity futures to hedge their inventory positions, and the retailers that will balance their sales loss in stage 2. The interesting point is that all the retailers participating in the super-capacity trading are price takers and there is only one market price at any particular moment. However, the profit margin, wholesale price, and lost sales penalty are more stable, or may even be fixed for the entire season. Hence, members in sets A and B are mobile according to the prices of super capacity. It is clear that the lower the values of H_i^{α} and H_i^{β}, the higher the incentive for a retailer to join the game and gain more profit from capacity trading.

Proposition 4.1. *The retailers joining the trading of super capacity to hedge inventory risk or balance their sales lost will not be worse off.*

Proof. From Lemmas 4.1 and 4.2, only the retailers that are in feasible sets A and/or B will join in the trading of super capacity. However, both sets guarantee that the players will not suffer because the cost of super capacity is always less than their losses. Therefore, the players belonging to sets A and/or B will make more profits from hedging their inventory risks. □

Clearly, trading of super capacity futures is an alternative for the retailer and the supplier to reduce their inventory and capacity risks, respectively. However, a supplier cannot control the cost of developing super capacity under the market price point, or if a retailer does not have a sufficiently high profit margin, they will not be involved in the trading of super capacity.

Proposition 4.2. *The game discriminates against the retailers that have lower profit margins, inventory costs, and lost sales penalties.*

Proof. Lemmas 4.1 and 4.2 exclude the retailers that have lower profit margins, wholesale costs, and lost sales penalties. In fact, a retailer whose values of H_i^{α} and H_i^{β}, are close to zero gain more benefits from trading super capacity. The retailer that has a stronger competitive position is the one that will gain more from this game. However, the elements in A and B will change according to changes in the prices of super capacity in stages 1 and 2, respectively. □

Proposition 4.2 reveals that the players that have lower costs and/or higher value products will benefit from trading super capacity. Thus, this game enables the players that have better operational performance to further enhance their businesses.

4.3 Biform Analysis

We apply a biform game that has two stages: the non-cooperative game in the first stage and the cooperative game in the second stage as an outcome to analyze the model. We consider a set of players $N = \{m+1, ..., m+n\}$ and a finite set Q_i of strategies, $i = m+1,...,m+n$, for each player i. There is also a subset S of N being a coalition. Let $Q = Q_{m+1} \times ... \times Q_{m+n}$, with element q, and let v be a map from Q to the set of maps from 2^N to $\Re$, where $v(q)(\phi) = 0$ for every $q \in Q$. Each player i has confidence index $\varepsilon_i \in [0,1]$. Therefore, $(Q_{m+1},...,Q_{m+n}; v; \varepsilon_{m+1},...,\varepsilon_{m+n})$ is a collection of N-newsvendor biform game.

4.3.1 The non-cooperative stage

In this study the suppliers offer super capacity futures in stage 1 at unit price h. The price depends on the market-clearing force, and neither the seller nor the buyer can dominate it. We assume that the total demand for the futures is determined according to an inverse demand function. Therefore, the futures market provides an opportunity for the suppliers to establish a price for the super capacity in advance of delivery, and they can choose an output level according to the conditions in stage 1, as discussed in Section 4.2.1. However, the suppliers will not become involved in the cooperative game because they do not build any unsold super capacity. Meanwhile, only the retailers that are qualified by the discrimination rule in Proposition 4.2 may join the game.

A non-cooperative game seeks a rational prediction of how individuals interact with one another in an effort to achieve their own goals. The Nash equilibrium is a common solution concept that recommends a strategy for each player whereby no one player can get a better payoff by switching to another strategy available in Q. In other words, if the overall net profit gained from this super capacity game, as well as its ensuing uncertainties, are to be distributed among these players according to a certain rule, then all the players will share the risk in an equilibrium condition.

Proposition 4.3. *Each retail player has a best response q_i^* that maximizes the player's payoff.*

Proof. In our model the retailer's problem is to choose a physical order quantity q^P in stage 1 and to adjust super capacity quantity q^C in stage 2 by trading the unbalanced volume of capacity in the market after demand realization. Hence, the total inventory for the entire season is $q^P + q^C = q$. In a two-stage problem, the retailer that receives a forecast update ζ at the start of the selling season also faces a standard newsvendor problem in stage 2 (Donohue 2000). Let $\pi\left(q \middle| q^P, \zeta\right)$ be the retailer's expected payoff, for each retailer i, i.e.,

$$\pi_i\left(q_i \middle| q_i^P, q_{-i}, \zeta_i\right) = p_i E \min\{q_i, D_i'(\zeta_i)\} - (\omega_i + \overline{h}_i)q_i + \overline{h}_i q_i^p - g_i E[D_i'(\zeta_i) - q_i]^+ . \quad (4.4)$$

Given that $\pi_i\left(q_i \middle| q_i^P, q_{-i}, \zeta_i\right)$ is strictly concave in q_i and letting $S\left(q_i \middle| q_{-i}, \zeta_i\right)$ be the expected sales, we have by the first order conditions

$$\frac{d\pi_i\left(q_i \middle| q_i^P, q_{-i}, \zeta_i\right)}{d(q_i)} = p_i S'\left(q_i \middle| q_{-i}, \zeta_i\right) - \left(\omega_i + \overline{h_i}\right) + g_i S'\left(q_i \middle| q_{-i}, \zeta_i\right) = 0 .$$

Rearranging terms gives $\quad S'\left(q_i \middle| q_{-i}, \zeta_i\right) = \dfrac{\omega_i + \overline{h_i}}{p_i + g_i}$.

Hence, $\quad F\left(q_i^* \middle| q_{-i}, \zeta_i\right) = \dfrac{p_i + g_i - \omega_i - \overline{h_i}}{p_i + g_i}$ or

$$q_i^*\left(q_{-i}, \zeta_i\right) = F^{-1}\left(\frac{p_i + g_i - \omega_i - \overline{h_i}}{p_i + g_i}\right). \quad (4.5)$$

Player i's optimal policy for super capacity quantity is

$$q_i^{C*} = \begin{cases} q_i^* - q_i^{P*} & if\ q_i^* \geq q_i^{P*} \\ 0 & otherwise. \end{cases} \quad (4.6)\Box$$

How the quantity is allocated between physical inventory and super capacity in stage 1 depends on the confidence index ε_i, the preference of player i that is influenced by the private market information, the super capacity futures price, and other factors. We assume a player's preferences satisfy the four standard axioms as stated in Appendix B of Brandenburger and Stuart (2007): order, dominance, continuity, and positive affinity. If the player anticipates the sales quantity to be $[q_i^L, q_i^U]$, where $q_i^L \leq q_i^U$, q_i^L is the lower bound quantity, and $q_i^U \leq q_i^*$ is the upper bound quantity of the sales forecast in stage 1, then we can treat ε_i as a weighting factor, such that player i will hold the physical inventory $q_i^{p*} = \varepsilon_i q_i^U + (1-\varepsilon_i) q_i^L$ and will buy $q_i^{C*} = q_i^* - \{\varepsilon_i q_i^U + (1-\varepsilon_i) q_i^L\}$ according to (4.6). Therefore, the higher ε_i is, the more physical inventory will be held before the selling season. In this case, the player believes it can capture more profit if the market is favourable to it because it only pays a minimum amount for the reserved super capacity but would consume others' residual super capacity in the cooperative stage in order to earn an extra profit, if the profit margin or penalty cost of lost sales is greater than the price of super capacity in stage 2. In other words, players in the biform game will change their actions and preferences based on the differences in the sales forecast. The confidence index is more or less dependent on how much the player believes the sales forecast and how the player analyzes the market in order to anticipate results in the cooperative game. If the player is optimistic, the confidence index will be close to one, while a pessimistic player will have a confidence index close to zero and will buy more super capacity in order to play safe (Brandenburger and Stuart 2007).

We know the strategy q^* that maximizes the payoff of player i is the best response for player i in case player i has no incentive to select any other strategy from q^* when all the other players play q_{-i}^*, i.e.,

$$q_i^*(q_{-i}) = \arg\max_{q_i} \pi_i(q_i, q_{-i}).$$

Since all the players are rational, it is reasonable to believe that they will also follow this best strategy, if this is the only best response for all of them. This means if $q^* = (q_i^*)_{i\in N} \in Q$ is a dominant strategy for all the players, it should have a Nash equilibrium if for all $i \in N$

$$\pi_i(q_i^*, q_{-i}^*) \geq \pi_i(q_i, q_{-i}^*) \quad \forall q_i \in Q.$$

Next we will show the best response mapping will be contracted globally to a fixed point on the whole strategy space by using the contraction mapping principle to prove the game has a unique Nash equilibrium.

Proposition 4.4. *There exists a unique Nash equilibrium to the N-retailer super-capacity game.*

Proof. From Proposition 4.3, we have the best response q_i^*, $i = m+1,...,m+n$, and it is a $R^n \to R^n$ mapping that maximizes each player's payoff. Then the matrix of derivation of the best response functions is defined as:

$$\Delta = \begin{pmatrix} 0 & \frac{\partial f_{m+1}}{\partial q_{m+2}} & \cdots & \frac{\partial f_{m+1}}{\partial q_{m+n}} \\ \frac{\partial f_{m+2}}{\partial q_{m+1}} & 0 & \cdots & \frac{\partial f_{m+2}}{\partial q_{m+n}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_{m+n}}{\partial q_{m+1}} & \frac{\partial f_{m+n}}{\partial q_{m+2}} & \cdots & 0 \end{pmatrix}.$$

Gabay and Moulin (1980) show the uniqueness by the mapping $f(y),\ R^n \to R^n$ is a contraction if and only if the spectral radius of the matrix Δ, $\rho(\Delta)$, is less than one everywhere. However, $\rho(\Delta) = \left\{ \max|\lambda| : \Delta x = \lambda x, x \neq 0 \right\}$ (Horn and Johnson, 1985). This means if λ is an eigenvalue of Δ, then $|\lambda| \leq \rho(\Delta)$ and at least one λ of Δ could be $|\lambda| = \rho(\Delta)$. Let $\|\cdot\|$ be an arbitrary matrix norm, since $\Delta x = \lambda x$, we have

$$|\lambda| \, \|x\| = \|\lambda x\| = \|\Delta x\| \leq \|\Delta\| \, \|x\|,$$

hence

$$|\lambda| = \rho(\Delta) \leq \|\Delta\|. \tag{4.7}$$

Therefore, we need to show $\|\Delta\| < 1$ by verifying that no column sum or row sum of the matrix Δ exceeds one. Then the best response mapping has a unique fixed point as in Netessine and Rudi (2003).

In the matrix of derivatives of the best response functions, $\Delta_{ij} = \partial f_i / \partial q_j, i \neq j, \Delta_{ii} = 0$. Each element in the matrix Δ represents the slope of a best-response function. We define the column sum as $\|\Delta\|_1 = \max_{m+1 \leq j \leq m+n} \sum_{i=m+1}^{m+n} |\Delta_{ij}|$.

Since the slope of the best response function is

$$\Delta_{ij} = \frac{\partial f_i}{\partial q_j} = \frac{\delta_{ji} f_{D'|D_j > q_j}(q_i^*) \Pr(D_j > q_j)}{f(q_i^*)}, \tag{4.8}$$

recalling that δ_{ji} is the substitution rate of the excess demand of firm j that will be substituted by the inventory of firm i, we can tell that the function is monotonic and the slope is between 0 and δ_{ji} in absolute value. Hence $\|\Delta\|_1 \le \sum_{i=m+1}^{m+n} \left| \frac{\partial f_i}{\partial q_j} \right| \le \max_{m+1 \le j} \sum_{i=m+1}^{m+n} \delta_{ji} < 1, \forall j$, so this mapping is contraction and it has a unique Nash equilibrium in the game. □

Therefore, the players can choose their strategies simultaneously in stage 1 to hedge their inventory positions in the selling season. In fact, the trading system also provides a platform for those with different kinds of risk preference to participate in the game using different risk behaviours, without affecting the other parties. From a risk management point of view, this system is much better than a single supply chain coordination strategy, which only allows the shifting of risk along a certain supply chain.

4.3.2 The cooperative stage

In our analysis, we are interested in the outcomes of the cooperation game (N,v) in the cooperative stage. Let (N,v) be a transferable utility (TU) cooperative game $v(z): 2^N \to \Re$. $U = (u_1,\ldots,u_n)$ is the utility function and is assumed to be concave. A function v, assigning a value $v(S)$ for every coalition $S \subseteq N$ with $v(\phi) = 0$, is called a characteristic function. A characteristic function $v(S)$ specifies a total (maximum) value, $\pi(S)$ created by any subset of players (retailers) in N. The retailers are free to form any coalitions that are beneficial to them to obtain their highest utilization. We use the core of the game as our solution concept for the cooperative game. The core of a TU cooperative game (N,v) is the set of payoff vectors

$$C(v) = \left\{ \pi \in \Re^N : \pi(N) = v(N); \pi(S) \ge v(S), \forall S \subseteq N \right\}.$$

Therefore, the core is the set of imputations under which no coalition has a value greater than the sum of its members' payoffs.

Following the approach of Slikker *et al.* (2005), if, after demand realization, a few retailers can form a coalition to cooperate in trading super capacity, then the residual super capacity γ^S in stage 2 can improve the profit of the coalition. Let $h_{ij} \in H$ be the price of one transferred unit of super capacity from i to j, $i, j \in N$, and let Γ^S be a collection of possible residual super capacity vectors of coalition S retailers, defined by:

$$\Gamma^S = \left\{ \gamma^S \in \Re^N : \gamma_i^S = 0,\ \forall\, i \in N \setminus S \text{ and } \gamma_{\mathrm{i}}^{\mathrm{S}} \geq 0,\ \forall i \in S \right\}.$$

Suppose the coalition S before capacity transferring has a residual super capacity vector $\gamma^S \in \Gamma^S$ and they face demand vector $D'^S \in \Re^N$ with $D'^S = 0$ for all $i \in N \setminus S$. In stage 2, m_{ij}^S is the volume of super capacity to be transferred from retailer i to retailer j and no transfer happening is represented by m_{ij}^S for $i = j$. A reallocation matrix of γ^S is

$$M^S = \left\{ M^S \in \Re_+^{N \times N} : m_{ij}^S = 0 \text{ if } i \notin S \text{ or } j \notin S,\ \sum\nolimits_{j \in S} m_{ij}^S(h_{ij}) = \gamma_i^S, \forall i \in S \right\}.$$

The additional profit of the coalition S is then

$$\Pi^S(\gamma^S, D'^S) = \sum\nolimits_{j \in S} (p_j - \omega_j + g_j) \min\left\{ \sum\nolimits_{i \in S} m_{ij}^S(h_{ij}), D_i'^S(\zeta_i) \right\} \geq 0. \quad (4.9)$$

The cost of super capacity is omitted in (4.9) because it is a zero-sum game in S. $\Pi^S > 0$ since $p_j > \omega_j$ and $\Pi^S = 0$ if no transfer of super capacity occurs in S. The expected additional profit of the coalition S depends on its super capacity residual γ^S and the stochastic demand faced by each retailer is $E[\Pi^S(\gamma^S, D'^S)]$ or $\overline{\Pi}^S(\gamma^S, D'^S)$, and the associated game is defined by

$$v(S) = \max_{\gamma \in \Gamma^S} \overline{\Pi}^S\left(\gamma, D'^S\right), \forall\, S \subseteq N. \quad (4.10)$$

If S and T are two disjoint coalitions, they can accomplish at least as much by joining forces as by remaining separate, since the pool of super capacity residual and the need to cover the lost sales will be greater in the situation where two coalitions join together. Hence, the game has the superadditivity property as stated in Lemma 4.3.

Lemma 4.3. *The game* (N, v) *has the superadditivity property that*

$$v(S \cup T) \geq v(S) + v(T) \quad \text{if } S \cap T = \phi$$

Proof. See Appendix A. □

Furthermore, we would like to know whether a stable solution can be found in such a game to allow a certain imputation that will dominate the others. Therefore, the retailers will choose a strategy among themselves to maximize their payoffs.

Proposition 4.5. *The game* (N,v) *has a non-empty core.*

Proof. Let $\{S_{m+1},...,S_{m+k}\}$ be a balanced collection of coalitions of N with balancing weights $\kappa_{m+1},...,\kappa_{m+k}$ such that for every $i \in N$, $\sum_{i \in S_j} \kappa_j = 1$. In addition, let the balanced condition be $\sum_{j=m+1}^{m+k} \kappa_j I_{S_j}(i) \equiv I_N(i)$,

where $I(i) = \begin{cases} 1 & if\ i \in S \\ 0 & otherwise \end{cases}$ is the indicator function of S. We have

$$v(S_j) = \max \overline{\Pi}^{S_j}\left(\gamma, D'^{S_j}\right) \leq \pi(S_i), \text{ where } i \in S_j. \tag{4.11}$$

Multiplying both sides of (4.11) by κ_j and summing from 1 to k, we have

$$\sum_{j=m+1}^{m+k} \kappa_j v(S_j) \leq \sum_{j=m+1}^{m+k} \kappa_j \pi(S_i). \tag{4.12}$$

However, the right-hand side of (4.12) is equal to $\pi(N)$ by being balanced. Hence

$$\sum_{j=m+1}^{m+k} \kappa_j v(S_j) \leq \pi(N) = v(N). \tag{4.13}$$

Since the inequality (4.13) holds, the core of the game (N,v) is non-empty (Shapley 1967). □

Next we discuss the existence of the residual capacity market in stage 2 for the coalitions to be activated in the following proposition.

Proposition 4.6. *There exists a reallocation matrix* M^{*S} *and a residual super capacity vector* γ^{*S} *that maximize the expected additional profit of coalition* S.

Proof. A retailer j should pay $h_{ij} \in H$ to retailer i to obtain one unit of the super capacity residual and the unit price is subject to the market force.

We can identify $h_{ij}^* \in H$ as the maximum value of the super capacity in S by Lemma 4.1 and Proposition 4.5. Since we have $h_{ij}^* \in H$, we can also find the matrix M^{*S}. Similarly, the retailers that have super capacity residual will try to sell it all to maximize their incomes due to $h_{ij}^* \in H$. Therefore, γ^{*S} exists. □

The retailers may have an incentive to cooperate in the game and form coalitions since a core exists whereby a coalition may achieve higher expected profits than the sum of the expected profits of the individual retailers. Such a coalition may be stable due to superadditivity, and the core is the set of undominated imputations among all the strategies in the game. Indeed, there is no group of retailers with a reason to form a coalition and replace the core. Therefore, the reallocation of capacity residual will decrease both excess inventory and demand among different supply chains, ensuring that supply matches demand more effectively in the sub-industry. The trading of super-capacity futures provides a new mechanism for different supply chains to cooperate and reduce the stochastic demand risk.

4.4 Numerical Examples

We consider two examples to illustrate the theoretical results in the cooperative stage. Let $N = \{A, B, C, D, E, F, G\}$ and the additional profit per unit of each player be $\varphi^N = (10, 7, 6, 3, 2, 1, a)$. We do not assign a numerical value to G in φ because we assume this retailer has excess super capacity that may be transferred to retailers *A* - *F*, that have insufficient super capacity in stage 2. Since the price of the super capacity residual is determined by market-clearing force, the capacity price in stage 2 equals the lowest additional profit margin that clears all the capacity residual. We assume that in a good year there will be less super capacity residual than in a bad year, so the unit capacity price will be higher in the good year.

Example 4.1. Let the excess demand of the retailers in a good year be $D'^{G} = (20, 80, 90, 100, 120, 600, 0)$. If the market price of the capacity residual is \$2, then retailer F will not participate in the game because it will lose in the trading. Therefore the coalition is $S = \{A, B, C, D, E, G\}$. In addition, retailer E will not make any money because its excess profit margin is the same as the price of the capacity residual, but the game allows it to offset it penalty cost of lost sales. The total coalition's additional profit is \$2,140. If there is less residual capacity, the price of the capacity will rise, e.g., if it increases to \$6 (which is the profit margin of retailer C), then the additional profit of the coalition will be \$1,600 and now the members in $S = \{A, B, C, G\}$. The additional profit in this case is lower because there is less capacity residual available. We also find that if the capacity price is low, the retailers that buy the capacity will receive a larger share of the profits, but if the price is high, the sellers will receive a larger share. In this example, retail G will gain \$860 when the capacity market price is \$2 and the rest of S will share \$1,260. However, when the capacity market price increases to \$6, G obtains \$1,320, leaving \$280 to the capacity residual buyers.

Example 4.2. Let the excess demand of the retailers in a bad year be $D'^{B} = (0, 20, 50, 80, 100, 200, 0)$. Retailer A will not join the game because it does not have any excess demand. If we assume A does not hold any excess super capacity either, then $S = \{B, C, D, E, F, G\}$. If the capacity price is only \$1, the entire coalition has \$1,080 additional profit. If the price climbs to as high as \$6, the coalition's profit will be reduced to \$440. Once again, if the capacity price is low, e.g., \$1, the retailers that buy the capacity will receive \$630 in total, a larger share of the profits than the seller that gains \$450. In contrast, if the price is high, the sellers will receive a larger share. Let the capacity price become \$6; the seller takes \$420 but the buyers only have \$20 as a whole.

These two examples demonstrate that the amount of additional profit of the coalition depends on the amount of super capacity residual available and the excess demand in the market. However, the additional payoff of each player is split according to the proportion to its added value to the coalition. But we do not know whether the sellers of capacity residual will cover their costs or even make a profit from the game, since the price of the super capacity futures in stage 1 is not known.

4.5 Conclusions

We showed in this chapter that the inventory and capacity risks in newsvendor-type supply chains can be mitigated among many different supply chains selling different products by trading super capacity futures game. Different supply chains can regulate their inventory positions by using a new tool that more accurately matches demand with supply. This game is efficient in both the non-cooperative and cooperative stages, as all the players will reach an equilibrium point. Hence, the sub-industry will increase their aggregate payoff as a result of trading super capacity futures, and they will also improve their performance in matching inventory with stochastic demand for short life-cycle products.

However, the restriction of not allowing speculators in the game that we impose in this chapter is unlikely to occur in real life because it is hard to tell whether a retailer is a hedger or a speculator, particularly when one is playing a dual role. Therefore, we will study a super capacity futures market that is open to both hedgers and speculators in the next chapter.

5 Transfer Mismatching Risk to the Public

In this chapter, we argue that the inventory and supply risks of the newsvendors due to demand uncertainty can be pooled and shared among different supply chains by means of treating reserved capacity as commodities and trading them as *futures*, and *options on futures* to hedge the risks. The risks will be further shared with, and transferred to the public, if speculators are allowed to be involved in the game.

5.1 Introduction

The previous models are further extended in this chapter to allow both hedgers and speculators to participate in the trading of super capacity. In this set up suppliers would still pool their capacities in a sub-industry to reduce demand variability. Retailers may exchange the super capacity residual after realization of the demand is observed in the selling season, similar to the situation outlined in Chapter 4. However, one more financial instrument, 'options on futures' has been incorporated into the models of this chapter to allow the players to hedge both quantity and price risks of trading super capacity. The aims of this chapter are to study whether a unique best response strategy exists for the game players, and whether the whole sub-industry is better off, even with external speculators engaging in the game.

We organize the rest of this chapter as follows: in the next section, we describe the set up of the extended models. The optimal policies of the game are studied in Section 5.3. We move to Section 5.4 and Section 5.5 to examine how risks are transferred from hedgers to risk-neutral speculators and risk-averse speculators, respectively. The optimal payoff for players in the game is identified in both sections. Pareto-improvement under this mechanism is shown as well. Section 5.6 shows that the whole sub-industry will gain extra profit even when speculators are in the game. Section 5.7 presents numerical examples to demonstrate how mismatching risks are transferred to the players in the game under different situations. We conclude the chapter in Section 5.8.

5.2 The Setup for $m \geq 1$ Suppliers, $n \geq 2$ Retailers and $\eta \geq 1$ Speculators

The players of the super capacity futures game in this chapter are basically formed by three types of investors, namely the supplier i, $i = 1,...,m$; the retailer i, $i = m+1,...,m+n$; and the speculator i, $i = m+n+1,...,m+n+\eta$. We define all retailers and suppliers as hedgers. Hedgers intend to make or take delivery of the futures market position, unless they suffer from inaccurate forecasting that the residual part of the futures position will be liquidated at some time prior to expiration. Speculators are the sellers of option contracts and/or they merely offset their positions at some point before the date set for the futures delivery. If a retailer or a supplier plays both roles, as hedger and speculator simultaneously, it is treated as two players in the game. However, if a player changes roles between a hedger and a speculator along the way, in our analysis we still treat this investor as only one player at one point in this game.

This super capacity futures game $Z = (\pi_i, \Lambda_i)_{i \in I}$ with a finite set of players $I = \{1,2,...,m+n+\eta\}$ has a collection of strategy set, $\Lambda_i = \Re_+$, $i \in I$. This game has payoff function $\pi_i : \Lambda_i \times \Re_+ \to \Re$ that is the profit due to super capacity trading that includes the extra income from merchandise sales and the reduction of mismatching costs after exchanging capacity residual among supply chains, in the forms $\pi(q_i^C, q^C)$, where $q_i^C \in \Lambda_i$, $q^C = \sum_1^{m+n+\eta} q_j^C$. The strategies of the players can be aggregated in an additive way and the payoff of each player is a function of the player's own actions. This super capacity futures game is therefore one of the aggregative games.

All the supply chains in our study face stochastic demands, $D_i \geq 0$, $i = m+1,...,m+n$ and the aggregate demand is $D > 0$. We denote $\mu_i = E[D_i]$ and $\sigma_i^2 = Var[D_i]$. Each supply chain demand also has a probability density function $f_i(y)$, and cumulative distribution $F_i(y)$, $i = m+1,...,m+n$. We follow Chapters 3 and 4 to let $\overline{F_i}(y) = 1 - F_i(y)$, and $F_i(0) = 0$ and assume that all distribution functions are continuous, invertible, and double differentiable. Since demand substitution between competitive retailers will not affect the analysis and the results of this set up, we therefore, without loss of generality, assume there is no demand substitution to simplify the notations. In stage 1, retailers determine order quantity, $q_i^P = \left(q_{m+1}^P,...,q_{m+n}^P\right)$ units, to build physical inventory and consider buying super capacity $q_i^C = \left(q_{m+1}^C,...,q_{m+n}^C\right)$ units to reserve a certain capacity to substitute inventory in order to gain a total inventory of $q_i = q_i^P + q_i^C$ units for the whole season. In stage 2, retailers will adjust their super capacity on hand by trading the unbalanced amount of capacity in the market after receiving a forecast update, ζ. The unit retail price to the consumer is $p_i = (p_{m+1},...,p_{m+n})$. Costs of the goods for suppliers are c_{ik} per unit, where $k = 1,...,m$ represents the supplier and $i = m+1,...,m+n$ represents the retailer who places an order with the supplier. ω_{ik} is the wholesale price per unit for both periods. If the retailers do not satisfy demands, they will incur a goodwill penalty cost g_i per unit, and the leftover inventory will have salvage value v_i per unit. We assume $p_i > \omega_{ik} > c_{ik} > 0$ in order to avoid unrealistic and trivial cases.

Let h denote the unit price of super capacity futures in stage 1 and h' in stage 2. We assume the price is transparent any time in the market. There is no extra cost of building super capacity but the supplier needs to prepare super capacity in stage 1. After the demand realization is known, retailers can place replenishment orders, provided they have the same amount of super capacity on hand; otherwise they have to buy an inadequate amount of capacity in the market for the replenishment orders, as described in Chapters 3 and 4.

However, we assume that all the available super capacities for trading in stage 2 come from retailers and speculators, and that suppliers do not hold any unsold capacity on hand in stage 2; otherwise we treat them as speculators.

Meanwhile, retailers and suppliers can also select options on futures as their financial instrument in the market. A call (put) option on super capacity futures gives the buyer the right, but not the obligation, to buy (sell) the super capacity futures contract at the strike price at any time prior to a specific date. Hence, buyers can also hedge the price risk, because they have an opportunity in stage 2 to decide whether or not to exercise the right after the capacity futures spot price at maturity, h' has been observed. Therefore, retailers and suppliers may depend on their forecast and risk preferences for put or call options on the futures simultaneously to hedge both quantity and price risks of the super capacity.

The basic results from financial economics provide the underlying analytical framework for derivative instruments in the commodity markets. The finance literature shows that in the presence of commodity price and yield uncertainty, firms use options as well as futures to moderate price and quantity risks (Sakong *et al.* 1993). In fact, the pricing of futures and futures options correlate highly with the underlying commodity, and they have the same behaviour in the market if interest rate is ignored (Black 1976). Therefore, we denote $\overline{h_i}$ as the net weighted average super capacity price for the capacity committed in stage 1 of an individual investor. Since $\overline{h_i}$ is a combination of different financial instruments and its value will be revealed in stage 2 after the options exercise is determined, it is not exactly the same for all investors who hold the same amount of super capacity even if a market price has occurred.

5.3 The Optimal Policies of the Game

Our study is different from the traditional capacity reservation in two-stage newsvendor models. We assume that a super capacity futures market exists and that anyone, including a speculator, can participate in it. The role of a speculator is to not only take up some of the risk, but also to eliminate the need for a perfect match of the retailers and the suppliers of a risk. We will analyze the transferring and sharing of the risk of stochastic demand among the players in the following sections.

In a two-stage problem, a retailer's problem is to choose a physical order quantity, q_i^P, and to hold a super capacity quantity, q_i^C, in terms of futures or options to build an inventory position as $q_i^P + q_i^C = q_i$ in stage 1. Recalling $(z)^+ = \max(z,0)$, we denote $\pi\left(q \middle| q^P, \zeta\right)$ to be the retailer's random payoff, and then the expected profit for each retailer i can be written as

$$\pi_i\left(q_i \middle| q_i^P, q_{-i}, \zeta_i\right) = p_i \min\{q_i, D_i(\zeta_i)\} - (\omega_i + \bar{h}_i)q_i + \bar{h}_i q_i^P + v_i[q_i^P - D_i(\zeta_i)]^+ \\ - g_i[D_i(\zeta_i) - q_i]^+ . \quad (5.1)$$

where $i = m+1,\ldots,m+n$. The first term is the revenue, the second and third terms are the costs of goods sold. The forth term is the income from leftover inventory, and the last term is the cost of lost sales. Given that $\pi_i\left(q_i \middle| q_i^P, q_{-i}, \zeta_i\right)$ is strictly concave in q_i, and the expected utility function is denoted by $u(\cdot)$. The retailer is assumed to be weakly risk averse and u is a concave function and can be double differentiable. The objective for the retailer is as follows:

$$\max_{q \geq 0} \pi_i(q) = E_i\left\{u_i\left(\pi_i(q \middle| q_i^P, q_{-i}, \zeta_i)\right)\right\}, \quad i = m+1,\ldots,m+n \quad (5.2)$$

where E denotes the expectation operator. If the retailer is risk neutral of $u'' = 0$, then (5.2) yields the classical solution

$$F\left(q_i^* \middle| q_{-i}, \zeta_i\right) = \frac{p_i + g_i - \omega_i - \bar{h}_i}{p_i + g_i - v_i} \quad \text{or} \quad q_i^*\left(q_{-i}, \zeta_i\right) = F^{-1}\left(\frac{p_i + g_i - \omega_i - \bar{h}_i}{p_i + g_i - v_i}\right) \quad (5.3)$$

The player's i optimal policy for super capacity quantity is

$$q_i^{C*} = \begin{cases} q_i^* - q_i^{p*} & \text{if } q_i^* \geq q_i^{p*} \\ 0 & \text{otherwise} \end{cases}, \quad i = m+1,\ldots,m+n$$

Again, we can use the principle in Section 4.3.1 to allocate the total quantity between physical inventory and super capacity in stage 1 in this model, the same as a biform game analysis.

A risk-neutral manager would have a linear utility function such that only expected outcomes matter. In contrast, risk-averse managers have concave utility functions, which reflect their higher sensitivity to a downside than to an upside. Therefore, if $u'' < 0$, then $F(q^* | q_{-1}, \zeta) < \frac{p + g - \omega - \overline{h}}{p + g - v}$. Thus, the risk-averse retailer, who dislikes volatility, orders fewer inventories in general[2].

However, the main purpose for a player to join the super capacity game is to gain extra income $\pi_i(q_i^C)$ from capacity trading and the corresponding profit from selling inventory. The payoff for a retailer i, $i = m+1, ..., m+n$ from the investment of super capacity is

$$\begin{aligned}\pi_i(q_i^C) = &(\omega_i - v_i - \overline{h_i})\left\{q_i^C - \min\left(q_i^C, [D_i(\zeta_i) - q_i^P]^+\right)\right\}^+ \\ &+ (p_i - \omega_i + g_i - h')\min\left\{[D_i(\zeta_i) - q_i]^+, \underline{q_i^C}\right\} + h'\left\{q_i^C - [D_i(\zeta_i) - q_i^P]^+\right\}^+.\end{aligned} \tag{5.4}$$

In (5.4), the first term is the saving of leftover inventory, the second term is extra payoff from buying capacity in the season, and the last term is the payoff from selling excess capacity. $\underline{q_i^C}$ is the amount of capacity residual available in the market for retailer i in stage 2.

For the supplier, the income of super capacity is revenue of selling super capacity and the profit from the replenishment order:

$$\pi_i(q_i^C) = (\overline{h_i} + \omega_i - c_i) q_i^C, \quad i = 1, ..., m \tag{5.5}$$

We assume the super capacity of any supplier in (5.5) would be fully utilized in stage 2 due to the fact that capacity residual will be cleared by market forces.

If there is no speculator involved in the game, the extra payoff for the whole sub-industry from the super capacity game is the sum of the expected savings of leftover inventory and the extra revenue of retailer and suppliers

---

[2] As mentioned in Section 2.3, Wu *et al.* (2009) show that a risk-averse newsvendor may order more than a risk-neutral newsvendor if the lost sales cost is considered with mean-variance trade-off.

from the residual capacity exchange in stage 2, as in the following:

$$\Pi^H(q^C) = \sum_{i=m+1}^{m+n}(\omega_i - v_i)\left\{q_i^C - \min\left(q_i^C, [D_i(\zeta_i) - q_i^P]^+\right)\right\}^+ \\ + (\bar{p} - \bar{c} + \bar{g})\min\left\{\sum_{i=m+1}^{m+n}[D_i(\zeta_i) - q_i]^+, \sum_{i=m+1}^{m+n}\left(q_i^C - [D_i(\zeta_i) - q_i^P]^+\right)^+\right\}. \tag{5.6}$$

Notation with an upper bar means the value is a weighted average in the sub-industry. The first term is the thrift from reducing leftover inventory, and the second term is the sum of extra revenue and the decreasing of lost sales from capacity residual. The cost of super capacity has not occurred in (5.6) because it is a zero sum game between retailers and suppliers.

However, the payoff for speculators that participate in the trading of super capacity is $\pi_i(q_i^C) = (\overline{h_i} - \overline{h'})q_i^C, \ \ i = m+n+1, \ldots, m+n+\eta$ (5.7)

and the total payoff for speculators is $\Pi^S(q^C) = \sum_{i=m+n+1}^{m+n+\eta}\pi_i(q_i^C)$.

Therefore, the total payoff of the whole super capacity game is then

$$\Pi^T(q^C) = \sum_{i=1}^{m+n+\eta}\pi_i(q_i^C). \tag{5.8}$$

Hence we can determine (5.8) by adding (5.4), (5.5) and (5.7) together.

5.4 Risk Transfer to Risk-neutral Speculators

In order to reduce quantity and price risks of super capacity, the risk-averse hedgers need to give up some of their payoffs in terms of risk premium. Speculators, on the other hand, should earn a risk premium for bearing the commodity demand and price volatilities. The trade-off is between risk and expected return. The individual players' attitudes towards risk will affect the sharing of risk among themselves. Here we study two different risk attitudes of speculators, i.e., risk neutrality and risk aversion, in order to understand how hedgers' risks can be transferred and shared by other parties.

Proposition 5.1. *The capacity risk of retailers and suppliers can be transferred totally to others if risk-neutral speculators are involved in the game.*

Proof. Assume that we have one risk-neutral speculator together with n

risk-averse supply chains who join the game. Under the expected utility model, the objective function of the i^{th} player is shown by $E_i u_i\left(\pi_i(q_i^C)\right)$. Therefore, the certainty equivalent of $E_i u_i(\cdot)$ is $E_i\left(\pi_i(q_i^C)\right) - R_i$, $i \in I$, where R_i is the Arrow-Pratt risk premium for the i^{th} player. Following Chavas (2004), we adopt an auxiliary function, benefit function $b_i\left(\pi_i(q_i^C), u_i\right)$ that is derived from a utility function to measure how much benefit that an individual player would be willing to abandon in order to reach the utility level U_i (Luenberger 1992). Let the i^{th} player satisfy $E_i u_i\left(\pi_i(q_i^C)\right) - b_i = U_i$. Therefore, the benefit of the i^{th} player is

$$b_i\left(\pi_i(q_i^C)\right) = E_i\left(\pi_i(q_i^C)\right) - R_i - u_i^{-1}(U), \; i \in I. \tag{5.10}$$

Hence the problem of the whole market is to maximize

$$\Omega(U) = Max\left\{\sum_1^{m+n+\eta}\{E_i\left(\pi_i(q_i^C)\right) - R_i - u_i^{-1}(U_i)\}\right\}. \tag{5.11}$$

If $\Omega(U) = 0$, the maximum aggregate benefit $b_i\left(\pi_i(q_i^C)\right)$ will be redistributed among all the players efficiently. From (5.11), the aggregate benefit will reach the maximum point if $\forall R_i = 0$. However, if the risk-averse suppliers and retailers shift their risks to the speculator, then they will face no risk and have a risk premium equal to zero; hence $\forall R_i = 0$ because the speculator is risk-neutral with $R_\eta = 0$. Therefore the effective transfer for the speculator is:

$\pi_\eta(q_i^C)^* = \Pi^H(q^C) - \sum_1^{m+n} K_i$, where K_i is a non-random number. (5.12)

In other words, the risk-neutral speculator pays each hedger K_i to buy the whole sub-industry risky payoff. Hence, K_i, $i = 1,\ldots,m+n$ is the monetary term of risk premium. If we assume the profit of the speculator equal to or greater than zero is a fair deal for both sides, since $\Omega(U) = 0$ and $\forall R_i = 0$, we gain $u_i^{-1}(U_i) = E_i\left(\pi_i(q_i^C)\right)$ from (5.11), then for the individual hedger its optimal return is $\pi_i(q_i^C)^* = K_i \le E_i\left(\pi_i(q_i^C)\right)$, $i = 1,\ldots,m+n$. □

Therefore, the hedgers can sell and transfer their private risks totally to the speculator by receiving K_i which is equal to or less than the expected value of their own payoffs. The transactions appear to mean either that the hedger and speculator of it have different expectations about the futures, or that they have different preferences in relation to the likely return and risk involved in holding the super capacity futures. In fact, the risk-neutral speculator plays the role of an insurance agent to bear all the risks in this case, but can enjoy a positive expected value in the long run. Hence, the game engages speculators.

5.5 Risk Sharing with Risk-averse Speculators

If all speculators as well as the hedgers are risk-averse, then we have to scrutinize how the commodity market provides a mechanism for them to share the hedging risk among themselves. Assume that each of the players wishes to invest $q_i^C \in \Lambda_i$ in the commodity market either to hedge their capacity risk or to make a profit from the futures trade. We wish to identify an optimal strategy and payoff rule for each player in Z.

Proposition 5.2. *Each player has a best response $q_i^{C*} \in \Lambda_i$, $i \in I$ to invest in the super capacity futures game that an optimal risk-sharing rule exists to maximize the player's payoff from the investment.*

Proof. Denote $\pi_i^d = \pi_i(q_i^C)$ as the decision rule giving the net profit made to the i^{th} player from the commodity market. In section 5.4 we define the certainty equivalent of $E_i u_i(\cdot)$ from the objective function of the i^{th} player as $E_i\left(\pi_i(q_i^C)\right) - R_i$, $i \in I$. If the entire super capacity futures market generates an extra net profit $\Pi^T(q^C)$, then $\sum_{i=1}^{m+n+\eta} \pi_i(q^C) \le \Pi^T(q^C)$. Let the i^{th} player have an ex-ante utility function of $u_i(\pi_i^d)$ and the benefit function of $b_i\left(\pi_i(q_i^C), u_i\right)$, which satisfies $u_i(\pi_i^d - b_i) = U_i$, $i \in I$ and U_i is the individual player utility function representing individual risk preference. Hence, the whole market net profit from the capacity futures trading is to maximize

$$\Omega(U) = \sum_{i=1}^{m+n+\eta} b_i(\pi_i^d, U_i), \text{ where } U = (U_1, \ldots, U_{m+n}, U_{m+n+1}, \ldots, U_{m+n+\eta}), \quad (5.13)$$

subject to $\sum_{i=1}^{m+n+\eta} \pi_i(q^C) \le \Pi^T(q^C)$.

Assume a linear distribution rule $\pi_i(q_i^C) = \gamma_i + \delta_i \Pi^T(q^C)$ to be adopted among the players, then $E(\pi_i) = E[\pi_i(q_i^C)] = \gamma_i + \delta_i \Pi^T(q^C)$,

and $$Var(\pi_i) = Var[\pi_i(q_i^C)] = \delta_i^2 Var[\Pi^T(q^C)].$$

Define $r_i \equiv \dfrac{-U_i''}{U_i'}$ as the Arrow-Pratt coefficient of the i^{th} player, thus an individual risk premium is $R_i = \dfrac{r_i}{2} Var(\pi_i)$, $r_i > 0$.

Assume that the i^{th} player has a mean variance preference function

$$u_i(\pi_i^d) = E(\pi_i) - \tfrac{1}{2} r_i Var(\pi_i).$$

Therefore, $$b_i = \gamma_i + \delta_i E\Pi^T(q^C) - \tfrac{1}{2} r_i \delta_i^2 Var[\Pi^T(q^C)] - U_i. \quad (5.14)$$

Hence, the problem of the players is to maximize

$$\Omega(U,r) = \sum_{i=1}^{m+n+\eta} (\gamma_i + \delta_i E\Pi^T(q^C) - \tfrac{1}{2} r_i \delta_i^2 Var\left[\Pi^T(q^C)\right] - U_i), \quad (5.15)$$

subject to $$\sum_{i=1}^{m+n+\eta} (\gamma_i + \delta_i E\Pi^T(q^C)) \le \Pi^T(q^C).$$

Using the Lagrangian for the maximization problem, the first-order conditions for the constraint become

$$\sum_{i=1}^{m+n+\eta} \gamma_i + \Pi^T(q^C)\left\{\sum_{i=1}^{m+n+\eta} \delta_i - 1\right\} = 0. \quad (5.16)$$

This can hold for all q^C only if $\sum_{i=1}^{m+n+\eta} \gamma_i = 0$ and $\sum_{i=1}^{m+n+\eta} \delta_i = 1$.

From (5.14) we have $$\frac{\partial b_i}{\partial \delta_i} = E\Pi^T(q^C) - r_i \delta_i Var\left[\Pi^T(q^C)\right] = 0. \quad (5.17)$$

Hence we obtain $$\delta_i = \frac{E\Pi^T(q^C)}{r_i Var\left[\Pi^T(q^C)\right]}, \text{ for } r_i > 0. \quad (5.18)$$

Therefore, $$\sum_{i=1}^{m+n+\eta} \delta_i = \sum_{i=1}^{m+n+\eta} \left\{ \frac{E\Pi^T(q^C)}{r_i Var\left[\Pi^T(q^C)\right]} \right\} = 1. \quad (5.19)$$

Based upon the result of (5.19), then (5.18) becomes

$$\delta_i^* = \frac{1}{r_i} \Bigg/ \left[\sum_{i=1}^{m+n+\eta} \left(\frac{1}{r_i}\right)\right] \quad (5.20)$$

Furthermore, we substitute (5.20) into (5.17) giving the sum of the players' payoffs equal to

$$E\Pi^T(q^C) - \frac{1}{\sum_i 1/r_i} Var\left[\Pi^T(q^C)\right].$$

Therefore, to maximize the total payoff will give an optimal action

$$q^{C*} = \arg\max_{q} \left[E\Pi^T(q^c) - \frac{1}{\sum_i 1/r_i} Var\left(\Pi^T(q^C)\right) \right]; \qquad (5.21)$$

and the optimal payoff to the i^{th} player from the capacity futures game is

$$\pi_i(q^C)^* = \gamma_i^* + \delta_i^* \Pi^T(q^c), \text{ where } \sum_i \gamma_i = 0;$$

or

$$\pi_i(q^C)^* = \gamma_i^* + \frac{1/r_i}{\sum_{i=1}^{m+n+\eta}\left(1/r_i\right)} \Pi^T(q^c). \qquad (5.22)\square$$

Therefore, each player has a best response strategy and an optimal payoff in the super capacity game to allow risk sharing of demand uncertainty occurring among the sub-industry and the public under this capacity pooling system. Moreover, the involvement of capacity trading from the public and speculators does not impinge upon any part of the original benefit of each supply chain. From a risk management point of view, this system is much better than supply chain coordination strategy that only allows the shifting of risk along a certain supply chain. We can apply comparative statics analysis on the optimal sharing rule to further understand the characteristic of the risk sharing in corollaries 5.1 and 5.2 below.

Corollary 5.1. *Higher degree of risk aversion of any individual player will produce a negative impact on the total extra payoff of the whole sub-industry from the super capacity game.*

Proof. See Appendix A. □

Corollary 5.2. *Given the optimal payoff* $\pi_i(q^C)^*$ *in Proposition 5.2, then a higher degree of risk aversion by the* i^{th} *player decreases the profit share to this investor, but the share to other investors in the group is unaffected.*

Proof. See Appendix A. □

From Proposition 5.2, Corollaries 5.1 and 5.2, we know that the trading of capacity is an efficient allocation of risk since the constraint is always binding. The trading system provides a platform for those with different kinds of risk preferences to participate in the game using different risk behaviour without affecting other parties. Therefore the capacity market plays a role in allowing for risk spreading to the public from any supply chain. The inventory risk and capacity risk are therefore shared by the public. Hence we obtain again Pareto-improvement under this mechanism. However, the next question is whether the equilibrium is unique.

Proposition 5.3. *There exists a unique Nash equilibrium to the super capacity futures game.*

Proof. See Appendix A. □

Therefore, we gain the solution concept of the game for both hedgers and speculators that are risk-averse.

5.6 The Gain of Sub-industry

One of our concerns is whether the net payoff of speculators will take all the profits in Z from the sub-industry. If so, there is no point to involve outsiders or even to allow a capacity futures market to exist. We analyze a general result of sub-industry players in the following proposition.

Proposition 5.4. *The extra gain of sub-industry from the super capacity game is greater than zero.*

Proof. Assume that all players are rational and will try to maximize their own profit at any time. For the supplier, both price uncertainty and demand uncertainty of super capacity in stage 2 will not make them worse off because the price of super capacity futures in stage 1 guarantees they will gain some extra profits by selling super capacity; otherwise they will not become involved in the capacity trading as discussed in Section 4.2.2.

On the other hand, a retailer i may buy super capacity futures and options in stage 1 if $(p_i - \omega_i) \wedge \omega_i > h$ to ensure no loss in the deal. Similarly, in stage 2, retailer i will buy capacity residual from the market to balance lost sales if $(p_i - \omega_i) \wedge g_i \geq h'$. However, a retailer will sell all the capacity residual on hand to increase his payoff. Therefore, the sub-industry payoff according to the optimal sharing rule,

$$\sum_{i=1}^{m+n} \pi_i (q_i^C)^* = \sum_{i=1}^{m+n} \left\{ \gamma_i^* + \frac{1/r_i}{\sum_{j=1}^{m+n+\eta} \left(1/r_j \right)} \Pi^T (q^C) \right\} > 0 \,. \qquad \square$$

Obviously, h plays a role in allocating the aggregate payoff that will reflect the degree of risk premium of players in Z. However, h also impacts the payoff sharing between suppliers and retailers in each supply chain.

5.7 Numerical Examples

Some of the behaviour and results of transfer mismatching risk to the public by the super capacity trading game would be easier to observe from numerical examples. We assume there are sufficient spectators to participate in the trading of super capacity in any stage to buy or sell futures and to sell option contracts as option writers. Since the set up includes any traders and uses two financial instruments, it is too complicated to consider all situations simultaneously. Therefore, we follow the main theme of this thesis to focus on how the risks are to be transferred from aggregate supplier hedger, and from aggregate retailer hedger separately to the sub-industry and the public as follows.

5.7.1 The performance of aggregated supplier

Suppliers who want to hedge the risk of super capacity would sell their super capacity futures to any players in stage 1. If they want to further insure the price risk of the futures, they may buy put options on futures to own the right to sell the super capacity at a specific price in stage 2 to share some of the benefit from the price rising. Therefore, they would mix these two instruments to match their risk preferences. However, if a certain supplier does not want to hedge super capacity in stage 1 but holds them until stage 2 for any reason, that supplier is in fact playing two roles. This is equivalent to a situation in which a supplier sells all the super capacity as futures to itself in stage 1, and then plays a speculator role to hold the futures and sell them in stage 2. Actually all suppliers hedge their super capacity in stage 1 with this viewpoint. We assume all the super capacity will be utilized by retailers in stage 2 due to the fact that capacity residual can be transferred among retailers and speculators. Two situations are considered in this section: the super capacity futures price is lower in stage 1 than the spot price in stage 2 and vice versa to demonstrate the benefit of the hedgers.

5.7.1.1 Super capacity price is lower in stage 1 than stage 2

If the price of futures of super capacity is \$1.20/unit in stage 1. The supplier would pay an option cost or premium to the option seller of \$0.2/unit to obtain a put option that would allow it to own a right, but not the obligation to sell super capacity in stage 2 at strike price of \$1.80/unit. If the spot price in stage 2 is \$2.2/unit, then the supplier who owns the put options will surrender exercising the option contract and will sell its super capacity to the spot market.

Assuming the aggregate super capacity that will be offered by all suppliers in the sub-industry is 100 units, we look at two combinations of the trading: 70% futures mixed with 30% options, and 30% futures mixed with 70% options as shown in Table 5.1.

Since the price of super capacity is higher in stage 2, no supplier will proceed with the put option contracts to sell any super capacity to the option seller. We notice that in both mixtures of the instruments, suppliers will make higher income than using only futures as a single tool. Obviously, a supplier can share more benefit in case the actual demand in the season is higher than expected. Moreover, on top of the super capacity income, a supplier can further make extra profit from the replenishment orders that have come from the capacity residual in stage 2.

Table 5.1: Income of suppliers from trading supplier capacity with two mixtures of financial instruments in case the capacity price is higher in stage 2

<table>
<tr><th rowspan="2"></th><th colspan="4">High Futures Proportion</th><th colspan="4">Low Futures Proportion</th></tr>
<tr><td>Futures</td><td>Options</td><td>Spot</td><td rowspan="7"></td><td>Futures</td><td>Options</td><td>Spot</td><td rowspan="7"></td></tr>
<tr><td>Quantity engaged in stage 1</td><td>70</td><td>30</td><td>0</td><td>30</td><td>70</td><td>0</td></tr>
<tr><td>Potential amount selling in stage 2</td><td>0</td><td>0</td><td>30</td><td>0</td><td>0</td><td>70</td></tr>
<tr><td>Unit price</td><td>$1.20</td><td>$1.80</td><td>$2.20</td><td>$1.20</td><td>$1.80</td><td>$2.20</td></tr>
<tr><td>Premium</td><td>-</td><td>$0.2</td><td>-</td><td>-</td><td>$0.2</td><td>-</td></tr>
<tr><td>Quantity settled in stage 2</td><td>70</td><td>30</td><td>0</td><td>30</td><td>70</td><td>0</td></tr>
<tr><td>Income</td><td>$84</td><td>-$6</td><td>66</td><td>$36</td><td>-$14</td><td>$154</td></tr>
<tr><td>Total gain from super capacity</td><td colspan="3"></td><td>$144</td><td colspan="3"></td><td>$176</td></tr>
</table>

5.7.1.2 Super capacity price is higher in stage 1 than stage 2

We use the same set up in Section 5.7.1.1, but the price of super capacity futures is at a higher level in stage 1. We assume the actual demand is lower than the forecast, and retailers who have wider profit margin or higher lost sales do not need all the reserved capacity in stage 2. Thus, the capacity spot price should be lowered to coincide with those lower profit margin retailers. It can be seen from Table 5.2 that the options on the futures do provide some protection to hedge the price risk, and the supplier can keep a high portion of the income from super capacity. We would observe that the aggregated income of the supplier is higher than purely selling capacity by futures.

The results of these two scenarios reveal that futures and options on futures can offer sufficient protection of both quantity and price for the suppliers. Under this mechanism, suppliers can share profit according to the value and contribution of their short lead-time capacities to reduce mismatching. However, their actual payoffs will be affected by the combination of the instruments that are influenced by the risk preference of the individual supplier.

Table 5.2: Income of suppliers from trading supplier capacity with two mixtures of financial instruments in case the capacity price is higher in stage 1

	High Futures Proportion				**Low Futures Proportion**			
	Futures	Options	Spot		Futures	Options	Spot	
Quantity engaged in stage 1	70	30	0		30	70	0	
Potential amount selling in stage 2	0	0	30		0	0	70	
Unit price	$2.20	$2.50	$1.2		$2.20	$2.50	$1.20	
Premium	-	$0.5	-		-	$0.5	-	
Quantity settled in stage 2	70	30	0		30	70	0	
Income	$154	$60	0		$66	$140	0	
Total gain from super capacity trade				$214				$206

5.7.2 The results of aggregated retailer

The situation of retailers is much more complicated than suppliers, because speculators will take the same actions as retailers in exchanging of capacity in both stages. Retailers who participate in the game to buy super capacity just for their retail business in stage 1 are hedgers. If any retailer holds super capacity in stage 1 for selling at any point along the horizon in order to make a profit from the price movement, we treat it as a speculator.

We also consider two statuses for the retailers as in Section 5.7.1; the first is that super capacity price is lower in stage 1, but has ended up higher in stage 2, while the second is the reverse.

5.7.2.1 Super capacity price is lower in stage 1 than stage 2

We assume there are 100 units of super capacity in the market, of which speculators hold 20 units. Let the price of capacity futures be $1.20 per unit in stage 1. The strike price is $1.30/unit for a call option that gives the holder the right, but not the obligation, to buy the super capacity at the strike price in stage 2. The premium costs $0.10/unit. Retailers who hold a call option

Table 5.3: Total payment of retailers in trading supplier capacity with two mixtures of financial instruments in case the capacity price is higher in stage 2

	High Futures Proportion				Low Futures Proportion			
	Futures	Options	Spot		Futures	Options	Spot	
Quantity engaged by hedgers in stage 1	55	25	0		25	55	0	
Unit price	$1.20	$1.30	$2.20		$1.20	$1.30	$2.20	
Premium	-	$0.1	-		-	$0.1	-	
Quantity settled in stage 2	50	20	30		25	45	30	
Cost of super capacity	$66	$28.5	$66		$30	$64	$66	
Income of selling capacity residual	$11	0	0		0	0	0	
Total paid for super capacity				$149.5				$160

contract because of a belief that the price may drop in stage 2 do not want to explore any risk to only buy capacity from market in stage 2. We examine the situation in which the retail business turns out to have a good year, and then the price of super capacity rises to $2.20/unit. Some retailers may still have some capacity residual for the market because they might own too many futures, or their business is not as good as others even in a good year. According to Table 5.3, retailers as a whole pay more than the cost of futures because they have paid a premium for price risk and some of the reserved capacity was not hedged. Hence retailers who buy super capacity in stage 2 share some of their profit from the saving of mismatching cost and the extra profit to suppliers with speculators and those retailers who have capacity residual. However, the sub-industry does not need to pay the market price in stage 2 due to protection from financial instruments.

5.7.2.2 Super capacity price is higher in stage 1 than stage 2

Assume the macro environment is favourable for business before the selling season and the retailer is willing to pay a higher super capacity price. Let the price of futures be \$2.20/unit, and the options premium be \$0.3/unit with a strike price of \$1.60/unit in stage 1. We assume again the whole market has 100 units of super capacity and speculators hold 20 units of them. If a natural disaster or any serious crisis that occurs in the beginning of stage 2 creates enough negative impact to shrink the aggregated demand, then the spot capacity price drops accordingly. We assume the capacity has the value of \$1.20/unit only in stage 2. Under this design, it can be seen from Table 5.4 that the whole sub-industry pays less than the futures, because the options take the risk of pricing that all call option contracts holders do not exercise the right to buy any capacity at a higher price point. Therefore, if the retailer as a whole owns a higher portion of call option contract than futures, the net total amount pay for the super capacity market is lower than the high futures portion. In contrast, the retailer will pay more if the price is moved from a lower end to a higher end.

From the above examples, we have found that the new mechanism offers an efficient protection for both retailers and suppliers as hedgers. The extra payoff of the whole sub-industry from the super capacity game in Eq. (5.6) is the main source to attract investors to join the game. The different combinations of futures and options on futures provide any degree of risk

Table 5.4: Total payment of retailers in trading supplier capacity with two mixtures of financial instruments in case the capacity price is higher in stage 1

	High Futures Proportion				**Low Futures Proportion**			
	Futures	Options	Spot		Futures	Options	Spot	
Quantity engaged by hedgers in stage 1	55	25	0		25	55	0	
Unit price	\$2.20	\$1.60	\$1.20		\$2.20	\$1.60	\$1.20	
Premium	-	\$0.3	-		-	\$0.3	-	
Quantity settled in stage 2	45	0	55		25	0	75	
Cost of super capacity	\$121	\$7.5	\$66		\$55	\$16.5	\$90	
Income of selling capacity residual	\$12	0	0		0	0	0	
Total paid for super capacity				\$182.5				\$161.5

exposure for any individual player in the game. Therefore, both supplier and retailer can protect themselves by transferring all or part of the mismatching risk to others by the financial instruments. We should note that the capacity futures market depends not only on the existence of uncertainty about the future, but on the existence of different utility functions as to what will happen, and individual preferences regarding risk and return. Hence there is an opportunity for speculators in the game to make a profit.

However, we are only concerned with whether speculators as a whole will gain more than the whole sub-industry in our study, but not with their payoffs in any particular situations. Therefore, we will not provide any numerical example to observe and explain their results.

5.8 Conclusions

The existence of the unique Nash equilibrium demonstrates that this new mechanism is an efficient means to let single supply chain risks be shared with, or be transferred to, the other supply chains and even to the public who are in the game. Each investor will share the payoff according to their own degree of risk aversion, which does not affect profit share of other investors. Our results in this chapter indicate that this new market-based risk transfer mechanism combines operational and financial hedging strategies which provide industry with a new way of meeting demand more efficiently to achieve a Pareto-improvement.

However, trading of super capacity is difficult to process if there is a lack of a common capacity unit for different products. We therefore develop a new measurement unit for super capacity of different products among different facilitators in the next chapter.

6 Capacity Unit for Trading

In order to allow the super capacity trading among different supply chains to be processed, a capacity unit that is common to the mix of products encountered should be created. We develop in this chapter a time-based, value-added capacity measurement model that is an output orientated input measure to satisfy the need. The capacity unit rates not only the overall production process, but also each product produced by the process. Our model measures process capacity based on our new concept that a process contains only value-added and non-value-added work, but no waste. We verify that value-added capacity is fully efficient. Therefore, its output level depends on the value-added process time that is required by each product.

6.1 Introduction

One predominant challenge in trading super capacities of different suppliers in the market is that there should be a common trading unit through which different outputs can be inferred, otherwise a market price is hard to be determined for such a commodity. Assuming the suppliers in a sub-industry adopt a prevailing technology, the same amount of capacity will produce various output quantities by different suppliers due to the inconsistent, manifold efficiencies that appear in different production lines. Our problem is whether there is a metric to measure super capacity such that retailers can convert it into their own physical product quantities without disturbance of capacity effectiveness. Therefore, the purpose of this chapter is to develop and define a capacity unit that is common to different products in a sub-industry. The capacity unit selected also allows super capacity trading between different supply chains to share the same market price.

As mentioned in Chapter 1, capacity is the maximum quantity of output per unit in a given amount of time that a stock of plant and equipment is capable of completing, provided that the availability of variable factors is not restricted (De Leeuw 1962). Van Mieghem (2003) defines capacity as a measure of processing abilities and limitations that stem from the scarcity of various processing resources, which are represented as a vector of stocks of various processing resources.

Capacity measurement has been studied since the early days of the last century, and has evolved into three streams of studies in the literature. The first stream is from an engineering point of view to measure capacity; the second direction is to consider the capacity measurement in cost accounting; and the third stream applies capacity measurement in economic terms.

6.1.1 Engineering approach

Measurement of industrial production can be constructed by output measures in physical units and input measures to the production process, from where output is inferred. Input factors include machine, people, and factory that are able to produce or do what the customer requires. In general, capacity is a measurement of time as an input measure (Church 1917). Output measures of capacity are better adapted to different processes, or the supplier provides a relatively small number of standard products. However, once the variety in the product mix increases, output-based capacity measures become less useful (Krajewski and Ritzman 2007).

Cachon & Terwiesch (2009) define capacity as the maximum flow rate that can be supported by a resource. The flow rate is the minimum of available input, demand and process capacity. Blackstone (1989) uses Wight's funnel to describe the capacity as the rate at which work is withdrawn from the system. He suggests capacity is equal to the result of production time available, efficiency and utilization. Management has to control the amount of work in a system by both input and output rates to maximize the output of capacity.

Obviously, output based capacity measures become less useful in our study because different products represent various amounts of capacity in production. It is inconvenient to convert output units among different products during the capacity trading process. Therefore, we will only focus on the input measures to develop a capacity unit on an equivalent basis that will be suitable for capacity trading.

6.1.2 Cost accounting approach

In the early of the twentieth century, papers and debates addressed capacity cost management. The key arguments took place between A. H. Church and H. L. Gantt on how to allocate the cost of idle capacity (McNair and Vangermeersch 1998). In fact, different allocation methods of the idle capacity cost come from different capacity measurements, whether 'normal' or 'practical' capacity is used (Watts *et al.* 2009). Normal capacity is the average of utilization of machinery over a 3-5 years period. But a practical capacity is the maximum amount of capacity reduced for unavoidable downtime. The problem of absorption of idle capacity cost, however, is still a problem that remains until today. The current attentions resemble the early twentieth century debates about the proper treatment of capacity, particularly the management of excess capacity (Hertenstein *et al.* 2006).

The Consortium for Advanced Manufacturing – International (CAM-I) based on the concept of H. L. Gantt to build a CAM-I Capacity Model in the 1980s' (Klammer 1996). This model classifies the 'Rated Capacity' in terms of 'Idle', 'Non-productive' and 'Productive'. Rated capacity is the maximum theoretical capacity, and it uses a time measure. Idle capacity includes marketable capacity that is usable, not marketable capacity that is excess, and off-limits capacity. Non-productive capacity is capacity not in a productive state such as standby, yield loss, setup and maintenance. Productivity capacity is capacity used to produce output or provide service. The advantage of CAM-I is to make the costs of non-productive capacity transparent to allow managers to take action in reducing these costs. In fact, the efficient use of capacity means the organization is positioned to reduce cost by managing capacity (Yu-Lee 2003).

We will make use of CAM-I model to further define a capacity deployment in Section 6.4 to develop our super capacity trading unit.

6.1.3 Economic approach

At a macroeconomic level, capacity measurement and rate of capacity utilization were discussed significantly between the 1950's and 1970's. Zabel (1955) summarized the measurements of capacity into two major categories, maximum output and optimum stock levels. In maximum output method the capacity is determined only if the conditions under which the output is produced are completely specified. The optimum stock levels method is useful for studies of investment behaviour. Zabel further studied the measures of industry capacity and defined three types of technological measures; namely, rated capacity, theoretical capacity and practical capacity (Zabel 1956). Rated capacity is output potential of machinery during a full period of 24 hours running time, without interrupted flow of labour and raw materials, and no machine downtime. Theoretical capacity refers to maximum output of working days in a given period, with the existing equipment and uninterrupted operation. Practical capacity adjusts the theoretical capacity by deducting the maintenance and repair time and considers the usual number of shifts worked in the industry.

Indeed, economic aspects of capacity measurement involve the production function and cost considerations. In a perfect competition market, with a full equilibrium position with zero profit, at the level of operation equating marginal cost with the marginal revenue, the output is equal to the full capacity because the point of minimum average cost may represent full capacity. Chamberlin (1948) implies that imperfect competition causes inefficiency in a firm, and hence incurs excess capacity. In fact, Chamberlin greatly furthers the study of capacity by linking the economic considerations of cost to the determining of capacity (Klein 1960).

However, definitions and measurement of capacity continued to be argued over in the middle of the last century, and some discrepancies existed in the major measures used by different organizations. In that period, four major sources of data were used to estimate capacity in the United States. Wharton School measured trend lines drawn between cyclical output peaks. It was hard to identify that output at each cyclical peak had the same degree of capacity utilization. McGraw-Hill gathered questionnaire data for two measures that were criticized for being too subjective and not being a random sampling. The National Industrial Conference Board used deflated capital stock as a measure of capacity that was based on an assumption that capacity bore a constant ratio to capital stock. (De Leeuw 1962; Phillips 1963).

Furthermore, Klein and Preston (1967) developed a measure of capacity and capacity utilization at the industry level from production functions to adjust the component of Wharton Index to reduce its bias. In their model, the full capacity real output was a function of man-hours employed and capital utilized together with a technical change factor. They added a disturbance factor to the full capacity function to work out an actual output.

In the next section, we examine components of a process that consist of only value-added and non-value-added work. We also provide a new concept of waste in the production process. Section 6.3 discusses fully efficient value-added input. Section 6.4 defines a capacity deployment according to the results in Sections 6.2 and 6.3. We therefore create a new super capacity trading unit in Section 6.5. Lastly, Section 6.6 is the chapter conclusion.

6.2 Capacity Management and Waste

One of the main focuses on capacity management in cost accounting is to identify and eliminate wasted resources in a system. Waste-based capacity cost management approach focuses on diagnosing capacity utilization, then changing polices to improve the capacity utilization (McNair and Vangermeersch 1998). They determine five sources of waste in capacity utilization, and point out hidden costs of capacity in unbalanced production, and obsolescence fixed assets on top of the non-used capacity. CAM-I capacity model takes a similar approach to deploy rated capacity, the theoretical maximum rate of non-stop production of a system in a given period. As we have mentioned in Section 6.1.2, rated capacity is divided into three components: idle capacity, non-productive capacity and productive capacity. Konopka (1995) has developed CUBES, the Capacity Utilization Bottleneck Efficiency System to deploy capacity into six categories, similar to CAM-I. However, CUBES emphasizes tool efficiency and identifies speed losses in tools and batch size in production by comparing the theoretical/plan and actual output data. CUBES determines that waste in productive capacity is the key distinction from CAM-I. Its concept is close to Toyota Production System (TPS) in searching for wastes in production.

TPS has revealed that the basic rule for improving productivity is to eliminate waste, which increases the proportion of value-added work in any operation process (Ohno 1998). The implementation of this simple rule usually involves the application of the famous “Seven Wastes” identified by TPS. We, however, propose that waste may not exist in a process and the concept of waste existing in a process is a blind spot in Just-in-time (JIT) and TPS. We provide the background and a new look at waste and value-added work that supports our argument in Appendix B of this dissertation. We will explain our concept in the following.

Discussion about the existence of waste in production processes can be studied in terms of the relationship between processing time and the value created by production, particularly if cycle time and/or lead-time reduction are taken into account. A production line comprises a sequence of processes. Each process has its own cycle time. The cycle time is the period required to complete one cycle of an operation or process. The value of a product is generated and accumulated along the production process until the required end point within a lead-time. The lead-time is defined as the horizon between when an order is placed and when the output arrives at an assigned point. Time as the unit of measurement, however, is arbitrary. When an order is placed, it is assumed that both the customer and the supplier know the requested value of the product itself, but the supplier may not be aware of how short the shortest lead-time is. In other words, the technical efficiency that would lead to the shortest lead-time is unknown to them. We will discuss the components of production function and waste in the next sub-section.

6.2.1 Work classification and waste

Considering the production process as a function of time, we assume that a series of production processes p_i forms a meaningful value-adding production function *P*. Let $p_i : [t_i, t_{i+1}] \to \Re$ be a piece of straight line segment, $p_i(t) = c_i t + d_i$, $t_i \le t \le t_{i+1}$, on which exist the points $\{ t_i : i = 0,1,...,m \}$ that satisfy

$$a = t_0 < t_1 < ... < t_i < ... < t_m = b,$$

where t_i are knots in *P*, and *P* is piecewise linear on $[a,b]$ of degree 1. We define *P*(*t*) as a continuous function, so $p_{i-1}(t_i) = p_i(t_i)$, or $c_{i-1}t_i + d_{i-1} = c_i t_i + d_i$, for $i = 1,2,...,m$ and *P(t)* is entirely determined by its nodal values, $p(t_i)$ and each knot $t_i, i = 0,1,...,m$ is the end point of the corresponding process.

We follow Gockenbach (2010) in that the standard basis for the space of all the continuous piecewise linear functions is $\{\phi_0, \phi_1, ..., \phi_{m-1}\}$, where

$$\phi_j(t_i) = \begin{cases} 1, & i = j, \\ 0, & i \ne j. \end{cases} \tag{6.1}$$

Therefore, the production function satisfies

$$P(t) = \sum_{i=1}^{m} p_{i-1}(t_i)\phi_{i-1}(t). \tag{6.2}$$

Hence it is a continuous piecewise linear function in terms of the nodal basis that comprises *m* processes in a production line, and the output is its production value. During the production process, the value of the output starts from zero and increases until the value is equal to $p_{m-1}(t_m) = z$.

If a process, which is part of the production line, has its cycle time reduced, it can produce the same value in a shorter period. The cycle time of the output will then be reduced by the same length of time. Therefore, attention should be paid to a certain process rather than studying all the production processes for that output.

Lemma 6.1. *The step size* $l_i = t_{i+1} - t_i$, $i = 0,1,...,m-1$, *in* $p_i(t)$ *can be reduced by applying an appropriate new operator in* l_i, *subject to the value of* $p_i(t_{i+1})$ *not being decreased.*

Proof. See Appendix A. □

The new operator is a function in $p_i(t)$ and can be a new method, new equipment, or new setup. The introduction of this new operator results from fulfilling or removing the hindrance in the original process, especially when the process is a non-value-added process as in case (ii) in the proof of Lemma 6.1. A hindrance is defined as any cause that prevents the elimination of a process or activity that is non-value-added. From Lemma 6.1, it can be said that the length of time cannot be shortened if there is no change in this process.

Moreover, every process may be decomposed into a sequence of activities. In fact, the decomposed process has the same characteristics as Lemma 6.1, too. The following lemma explains this point.

Lemma 6.2. *There exists a piecewise polynomial between knots* t_i *and* t_{i+1}, $i = 0,1,...,m-1$, *that has the same characteristics as Lemma 6.1.*

Proof. See Appendix A. □

Many performance improvement opportunities exist throughout the whole process, from receiving an order to the delivery of the manufactured goods or services. Lemma 6.2 states that lead-time reduction can come from both the processes in production and the sub-processes within a process. The improvement in the non-value-added movement of an operator on the work floor mentioned by Ohno (1988, p.57) is due to a reduction in the cycle time in a sub-process. The suggestions Shingo (1989) make about eliminating waste in transportation, delay, and inspection are indeed either process or sub-process improvements that depend on various situations. If the components of the production function are examined, only positive/negative value-added processes and non-value-added processes can be identified, as stated in Proposition 6.1.

Proposition 6.1. *A piecewise polynomial value-added production function P contains only positive/negative value-added processes and non-value-added processes, not waste.*

Proof. Since $P(t)=\sum_{i=1}^{m} p_{i-1}(t_i)\phi_{i-1}(t)$, $a \le t \le b$ and $p_0(t_0)=0$, $p_{m-1}(t_m)=z$. Consider c_i in each line segment $p_i(t)=c_i t+d_i$, $t_i \le t \le t_{i+1}$, $i=0,1,...,m-1$, and note that c_i has only three different types of behaviour:

(i) If $c_i > 0$, then $p_i(t_i) < p_i(t_{i+1})$. Thus, $p_i(t)$ is a positive value-added process.

(ii) If $c_i = 0$, then $p_i(t_i) = p_i(t_{i+1})$. Thus, $p_i(t)$ is a non-value-added process.

(iii) If $c_i < 0$, then $p_i(t_i) > p_i(t_{i+1})$. Thus, $p_i(t)$ is a negative value-added process.

Now let $e(t)$ be the value of the waste and assume that $P'(t)$ contains both $P(t)$ and $e(t)$. Then

$$P'(t)=P(t)+e(t) \quad \text{or}$$

$$e(t)=P'(t)-P(t). \tag{6.3}$$

However, (6.3) is equal to zero because $e(t)$ does not produce any value to the process. Therefore, $P(t)=P'(t)$. Hence, waste is in between two processes. $P(t)$ only comprises the three different processes. □

Therefore, waste is not part of the process and waste elimination is the result of improvement in the process. In fact, waste occurs after a new process or a new sub-process is put in place within an operating process. In order to reduce cycle time and lead-time, the hindrance or unfavourable issues in any area along the process or sub-process must be identified. If the hindrance can be removed, then the lead-time of an output will be correspondingly reduced. The results of cycle time and lead-time reductions can be obtained piece by piece along the horizon in a manner of continuous improvement.

Proposition 6.2. *The production lead-time/cycle time of a production function* P *with a piecewise polynomial can be reduced by continuous improvement in some segments and sub-segments by defining new operators for those segments and sub-segments.*

Proof. This proposition follows Lemma 6.1 and Lemma 6.2. □

Proposition 6.2 states that change is a must in any improvement effort. Such an improvement is indeed to reduce technical inefficiency in production. The change will be even more worthwhile if it also increases productivity.

The results of the analysis in this chapter, therefore, allow different work within a process to be defined as follows:

1 Value-added work
   - 1.1 Positive – work that creates value as perceived by the customer;
   - 1.2 Negative – work that decreases the accumulated value of the process.

2 Non-value-added work – work that is necessarily carried out to cope with the hindrance and maintains the accumulated value of the process. Each piece of non-value-added work is associated with one or more hindrances.

3 Waste – work that has been eliminated following the introduction of a change in the process.

According to these definitions, there is no waste but only value-added and non-value-added work in any running process. Each piece of non-value-added work is associated with one or more hindrances.

6.3 Value-added Input and Efficiency

According to the analysis in Section 6.2, capacity efficiency can be expressed as the ratio of value-added time to the productive time since non-value-added time is inefficient. Therefore, positive value-added time can be treated as full efficient work. We study value-added work in the following to develop the super capacity trading unit.

Let a production function $P:\Re_+^n \to \Re_+$ to produce output $q \in \Re_+$. We denote the variable input which includes value-added and non-value-added work by the vector $\mathrm{x} = (x_1, x_2, \ldots x_n) \in \Re_+^n$. The input requirement set is $V(q) = \{\mathrm{x} \in R_+^n : (\mathrm{x}, q) \in Q\}$, where Q is the production possibility set of a supply chain. The input requirement set is the set of all input bundles that produce at least q units of output.

Lemma 6.3. $V(q)$ *is a convex set.*

Proof. Let x_1, $x_2 \in V(q)$ and $0 < \alpha < 1$ for $\forall \alpha$, since $\mathrm{x} = (x_1, x_2, \ldots x_n) \in \Re_+^n$ is the positive input, so the point $\alpha x_1 + (1-\alpha)x_2 \in V(q)$. □

In the case that x* contains only value-added input to produce $q_0 = 1$ unit of output, we can define a value-added input set as

$$V^*(q_0) = \{\mathrm{x}^* \in \mathrm{R}_+^n : \mathrm{x}^* \in V(q_0) \text{ and } \mathrm{x} > \mathrm{x}^*\}. \tag{6.4}$$

Therefore, the non-value-added input set is the difference between $V(q_0)$ and $V^*(q_0)$.

Lemma 6.4. $V^*(q_0)$ *is an efficient set and there exists a unique* $\mathrm{x}^*(q_0)$ *that represents the minimum input for the output.*

Proof. Suppose the value-added input set has two elements, $\mathrm{x}^* = (x_1^*, x_2^*) \in \Re_+^n$ to produce an output q_0. There exists the minimum point $\mathrm{x}^*(q_0)$ by Lemma 6.3. We follow the expression of productive efficiency in Farrell (1957) to construct an input-orientated measure in Figure 6.1. The unit input isoquant is represented by the curve VV'. A supplier uses quantities of inputs, defined by the point B, to produce a unit of output, the value-added input $\mathrm{x}^*(q_0)$ is represented by the distance OA and the

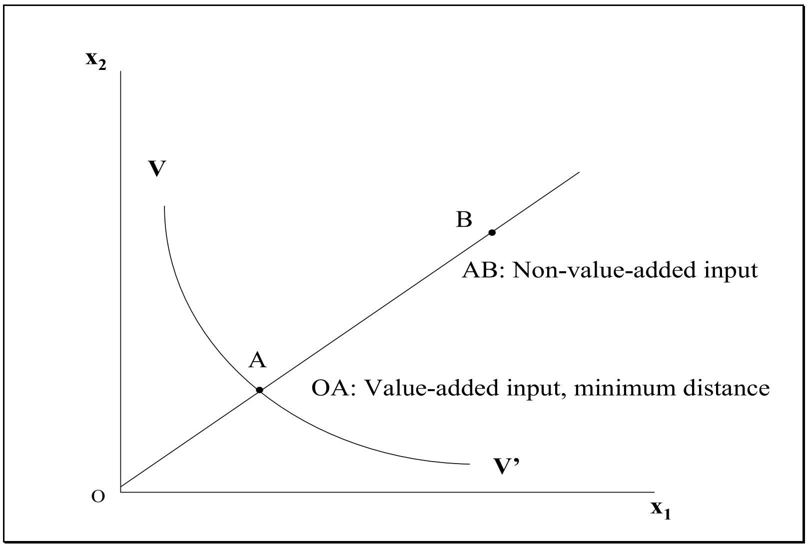

Figure 6.1: Input-orientated measures by distance function

non-value-added input is the length of AB. Since $V^*(q_0)$ is the value-added input and we could not reduce the distance OA by definition, VV' is the fully efficient isoquant and $V^*(q_0)$ is an efficient set. Therefore, the supplier could only improve its production efficiency by reducing the length of AB. □

We therefore could consider a time-based value-added input to construct a value-added capacity for super capacity trading as shown in section 6.5.

6.4 Capacity Deployment

Before we develop the value-added capacity, we need to study the components of capacity. Since we have a production function $P:\Re_+^n \to \Re_+$ with $\mathrm{x} = (x_1, x_2, \ldots x_n) \in \Re_+^n$ to produce output $q \in \Re_+$. Denote the capital input by the vector $K = (K_{n+1}, \ldots, K_{n+m})$. Hence

$$q^t = P(\mathrm{x}^t, K^t) \qquad K^t \le K \qquad t = 1, \ldots T .$$

Such that if a firm employs x^t variable inputs and utilizes K^t capital inputs in period *t*, the firm will have a level of q^t output. K is the size of plant and is associated with a capacity $h(K)$, which cannot be exceeded, regardless of the amount of variable inputs employed (Panzar 1976). Therefore, the rated capacity of a plant is defined as $h(K) = \max_{x^t} P(\mathrm{x}^t, K)$ and $h(K)$ can be deployed into available and unavailable capacities. Unavailable capacity includes non-saleable capacity, scheduled holidays and maintenance, unscheduled material or labour shortage.

Then the available capacity would be deployed into operative and inoperative capacities. Inoperative capacity exists because a production line needs to be set up and developed, or there is short idle time caused by training, absenteeism or breakdown. However, the operative capacity includes effective and damaged capacities. Effective capacity has both value-added and non-value-added capacities. Damaged capacity is the production losses. Figure 6.2 depicts the deployment of the capacity.

We define the available, operative, effective and value-added capacity in the following subsections.

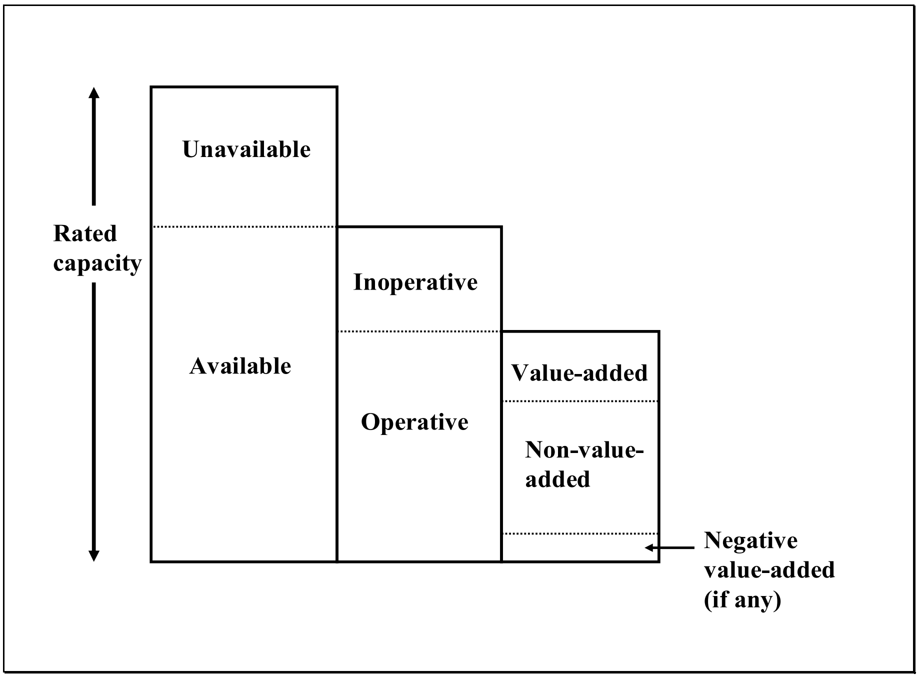

Figure 6.2: Capacity deployment

6.4.1 Available capacity

Available capacity is the expected maximum rate of production of a system except for unscheduled downtime, and is defined as

$h(K^t) = \max_{x^t} P(x^t, K^t)$, and $h(K^t) \le h(K)$, if only K^t capital inputs are available for production.

6.4.2 Operative and effective capacity

Operative Capacity is the maximum rate of production of a process or a system without idle and on-line preparation time. In a given period t that produces p different products, operative capacity is

$$h(\underline{K^t}) = \max_{x^t} P(\underline{x^t}, \underline{K^t}) = \sum_{p=1}^{P} \sum_{i=1}^{r} v_i^{p_t} q^{p_t} \left(\frac{1}{1-\eta_{p_t}} \right) + \sum_{p=1}^{P} \sum_{j=r+1}^{s} n_j^{p_t} q^{p_t} \left(\frac{1}{1-\eta_{p_t}} \right) \quad (6.5)$$

where $\underline{x^t}$ and $\underline{K^t}$ are inputs under operation in period t, q denotes the output quantity, v is the value-added work and n is the non-value-added work; η_{p_t} is the rate of damaged capacity and $h(\underline{K^t}) \le h(K^t) \le h(K)$. We define $\eta_{p_t} = \dfrac{\text{Operative Capacity - Effective Capacity}}{\text{Operative Capacity}}$, where effective capacity is the sum of positive value-added and non-value-added capacity:

$$\text{Effective Capacity} = \sum_{p=1}^{P} \sum_{i=1}^{r} v_i^{p_t} q^{p_t} + \sum_{p=1}^{P} \sum_{j=r+1}^{s} n_j^{p_t} q^{p_t} \quad (6.6)$$

Damaged capacity refers to the negative value-added work in Section 6.3 due to output failure.

6.4.3 Value-added capacity

Value-added capacity is the minimum capacity required to produce output under the constraint of value-added input set. We define the value-added capacity as

$h(K^{t^*}) = \max_{x^t} P(x^{t^*}, K^{t^*})$, where $x^{t^*} \in V^*(q)$ contains only value-added input.

6.5 Capacity Trading Unit

In Section 6.3, we know that the value-added input is the minimum input required by a production function for an output, and it is independent from production lines or facilities under the same technology level. Therefore, in a sub-industry the request of value-added input of a product is independent of any supplier. We can extend this concept to the value-added capacity, in a given unit of time as measurement, in the following propositions.

Proposition 6.3. *The value-added capacity* $h(K^{t^*}) = \max_{x^t} P(x^{t^*}, K^{t^*})$ *is an efficiency capacity.*

Proof. The value-added input $x^{t^*} \in V^*(q)$ is an efficient set by Lemma 6.4. Therefore, $h(K^{t^*})$ follows the lemma to be the minimum input. □

Proposition 6.4. Value-added capacity is an effective trading unit for super capacity.

Proof. From Lemma 6.4, we know that there exists a minimum input in a value-added input set for an output that is common to all supply chains in the sub-industry. Moreover, value-added capacity is efficient by Proposition 6.3. Hence each retailer can calculate the needed amount of value-added capacity for his order. Moreover, the supplier's value-added capacity is clearly defined in Section 6.4. □

Hence the value-added capacity is a full efficiency capacity and will not be affected by different production functions. It is an output orientated input measure unit. Therefore, it can be a trading unit in the super capacity market.

Furthermore, the supplier can estimate saleable value-added capacity from capacity planning and an available value-added capacity ratio. We define the available value-added capacity ratio as the following:

$$\text{Available value-added capacity ratio} = \frac{\text{Value - added capacity}}{\text{Available capacity}} = \frac{h(K^{t^*})}{h(K^t)} \quad (6.7)$$

This ratio is worked out from historic data and is therefore only an approximate figure. However, we do not need a precise quantity in the capacity requirement planning stage because super capacity is reserved for the coming season, and the supplier can eliminate or reduce the discrepancies between the actual capacity required by the retailer and the available capacity by adjusting capacity from overtime production or subcontracting.

6.6 Conclusions

We develop in this chapter a time-based, value-added capacity measurement model for super capacity trading. The model measures process capacity based on our new concept that a process contains only value-added and non-value-added work, but no waste. Since the value-added capacity is fully efficient, there is no difference between different production lines. Therefore, it can be easily inferred to the output quantity of a retailer's order if the requested value-added input of the product is known.

A supplier's available capacity and the correspondingly available value-added capacity ratio of a production line can be converted into the supplier's value-added capacity for selling in the market. Then the value-added capacity unit also allows super capacity trading between different supply chains to share the same market price.

7 Conclusions

This study sought to investigate mitigating mismatching risk of newsvendor-type retailers due to demand uncertainty by pooling and sharing the risk among different supply chains in a sub-industry, and transferring it to external speculators via financial instruments. We used super capacity as the means for risk pooling to reduce demand variability and share inventory and supply risks among players to build a theoretical framework for our study. The following chapter offers discussion and conclusions of our study. We summarize our major findings in Section 7.1. From there we discuss how our postulation fits the operational hedging strategies in Section 7.2. In Section 7.3 we talk about the concern of short lead-time capacity. Section 7.4 provides an example of capacity trading in history. In Section 7.5, we have discussed thoroughly about the managerial insights and potential applications of our theoretical results. We then inspect the limitations of our study and suggest future research possibilities in Sections 7.6 and 7.7. Section 7.8 forms concluding remarks.

7.1 Summary of Major Findings

We propose in this dissertation a co-opetition game in which we treat super capacity as an independent product. Retailers can then reduce and hedge against demand uncertainty by buying super capacity as futures as an alternative inventory for single-period goods. Hence retailers can estimate the amount of the available super capacity they wish to convert into inventory later, when more updated market information is available during the season. Moreover, retailers can improve flexibility by exchanging their super capacity holdings as a commodity with their competitors in a sub-industry. From the supplier's viewpoint, other than income from selling capacity futures, super capacity trading can increase capacity utilization and help create a more stable production environment during the selling season. The supplier will also benefit by shifting the volume of production from the pre-selling season to the selling season. In so doing, peak production period can be prolonged to give a smoother production schedule throughout the year. In this way the supplier and retailers can hedge against capacity and inventory risks, respectively, by trading super capacity at any time before the super capacity expires.

Furthermore, we allow speculators to become involved in the game and the financial instruments that include futures and options on futures for hedgers to protect the player's quantity and price risks in the trade. In the next subsections we present the major findings that answer our research questions in Section 2.5.

7.1.1 Trading super capacity between two supply chains

We began our study by examining two supply chains to compare two scenarios in which capacity residual is allowed and is not allowed to be traded between the retailers in stage 2. We have shown that the pooling mechanism of trading super capacity futures drives Pareto-improvement since both retailers and suppliers improve their profit from this mechanism. We also found that the price of super capacity has a strong effect on retailers in modifying their inventory levels in period 1. Retailers will hold less inventory if the price of super capacity is low compared with their profit margins. Nevertheless, if retailers are allowed to exchange the capacity residual in stage 2, they will reduce their optimal physical inventory levels in stage 1.

On the other hand, we realize that the supplier would consider profit margin and cost of unutilized super capacity, and the predicted percentage of unutilized super capacity in stage 2 to determine the amount of super capacity to be reserved.

7.1.2 Trading capacity futures among n-newsvendor

We then extended our setup to a group of players that contained *m* suppliers and *n* competitive retailers. We employed a biform game to analyze the risks and payoffs to retailers in both stages. We showed that the game is efficient in both the non-cooperative and cooperative stages, as all the players will reach a unique equilibrium point. Hence, the sub-industry will increase their aggregate payoff as a result of trading super capacity futures and will also improve performance in matching inventory with stochastic demand for short life-cycle products. The coalitions that are formed by retailers to transfer the residual capacity futures among themselves in stage 2 allow the retailers to maximize their expected additional profit from the game. Therefore, the game is a Pareto-improvement too.

Moreover, we have identified the entry barrier for retailers: the game discriminates against retailers who have lower profit margins, inventory costs, and lost sales penalties. Nevertheless, the retailers joining the trading of super capacity to hedge inventory risk or balance their sale loss will not be worse off. We also observed that the price of super capacity helps the balancing of demand and supply of capacity to distribute the additional payoff to all the players in the game.

7.1.3 Transfer mismatching risk to the public

We have further released our setup to include any speculators wishing to join the game to act together with retailers and suppliers. We have proven the existence of the unique Nash equilibrium to demonstrate that this new mechanism is an efficient means to let a single supply chain risk be shared with or transferred to the other supply chains, and even to the public who are in the game. The game also engages speculators to share the risk premium of hedgers. The different preferences in payoff and risk between hedgers and speculators stimulate the transactions of super capacity futures. Our work on risk transfer to risk-neutral and risk-averse speculators allows us to understand that the performance of the game would provide a fair hedge for retailers and suppliers. An optimal payoff for risk-averse players has also been identified. The results indicate that this new market-based risk transfer mechanism combines operational and financial hedging strategies which offer industry a new way of meeting demand more efficiently to achieve Pareto-improvement, even with the occurrence of speculation.

The trading of super capacity as futures creates a multiple order replenishment environment for retailers during both the early and post-early seasons. Under this new strategy, the profit improvement for individual traders as well as for the aggregate sub-industry mainly comes from better matching between supply and demand, because leftover inventories and lost sales are decreased by both risk pooling and order postponement.

As a whole, our findings showed that there are some inherent advantages in using super capacity futures as a substitution for inventory. First, the trading of super capacity futures makes it worthwhile for suppliers to develop shorter lead-time capacities, because the value of their scarce ability and extra risk can be reflected and therefore compensated under the market mechanism. Second, the trading of super capacity as a commodity in the futures market can provide an efficient mechanism to reduce demand variability for the supplier and shift both capacity and inventory risks from a certain supply chain to a sub-industry. Third, the futures market increases the liquidity of super capacity between players with different risk and time preferences. Fourth, the trade also provides an opportunity for retailers to adjust their inventory positions during the selling season, not only upwardly but also in a bidirectional way.

7.1.4 Capacity unit for trading

We also developed in this dissertation a time-based, value-added capacity measurement unit that is common to a mix of products in different supply chains. Therefore, this new capacity unit allows super capacity trading among different supply chains involving various products being processed.

7.2 Operational Hedging Strategies

From the major findings in the last section, we recognize that our super capacity trading game is, in fact, highly compatible with the operational hedging strategies that are proposed by Van Mieghem (2008). Van Mieghem borrows the risk mitigation means from the insurance industry and adds one operations management technique to formulate four generic operational hedging strategies for mitigating operations risk. In the insurance industry, the belief is that risks over many clients can be pooled and used to build reserves to meet claims from policy holders. The industry further protects itself by using contracts to transfer remaining risk to reinsurers. For operational risks, Van Mieghem suggests reducing or even eliminating the root causes by postponement with OR, supplier collaboration and improvement, robust product and process design, as well as quality improvement.

Obviously, super capacity pools both different demands and different supplies together into a single source. This pooling effect can reduce the variances of demand and supply to mitigate expected mismatch cost while improving service. The financial instruments that we suggest for trading super capacity can further share and transfer the risk among a sub-industry and outside speculators too. We have used expected utility model to demonstrate that the differences in utilities is the main generator for super capacity trading to be actuated. Nevertheless, super capacity, as an alternative inventory that is decoupled from the physical goods, has its own market price, and can be transacted independently, which would induce suppliers to build short lead-time capacity that can become reserved capacity for the whole sub-industry. Moreover, the capacity management concept we assert in Chapter 6 is a fundamental strategy to reduce and eliminate the root causes of long cycle time/lead-time that allows postponement strategy to be implemented. Therefore, our mechanism adopts all four operational hedging strategies. However, there may be some concerns in implementing the hypothesis of the availability of short lead-time capacity in this study.

7.3 Short Lead-time Capacity

Nevertheless, people may query whether short lead-time capacity of this game can be realized in industry, since this novel mechanism has never been executed in the real world.

We know retailers would like to have a very short or even close to zero lead-time to give maximum flexibility so as to be able to respond to varying demand. Dell and Zara, for example, make their products in five days (Handfield and Nichols 2002) and fifteen days (Ferdows *et al.* 2004), respectively. Wainwright Industries, a manufacturer in the U.S.A., and the 1994 Baldrige Award recipient who supplies parts for the automotive and aerospace industries, reduced its lead-time from 8.75 days to 15 minutes (Verespej 1996). However, as we have argued in Chapter 1, short lead-time might spoil the supplier's benefit because part of the mismatching risk is shifted along a single supply chain from retailer to supplier. Therefore, efforts to achieve shorter lead-time, such as those made by Dell and Zara, are not common practice in industries that provide services to retailers.

Suppliers prefer a longer lead-time as a buffer against business risks such as the late shipment of raw materials, quality problems in the production line, or varying productivity. Meanwhile, a longer lead-time would permit the supplier to accumulate enough orders to justify an "optimum" production schedule, or at least a fuller utilization of the production capacity. But lead-times might be very short if the managers of a supply company are willing and able to make them so. Proposition 6.2 in Chapter 6 explains why a lead-time or cycle time of a production function can be reduced. Furthermore, we demonstrate how lead-time of a supplier could be shortened in Appendix B with a couple of real world examples.

The main concept of Section 6.2 and Appendix B is that any process has no waste but only value-added and non-value-added work. If we want to shorten cycle time/lead-time, we have to solve any hindrances in non-value-added work. The case in Section B.3 of Appendix B that depicts a trading firm in reducing lead-time of finished product delivery demonstrates how a 7 days lead-time can be shortened to a few hours by using this concept. The example in Section B.5 shows how a cycle time was saved by as much as 57% in a supplier in Thailand by applying Proposition 6.2. In both cases, different hindrances were identified in the processes and the managers took actions to eliminate and settle them. Manifestly, the performance of shortened lead-time depends on the ability and the effort of managers to solve the hindrances in their processes. Therefore, suppliers must be given an incentive that would compensate them for their additional endeavour, and a mechanism for managing their risks. Subsections 3.5.1 and 4.2.1 of this thesis have proved that a super capacity futures game is an appropriate induction for suppliers to offer short lead-time to retailers. In our analysis, if the super capacity market has been formed, then a good opportunity might exist for short lead-time capacity to be provided.

7.4 An Example of Capacity Co-opetition Game

Even though the proposed super capacity trading game has never appeared in any part of the world, we can refer to a close example to perceive the performance of the game. The typical example of a capacity market is quota trading in textile and clothing, used by some countries between the 1960s and 2004. During that period, many undeveloped and developing countries in Asia faced quantity restrictions on their export of textiles and apparel articles to the U.S.A., Canada, and the European Union under the Long-term Agreement on International Trade in Cotton Textiles, and later the Multi-Fibre Agreement (Hinkelman 2002). The quantity restriction was on a categorical base. A garment or textile was classified into different categories according to its style and the contents of material. Therefore, a quota list, for example for importing into the U.S.A., would be about 100 articles. Some exporting countries, such as Hong Kong, Korea, and Indonesia, allowed firms that held allocated quotas to transfer their quota usage to other firms. Thus, the quota had a value at any time before the right of holding it expired. Different quota markets were subsequently formed in the exporting countries during that period. The price of quota fluctuated according to the aggregate supply and demand of such category. Garment exporters quoted the selling price to importers with a two-part tariff: the price of the goods and the price of the quota. The quota utilization rates in the previously mentioned countries were consistently more than 90%, while in other countries where quota transfer was not allowed, such as India and Pakistan, quota utilization was relatively low[3] (Krishna and Tan 1998).

The main reason for supporting the high quota utilization is that the importer did not need to search for any exporters who had such quota quantity for their orders. The importers only concerned themselves with the right suppliers for their orders, and the prices were affordable. The quota market provided a function to match the right quota with the right supplier. However, in India and Pakistan, they might not have been able to get a certain order only because they were not allocated for such a quantity of appropriate quota. The inflexible quota allocation system, therefore, erected a barrier to sourcing activity.

This example of quota trading provides a good reference for the trading of super capacity, illustrating that traders are willing to shift spare capacities to competitors to generate additional benefits.

7.5 Managerial Implications

The analytical results in this dissertation may underline the usefulness to managers at different levels for decisions making on coping with demand

[3] It is difficult to trace the actual quota utilization levels for such poor performing countries because the quotas of certain categories would have been suspended before the end of the period due to low utilization rates.

volatility. We have emphasised the super capacity should be decoupled from physical inventory to become an independent product and has its own market price. Managers need to notice that the increase of unit cost of inventory paid for super capacity in fact will increase the overall payoff since mismatching cost will be mitigated. The super capacity price therefore becomes a tool to allocate saving of mismatching cost among different players. Basically managers could have potential applications of our theoretical results of this study in the following two levels.

7.5.1 The strategic level

Our propositions show that retailers should not solely use postponement strategy but also blend it well with risk pooling and sharing strategies to alleviate mismatching risk due to demand uncertainty. At the strategic level, we advocate that retailers might co-operate with their competitors to pool their reserved capacity together. However, Proposition 4.2 tells us that the game has an entry barrier. Therefore, managers should select those supply chains with similar profit margins, inventory costs and lost sales penalties to involve in the game. Otherwise they will be discriminated.

Our research results also suggest that an individual retailer will benefit from the super capacity game if it suffers from poor sales revenue in a good business year of the sub-industry because it can sell the capacity residual in a better price. Therefore, managers might consider holding certain amount of super capacity futures to insure their payoffs, no matter how confident their forecast to hedge demand risk on.

Additionally, managers do not need to avoid or exclude any speculator to involve in the game because we confirm that the demand uncertainty risk can be transferred to outsiders. In our study, the trading system allows different kinds of risk preferences to join in the game. Individual benefit will not be affected by the behaviour of other participants. However, managers should be aware that they will play a dual role, both as hedgers and speculators, in case they intend to make extra profit from trading super capacity itself.

Apart from the above, this study proves that Pareto-improvement would be attained in the super capacity game. Therefore, suppliers are highly suggested to develop shortening production lead-time and sell their super capacity to futures market. Obviously, a supplier would sell all or part of its super capacity in stage 1 according to its risk preference to maximize its payoff.

7.5.2 The tactical level

At the tactical level, our research finds out that the acceptable price of super capacity futures for a supplier will not be higher than the sum of supplier's profit margin and the cost of super capacity. Supplier will ask for a higher price of super capacity futures if it predicts the utilization of reserved capacity will be low in the season and vice verse. Therefore, managers of retailers would ask for a lower price in their first move.

Nevertheless, we also understand that supplier can manipulate stage 1 inventory order quantity by super capacity futures price, no matter how large the super capacity is. Hence, supplier would influence retailer to place an order that the quantity is favoured to it by adjusting the capacity futures price. But retailer could use market force to dilute the influencing effect of supplier since capacity is a commodity and its price is determined by demand and supply of the market. From retailer's point of view, therefore, capacity trading will be better to involve more than one supplier.

In contrast, supplier could develop short lead-time capacity by the policy we introduce in Proposition 6.2 to solve and eliminate any hindrances in non-value-added work along process. Managers of suppliers are advised to select their lead-time and productivity improvement project by the criterion that investment should be less than saving of one period divided by rate of return, otherwise they will make a loss in that project.

7.6 Limitation of the Study

This study has focused on the trading of super capacity in different market structures only. Unfortunately, we have not analyzed an optimum size of coalition for the cooperative game. We have not considered the optimal super capacity investment plans of supplier neither. The assumption of fixed product wholesale and retail prices could destroy the set up due to the correlation between prices of super capacity, and product retail prices have not been considered. Moreover, the assumption that transaction costs are negligible in this dissertation might ignore different actions in the players' responses to the mechanism in reality (Coase 1937).

Despite much effort, it remains that speculator behaviour has not been studied in this dissertation. Indeed, we do not know if much negative impact would be brought by speculators to the super capacity market in actual operation.

7.7 Recommendations for Future Research

Future research could make several extensions of the current study. We suggest that additional work could be done on the minimum size of sub-industry that has an effective market price for super capacity, such that a failure of super capacity exchanging mechanism in the starting period due to insufficient players in the game might be prevented. Moreover, the optimum size of coalition for trading capacity residual in stage 2 could be investigated to maximize the profit of hedgers.

Further more, it might also be worth studying the payoffs of hedgers under different actions of a speculator in trading super capacity. We have yet studied the behaviour of speculators in our thesis; nevertheless, we would avail from understanding whether circumstances in which capacity price is manipulated by speculators will damage the benefit of hedgers in the market.

Besides, future studies can also contribute to the literature by exploring how demand volatility and degree of competition among supply chains affect the results of super capacity futures trading. Even we know that sub-industry will gain from the super capacity game, we may want to know how attractive the game would be to retailer that is facing different amounts of demand fluctuating and/or different levels of product substitution in the market. It will help managers to select appropriate sub-industries that have the right applicable nature characters to adopt the super capacity trade.

Furthermore, investigation of suppliers' investment strategies for super capacity might also broaden the scope of this study. Future research may examine supplier's capacity optimal investment which depends on demand uncertainty, and the conditions that suppliers can improve their incomes by super capacity trading. Extensions of current study to include negative super capacity spot price in stage 2 would be interesting. Under this situation, suppliers can make use of super capacity price as a discount tool to attract buyer's order in a tough business environment. However, we might also want to know how the inventory pricing system would be affected by the negative capacity price.

7.8 Concluding Remarks

Our work contributes to the literature by showing that short lead-time capacity can become a commodity that alleviates mismatching risk of newsvendor-type supply chains by blending four operational hedging strategies. These risk mitigation strategies are to pool, reduce, avoid and hedge against demand uncertainty. The price of super capacity as an efficient capacity allocation and risk redistribution means in the co-opetition game is also a new idea in the literature.

Moreover, our study also provides new academic context by extending a single-period mismatching risk problem from one supply chain or one supply network to a group of heterogeneous supply chains in a sub-industry.

Furthermore, we create a new value-added capacity unit measure that can contribute to the literature and help examine mixed products or mixed supply chain capacity planning and scheduling. The concept behind the value-added capacity unit that waste does not exist in any process could be a breakthrough concept in both academia and industry.

We believe our work can benefit industry in a rational way by improving the performance of operations in a supply chain. It could be very exciting if in the future any practitioners in the industry are able to establish a super capacity market to realize the results of this dissertation.

Appendix A

Mathematical proof in Chapter 3

Proof of Proposition 3.2.:

Let $L(q_2)=[D_2-q_2]^+$ and $L(q_3)=[D_3-q_3]^+$. Comparing Scenario 2 with the one-stage problem, we see that the supplier's income will increase by selling super capacity and obtaining profit from the traded super capacity in stage 2. Let I_i be the increased profit. We have

$$I_i = h[\alpha L(q_2^0)+(1+\beta)L(q_3^0)]+(\omega_2-c_2)\beta' L(q_3^0)$$

or

$$I_i = h[\alpha L(q_2^0)+(1+\beta)L(q_3^0)]+(\omega_3-c_3)\beta' L(q_3^0).$$

Meanwhile, the supplier's profit will decrease by the amount of super capacity that replaces the leftover inventory and the cost of unutilized super capacity. Let I_d be the decreased profit. We have

$$I_d=(\omega_3-c_3)\beta L(q_3^0)+s(\beta-\beta')L(q_3^0).$$

Since the supplier will be motivated if $I_i > I_d$, which implies

$$h \ge \frac{(\omega_3-c_3+s)(\beta-\beta')L(q_3^0)}{\alpha L(q_2^0)+(1+\beta)L(q_3^0)}.$$

Let $\gamma=\dfrac{(\beta-\beta')L(q_3^0)}{\alpha L(q_2^0)+(1+\beta)L(q_3^0)}$. It follows that $0\le\gamma<1$, since $\beta\ge\beta'$, $\alpha>0$, $\beta>0$ and $\beta'>0$. Hence,

$$h\ge\gamma(\omega_3-c_3+s).$$

Mathematical proofs in Chapter 4

Proof of Lemma 4.1.:

Let the expected profit $g_i(y)$ of retailer i be

$$g_i(y) = E\{p_i \min(D_i, y) - \omega_i y\} = p\mu_i - \omega_i y - pE(D_i - y)^+ . \qquad \text{(A4.1)}$$

If retailer i substitutes k units of super capacity with the price h per unit, then the expected profit becomes

$$g_i(y+k) = E\{p_i \min[D_i, (y+k)] - \omega_i y\}$$

$$= p\mu_i - \omega_i(y+k) - pE[(D_i - (y+k)]^+ . \qquad \text{(A4.2)}$$

The retailer is willing to buy super capacity under the condition that $g_i(y+k) > g_i(y)$.

We have $(D_i - y)^+ = \mu - S(y)$, where $S(y) = y - \int_0^y F(x)dx$.

From (A4.1) and (A4.2), $pS(y) + pk - \omega k - hk > pS(y)$

or $$h < p - \omega . \qquad \text{(A4.3)}$$

Another condition for the retailer to hold super capacity is $h < \omega$. So the retailer is willing to hold super capacity if

$$\frac{h}{(p-\omega) \wedge \omega} < 1, \text{ where } (p-\omega) \wedge \omega = \begin{cases} p-\omega, & if\,(p-\omega) < \omega \\ \omega, & if\,(p-\omega) \geq \omega \end{cases} \qquad \text{(A4.4)}$$

In stage 1, if $h = 0$, the suppliers will not trade the super capacity futures, so $0 \notin \alpha_i$.

Therefore, $\dfrac{h}{(p_i - \omega_i) \wedge \omega_i} < 1$ and so $H_i^\alpha \in (0,1)$. □

Proof of Lemma 4.:

By the definition of the characteristic function $v(S)$ that the total additional payoff to the retailers in S is as least $v(S)$ and may have a maximum value as in (4.10) when a joint strategy for S is adopted. Similarly, T also holds the same statement. Therefore, the union of S and T will have a total additional payoff of at least $v(S)+v(T)$. However, the maximum value for the union might be even larger, which is the reason to induce the players to enhance the size of the coalition.

Mathematical proofs in Chapter 5

Proof of Corollary 5.1.:

From (5.17), we have $E\Pi^T(q^C) = r_i\delta_i Var\left[\Pi^T(q^C)\right] = \sum_{i=1}^{m+n+\eta}\frac{1}{r_i}Var\left[\Pi^T(q^C)\right]$

Hence,
$$\frac{\partial[E\Pi^T(q^C)]}{\partial r_i} < 0 \qquad \text{(A5.1)}\square$$

Proof of Corollary 5.2.:

We can work out the derivatives of the choice function directly to obtain the result from (5.22) since we have $\pi_i(q^C)^* = \gamma_i^* + \delta_i^*\Pi^T(q^c)$.

Consider
$$\delta_i^* = \frac{1/r_i}{\sum_{i=1}^{m+n+\eta}\left(1/r_i\right)} = \frac{r_i^{-1}}{\sum_{i=1}^{m+n+\eta} r_i^{-1}},$$

therefore,
$$\frac{\partial\delta_i^*}{\partial r_j} = \begin{cases} \dfrac{-r_i^{-2}\left(\sum_{i=1}^{m+n+\eta} r_i^{-1} - r_i^{-1}\right)}{\left(\sum_1^{m+n+\eta} r_i^{-1}\right)^2} < 0 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases},$$

hence
$$\frac{\partial\pi_i(q^C)^*}{\partial r_j}\begin{cases} < 0 & \text{for } i = j \\ = 0 & \text{for } i \neq j \end{cases}. \qquad \text{(A5.2)}\square$$

Proof of Proposition 5.3.:

If the Nash equilibrium is unique, the payoff function of each player should be twice continuously differentiable, and the marginal payoff of each player is strictly decreasing on q^C and on q_i^C, $\forall i \in I$ (Corchon 2001). We denote $MP = MP_i(q_i^C, q^C)$ as the marginal extra profit of player i, then we define:

$$MP = \frac{\partial \pi_i(q_i^C, q^C)}{\partial q_i^C} + \frac{\partial \pi_i(q_i^C, q^C)}{\partial q^C}. \tag{A5.3}$$

If payoffs are maximized, marginal revenue is equal to marginal cost, that means the optimal marginal profit is equal to zero. Hence we will get

$$MP = \frac{\partial \pi_i(q_i^{C^*}, q^{C^*})}{\partial q_i^C} + \frac{\partial \pi_i(q_i^{C^*}, q^{C^*})}{\partial q^C} = 0.$$

Therefore, $MP = MP_i(q_i^C, q^C) < 0$, the marginal extra profit of player i is strictly decreasing on q_i^C and on q^C, $i \in I$.

Moreover, the payoff function $E(\pi_i(q_i^C))$ is a twice continuously differentiable function; therefore, the Nash equilibrium in Proposition 5.2 is unique. □

Mathematical proofs in Chapter 6

Proof of Lemma 6.1.:

Let $p_i(t) = c_i t + d_i$, where $t_i \le t \le t_{i+1}$, $i = 0,1,\ldots,m-1$. (A6.1)

Case (i): $c_i \ne 0$

The length L_i of $p_i(t)$ between t_i and t_{i+1} is given by

$$L_i = \sqrt{l_i^2 + (c_i l_i)^2} = \sqrt{(1+c_i^2)}l_i \text{, where } l_i = t_{i+1} - t_i . \quad \text{(A6.2)}$$

When c_i increases, as the contribution value $(c_i l_i)$ is kept unchanged, so L_i will decrease and $l_i = \dfrac{L_i}{\sqrt{(1+c_i^2)}}$ decreases. (A6.3)

Thus, l_i will decrease if there is an appropriate new operator g_1 to increase c_i by rotation or transformation and the value of $p(t_{i+1})$ will not decrease, whereby

$$g_1(p_i(t)) = c_{gi} t + d_{gi} \quad \text{(A6.4)}$$

and $c_{gi} > c_i$. If $l_i = 0$, it means $t_0 < t_1 < \ldots < t_{i-1} < t_i = t_{i+1} < t_{i+2} < \ldots < t_m$.

Case (ii): $c_i = 0$

Let $$p_i(t) = c_i t + d_i = d_i \quad . \quad \text{(A6.5)}$$

Then $$L_i = (\sqrt{(1+c_i^2)})l_i = l_i . \quad \text{(A6.6)}$$

Hence, if there is a new operator g_2 to rotate or transform, then it will have a certain amount of contribution as in Case (i). □

Proof of Lemma 2.:

Let $p_i(t) \in P(t)$ be a line segment in the interval $[t_i, t_{i+1}]$. Insert *n*-1 knots between t_i and t_{i+1} such that $t_i = t_{i,0} < t_{i,1} < \ldots < t_{i,n} = t_{i+1}$. Let $P_{i,j} : [t_{i,j}, t_{i,j+1}] \to \Re$ be a straight line, $p_{i,j}(t) = \gamma_{i,j} t + \delta_{i,j}$, and the standard basis is $\{\phi_{i,0}, \phi_{i,1}, ..., \phi_{i,n-1}\}$, where

$$\phi_{i,k}(t_{i,j}) = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases} \tag{A6.7}$$

So $P_i(t) = \sum_{j=1}^{n} p_{i,j-1}(t_{i,j})\phi_{i,j-1}(t)$, where $t_i \leq t < t_{i+1}$. (A6.8)

Thus, $p_i(t)$ is a piecewise polynomial, which satisfies Lemma 1, too. □

Appendix B
A New Look at Waste and Value-added Work in the Toyota Production System

This appendix is a supplementary study of Section 6.2. It investigates the components of the production function and explores the reasons why non-physical waste in processes is hard to be identified. Analyzing more than 50 process improvement cases, we find that a process contains only value-added and non-value-added work, not waste. Value-added work can be either positive or negative. Negative value-added work is work that decreases the accumulated value of the process. Non-value-added work is unavoidable when there is hindrance in the process. Waste will only be revealed after changes are made to the process to eliminate the corresponding hindrance. We argue that the concept that waste exists in any process is a blind spot in the Toyota Production System because waste only manifests when hindrance is removed. To improve production efficiency, firms should take steps to identify and eliminate the hindrance behind their non-value-added work.

B.1 Introduction

The Toyota Production System (TPS) has been around since the 1970s. Application of the TPS' fundamental concept of "absolute elimination of waste" to reduce costs and obtain better efficiency is well known both in academia and industry. Shingo (1989) describes this concept as being so powerful that it could squeeze water even from a dry towel. Since the 1990s, Lean Manufacturing has promoted this concept beyond manufacturing to service industries and its implementation from the production floor to the whole enterprise.

Despite the length of time that TPS has been around, it seems that academics and industrialists are not yet able to fully grasp its philosophy. In the late 1990s, researchers focused on studying the spirit of TPS (Spear and Bowen 1999). In 2004, Liker (2004) presented a model to explain the concepts of TPS and its corporate culture, which provided a deeper analysis of TPS than had previously been offered. A few observers comment that Toyota just could not tell people in words what they were doing, not even in Japanese (Holweg 2006). Nevertheless, some TPS advocates emphasise that although the concept of waste can be easily understood, it is hard to be identified or sought out in reality (Ohno 1988; Shingo 1989; Imai 1997; Liker 2004; Liker and Meier 2006). In fact, Katsuaki Watannabe, Toyota's former president, admitted in an interview in 2007 that two or three months were not long enough for anyone to understand the "Toyota Way" (Stewart and Raman 2007). It is this apparent difficulty that people have in explaining and understanding TPS that motivates us to conduct this study that takes a new look at the fundamental concept of TPS. We set out to investigate the components of the production function and explore the reason why non-physical waste in processes is hard to identify.

In Section B.2 we review the dissemination of the concept and practice of TPS in the Western world and discuss the notion of waste estimation. In Section B.3 we argue that waste does not exist in any process. In Section B.4 we explain that waste results from the relationship between process timing and the value created by a process. We derive the conditions under which productivity will improve from shortening lead time and cycle time. We present a real-life example that illustrates how a Thai manufacturer adopts the ideas in Section B.5 to achieve efficiency improvement in its production process. In Section B.6 we conclude the appendix and suggest topics for future research.

B.2 TPS and Waste Elimination

Toyota has been developing TPS since the 1940s. In the 1950s, Toyota began to teach TPS to its suppliers, but the method was not formally documented until 1965 when Kanban systems were rolled out to suppliers. However, the West did not know about TPS until the late 1970s (Schonberger 2007). In this section we review the history of the introduction of TPS to the West in the 1970s and discuss the concept of waste elimination.

B.2.1 Dissemination of TPS

In 1977, two articles on TPS appeared in English (Ashburn 1977; Sugimori *et al.* 1977), one of which was written by Toyota's managers. In 1978, Taiichi Ohno, the father of TPS, published *Toyota Production System* in Japanese, which was translated into English in the late 1980s (Ohno 1988). Shigeo Shingo published an English version of *A Study of the Toyota Production System from an Industrial Engineering Viewpoint* in 1989 (Shingo 1989). Both Ohno and Shingo tried to explain in their books the background of TPS and developed a theory to explain the Just-in-Time (JIT) concept, which is central to TPS. In the early 1980s, Monden published a series of three articles on TPS in the *Industrial Engineer* (Monden 1981a, b, c) and later published a book entitled *Toyota Production System* (Monden 1983). In this book, he devotes a considerable portion to explain the Kanban system and other production methods and tools that are used in Toyota. During the same period, Schonberger (1982) and Hall (1983) published books based on their own observations of JIT. These books were accompanied by several key articles in academic journals and played a major part in disseminating the JIT message to the Western world (Holweg 2006). However, the views expressed in these books are different from those expressed by Ohno and Shingo. Schonberger (1982), Hall (1983), and Monden (1983) focus on the descriptions of the methods used in production lines and the tools that are applied in Toyota, although they take a similar approach to that of Ohno and Shingo in emphasizing the advantages of TPS over the traditional American production method. For example, they use JIT purchasing, Kanban, zero inventory, pulling system, setup time reduction, and employee involvement to describe the Japanese production system.

Americans who adopted the new production concept found that the implementation of TPS and JIT did not always produce the desired results. Some researchers propose that a lack of the Japanese culture is the key barrier to TPS' effectiveness in non-Japanese organizations and suggest that American companies should not just copy Japanese factories (Bolwijn and Brinkman 1987; Heiko 1989). However, upon examining their arguments closely, it is clear that they treat TPS and JIT merely as tools rather than a philosophy. For example, Ansari and Modarress (1986) identify that the major problems of JIT sourcing are low product quality, and lack of support from suppliers, top management, employees, and carrier companies, problems that would not have existed if the practitioners had truly understood the philosophy of JIT. In fact, the problems identified by Ansari and Modarress are issues that JIT could help to solve. A key concept of JIT is producing only what is needed, when it is needed, and in the amount needed (Ohno 1988, p.4). A crucial strategy of JIT is to continuously identify the root causes of problems, which in practical terms involves asking "why" five times. This is known as the "5 Whys". This approach is designed to effectively discover the root cause of a problem. The 5 Whys can be applied to all the problems identified in Ansari and Modarress (1986). In fact, in TPS, the 5 Whys is a means by which the hidden but real cause of a problem is revealed (Ohno 1988, p.17).

The cultural concern has been underlined in the book *The Machine that Changed the World* (Womack *et al.* 1990). They argue that national culture is not a barrier to implementing JIT. They use a new label – "Lean" – to promote TPS and JIT, a term that became popular in different industries in the 1990s. The value stream technique demonstrated in the book caused manufacturers to pay attention to TPS again. Womack and Jones (1996) summarize the theory of Lean as follows: "*Precisely specify value by specific product, identify the* value stream *for each product, make value flow without interruptions, let the customer pull value from the producer, and pursue perfection."* This summary suggests that Lean focuses on the concept of flow to eliminate waste in operating processes. This approach is easier for both manufacturers and service providers to understand and accept. Lean, however, only emphasizes the value stream and does not take the human factor into consideration.

In 2004, *The Toyota Way* (Liker 2004) introduced to industries another new tide of production concepts. He presents the 4P (philosophy, process, people/partners, and problem-solving) model and provides a more comprehensive explanation of how Toyota, as distinct from Lean, which considers only process improvement, succeeded. He develops 14 management principles from this 4P model to describe TPS and the Toyota culture. He emphasizes that if companies adopt only process improvement, this will do little more than tinkering with problems, because such improvements do not engage the heart and the intellect. Therefore, the performance of such companies would continue to lag behind that of companies that adopted the values of continuous improvement (Liker 2004).

B.2.2 Waste elimination

The Toyota Way, in Toyota's own words, is built on two pillars, namely Continuous Improvement and Respect for People (Liker 2004; Stewart and Raman 2007). Kaizen, the main purpose of which in TPS is to reduce waste, is the vital element in Continuous Improvement and defines Toyota's basic approach. Ohno (1988) explains that TPS is fundamentally based on the tenet of absolute elimination of waste and classifies the movement of workers as being either waste or work. Waste, by his definition, is needless, repetitious movement that must be eliminated immediately, while work includes both non-value-added work and value-added work. Non-value-added work is work that does not add value but must be done under existing work conditions, while value-added work is work that contributes to a product and therefore generates added value. Ohno emphasizes that work efficiency will increase if waste is eliminated.

In addition to the seven wastes that Ohno suggests, namely over-production, waiting, transportation, over-processing, inventory, movement, and defects, Shingo (1989) observes that an operation process should consist of four components, namely processing, inspection, transport, and delay operations, and indicates that only the processing adds value while the other three components are wastes. In *Lean Thinking*, Womack and Jones (1996) use the Japanese word "muda" to replace the English word "waste". They also indicate that the value in a process can be extended beyond the work or movement of a worker. These authors apply Ohno's value and waste concepts to the whole production process and create a new tool, which they call the Value Stream. In the Value Stream, the actions include design and order, and the making of a specific product is sorted into three categories: (a) action that creates value that is perceived by the customer; (b) type one muda – action that creates no value but is currently required by the system and so cannot yet be eliminated; and (c) type two muda – action that does not create value as perceived by the customer and can be eliminated immediately (Womack and Jones 1996).

Liker (2004) declares the elimination of waste to be the heart of the TPS. However, his definition of waste is confused with non-value-added work. He renames Ohno's seven wastes as seven non-value-adding wastes (Liker 2004, p.28). Furthermore, Liker adopts Fujimoto's (1999) approach by classifying the processing time into value-added time and non-value-added time (i.e., waste) in order to integrate them with the Value Stream tool. Unfortunately, he fails to define what types of non-value-added steps are nevertheless necessary.

Some researchers define waste in a broader sense. For example, Russell and Taylor (1999) provide a definition of waste as "anything other than the minimum amount of equipment, materials, parts, space, and time that are essential to add value to the product." Liker and Morgan (2006) define waste as "what costs time and money and resources but does not add value from the customer's perspective". Hicks *et al.* (2004) provide a similar holistic view of waste, stating that anything that adds cost but not value is waste. Nicholas (1998) includes space, energy, material, and time in the waste list. Obviously, these definitions make it difficult to distinguish clearly between waste and non-value added work. However, in this study we focus on the non-physical waste that is emphasized by TPS in order to analyze the fundamental concept.

Nevertheless, all these researchers claim that waste does exist in any process and can be eliminated immediately. They also admit that it is not easy for people to identify waste in their own processes. The practitioner's view, however, echoes an analogous conception that JIT is not appropriate for indirect operations that are "not-easy-to-visualize" activities (Adachi *et al.* 1995). As a result, as many as 31 JIT waste elimination techniques have been developed to help companies to eliminate waste (Hallihan *et al.* 1997). In this study, however, we argue that Ohno, Shingo, Womack *et al.*, and Liker all share a common blind spot. The so-called "wastes" in their books and papers cannot be eliminated without applying corrective actions to solve one or more issues within processes. In their proposition, they postulate that a certain change will happen to make an improvement occur. One example is that Shingo suggests using poka-yoke, a mistake-proofing mechanism, to replace inspection stations, which would render as waste the work done at inspection stations. This means that waste cannot be eliminated unless a change is made to the process. But Hyundai Motor Company adopts the poka-yoke concept stressing the prevention of faulty operations among workers instead of using Shingo's approach that seeks to make improvement through changes too (Lee and Jo 2007). Similarly, the design rules for TPS implementation advocated by Black (2007) to eliminate sources of variation in the system to remove waste also emphasize change is the key, not waste. In fact, a manager cannot avoid his or her workers having to move the work-in-process (WIP) if the layout of the production line has is not changed, since otherwise production cannot continue. The moving of WIP is not a waste, but is necessary work unless a change takes place. If people cannot think of a viable alternative to their ways of working, it is hard for them to recognize "the moving of WIP" as waste. Therefore, waste is not blindly eliminated (Towill 2007).

Liker (2004, p.30) describes an astonishing example where engineers and managers could not identify wastes in the manufacturing of steel nuts until he pointed them out. Shingo (1989, p.80) also came across people who just did not agree that waste existed, and asserted that "It has to be done this way".

Shingo suggests that the focus needs to be on those wastes that people take for granted or do not consider as problems. Moreover, the method used by Ohno to train new members to identify waste by “standing in the circle” for eight hours or more in a production floor environment illustrates how hard it is to detect waste (Liker and Meier 2006, p.60). The question we ask in this study is: “why is waste so hard to be seen?” We propose taking the approach of examining the process from the viewpoints of continuous improvement and cycle time reduction to find out where the waste is.

B.3 A New Look at Waste and Value-Added Work

In the mid-2000s, the author of this dissertation ran dozens of workshops to help different companies in a number of Asian countries to improve the efficiency of their processes. The participants involved in the processes were asked to classify the value-added work, the non-value-added work, and the waste in their processes. Then, for each non-value-added work, they were asked to identify the circumstances (hindrance) that prevented the elimination of that work. As expected, the participants found it hard to identify waste in the first place, but when the corresponding hindrance was removed, waste was more easily seen. In this way more than 50 examples of waste elimination were collected.

The next stage was to analyze the process that had been discussed with the process stakeholders, e.g., the people involved in the process or those in the preceding or succeeding processes. It was discovered that all of the cases had the common character that the waste was a necessity for the corresponding process. If it were simply removed, then all of the processes would be suspended and production could not continue. In fact, all successful waste elimination cases involved process improvement exercises. For example, it was common for the manufacturers to take a couple of days to process orders. This may be due to the need for manufacturers to have frequent communication with their customers first to clarify the details of an order before the computer ordering system can accept it. Order administrators were usually very surprised to find out that their value-added work amounted to less than 1 - 2% of the whole ordering process time. But all of them readily admitted that the rest, 98 - 99% of their time, was non-value-added. They pointed out that most non-value-added work resulted from customers not understanding their needs or making careless mistakes, or from administrators who did not have sufficient product knowledge or lacked the necessary information to process orders. These are the main reasons why people take non-value-added work for granted or do not consider it as a problem. However, once guidelines were provided to customers on how to place orders, or when order administrators had sufficient product knowledge or could get the necessary information in a timely fashion, some of the non-value-added work would become waste. In other words, waste could only be recognized after the hindrance in the non-value-added work is removed. In those workshops a new approach was taken whereby work was only classified as being either value-added or non-value-added and the participants found it much easier to "construe up" some of the waste during the discussion. Nobody further insisted that "it has to be done this way". However, if a mistake is made during the order process, the accumulated value of the product will be destroyed or decreased. This work is treated as a piece of negative value-added work. Of the seven wastes in the TPS, all, except for defects, are non-value-added work or processes. The production of defective products is negative value-added work because it reduces the value of an output, despite the fact that it could be repaired, rejected, or sold as seconds. In the worst-case scenario, a defective product will give rise to claims or requests for indemnification from the customer that will incur a cost higher than the selling price. Hence, producing defective products is not non-value-added work.

Therefore, all the above cases support the analysis in Section 6.2, and the following definitions of the different work of a process in Chapter 6 might accommodate them perfectly:

1 Value-added work –

1.1 Positive – work that creates value as perceived by the customer;

1.2 Negative – work that decreases the accumulated value of the process.

2 Non-value-added work – work that is necessarily carried out to cope with the hindrance and maintains the accumulated value of the process. Each piece of non-value-added work is associated with one or more hindrance.

3 Waste – work that has been eliminated following the introduction of a change in the process.

According to these definitions, there is no waste but only value-added and non-value-added work in any running process. Each piece of non-value-added work is associated with one or more hindrance. A hindrance is defined as any cause that prevents the elimination of a process or activity that is non-value-added. Hindrance can be divided into several categories, such as resource hindrance, market hindrance, and culture and policy hindrance. When hindrance is identified and removed in a revamped process, the related non-value-added work will immediately become waste. The 5 Whys and some other QC tools, such as affinity diagrams and relation diagrams, can help managers to identify the root cause of hindrance. Once the root cause is revealed, corresponding actions to eliminate the hindrance can be implemented. Improvement is therefore obtained by eliminating hindrance, not by eliminating waste. The concept of “waste elimination” can therefore be re-cast as “hindrance elimination”. Obviously, in some cases, the hindrance can be easily found and eradicated. But for some non-value-added work hindrance is very hard to eliminate and has to be accommodated in the process. Waste will be exposed and eliminated only after a new and improved process is established. The following is a real-life case that demonstrates this concept.

This case concerns a trading firm that wishes to reduce the time to deliver a finished product, i.e., the time that the product reaches the end of the production line to when it arrives at an assigned port. The process starts with the manufacturer informing the trading firm when the goods for a certain order are ready. On the same day or the next day, the trading firm provides instructions to the manufacturer about the packing requirements for the goods according to the output quantity. It takes one to three days for the manufacturer to arrange the packing of any order. The trading firm then sends a truck to pick up the goods one day after being informed that the packing has been done. According to records, therefore, it generally takes five to seven days to complete this simple process. After analyzing this work flow, the trading firm identifies that the hindrance lies mainly in its own employees' mindsets because they assume that the manufacturer could not understand their complex packing instructions correctly, which also points to a lack of trust on their part. It therefore provides a brief period of training to the workers of the manufacturer. The result is that the finished goods are packed immediately at the end of the production line with no need for further instructions from the trading firm and the goods are sent to the port by the manufacturer the following day. As a result the need for the communication that previously existed in relation to packing and trucking arrangements between the trading firm and the manufacturer is eliminated. The elimination of this communication is not because the communication is a waste in itself but because the manufacturer fully grasps the requirements of the customer. The investment needed to realize this change only involves one half-day training session - a negligible cost in comparison with the saving in the non-value-added process for managing the finished goods for every order and in terms of the reduced lead time. The trading firm identifies the hindrance easily from the non-value-added work and does not insist that the packing instructions have to be done in that way. If the trading firm had been asked to identify any waste in the process, it might have been very hard for it to imagine that the waste in fact lies in the unnecessary communication and that it could be eliminated if the manufacturer is provided with training on understanding the packing instructions. They might only have thought that it was wasteful of the manufacturer to take three days instead of a couple of hours for the packing. Therefore, the case also manifests that the problem is not a pure technical issue. The solution is a result of interactions between the technical and social systems (Lander and Liker 2007).

B.4 Productivity Improvement

Any elimination of non-value-added work that shortens cycle time or lead time is not equivalent to improving productivity because additional resources might be needed to reduce times. It is known that productivity is the ratio of output to input and total factor productivity is the ratio of the total quantity of outputs to total quantity of different inputs. It is a factor that contributes to the profitability of a firm (Van Loggerenberg and Cucchiaro 1981). In the last section we argued that any improvement involves a change in the production process. However, some of the changes may not guarantee that productivity improvement will happen because the additional input that may be needed to bring about a change may not be counterbalanced by the extra output. For example, the dwindling of WIP inventory levels by the JIT manufacturing method seldom improves productivity (Lieberman *et al.* 1990). However, we do find that cost reduction is a condition for productivity enhancement in process improvement.

Lemma B.1. *Process improvement resulting from a reduction in non-value-added work will contribute to productivity improvement on condition that the new process leads to cost reduction.*

Proof. Suppose a firm has taken a process improvement effort to eliminate some non-value-added work in the original process. Let y_1 be the input vector of the original process producing output q and y_2 be the input vector of the improved process producing the same output q. The corresponding input price vectors are (w_1, w_2). Before the process change, the process cost is given by $c_1 = w_1^T y_1$ and after the change, the process cost becomes $c_2 = w_2^T y_2$. Assume y_1, y_2, w_1, w_2, and q are all greater than zero. Then the cost ratio of the two processes is given by

$$\frac{c_2}{c_1} = \frac{w_2^T y_2}{w_1^T y_1} = \left(W^T\right)\frac{y_2}{y_1}. \tag{B1}$$

Multiplying both the numerator and denominator of (B1) by q yields

$$\frac{c_2}{c_1} = \left(W^T\right)\frac{q/y_1}{q/y_2}. \tag{B2}$$

We assume W is a constant and will not affect the process change, so $q/y_2 > q/y_1$ if only $c_2 < c_1$. □

Lemma B.1 is conditional on cost being reduced in order to gain any productivity improvement. In practice, the cost reduction resulting from a process improvement may be verified by comparing the saving from the changes in the operation and the investment associated with those changes.

Since the saving obtained by reducing lead time/cycle time represents a continuous income, it is useful to consider the present value of the continuous cash flow. Let $F(x)$ denote an instantaneous rate of cash flow expressed in dollars per year, where x is continuous time. The interest rate, r, is selected by the investors/management as the one that gives the minimum attractive rate of return. When $F(x)$ is discounted at a continuous rate r, its contribution to present worth is $F(x)e^{-rx}$, and the sum of the contributions over a period takes the form

$$PV = \int_0^u F(x)e^{-rx}dx\,, 0 \le x \le u\,, \tag{B3}$$

where PV stands for the present value. Net present value is a traditional project valuation methodology and it is easily applied to internal improvement projects. If the present value of the saving is greater than the investment, I, i.e., $PV > I$, it reflects that productivity improvement is secured by shortening cycle time or lead time because the input is less than the output.

Proposition B.1. *The lead time/cycle time and the productivity of the production process can be improved simultaneously, if*

1.there exists at least a change in a running process or sub-process in the production line to eliminate its hindrance; and

2.the new process or sub-process is incorporated into the production line without decreasing the total value of the sequence of processes, subject to the present value of the net saving from the change being positive, i.e.,

$$\int_0^u F(x)e^{-rx}dx \ge I\,,\quad 0 \le x \le u\,,$$

where I = *investment in improvement;*

$F(x)$ = *an instantaneous rate of cash flow due to saving resulting from a new process or sub-process;*

r = *the minimal attractive rate of return.*

Proof. Propositions 6.1 and 6.2 cover the components of value-added and non-value-added processes in the production function, and lead time/cycle time reduction. The present proposition emphasizes that any new process or sub-process should not create new issues such that the total value of the sequence of processes is decreased. It should be noted that the production system should not reject the change, otherwise performance improvement will not be guaranteed. Moreover, we consider saving comes from the input of labour, $L(x)$, and capital, $K(x)$, where capital includes everything except labour, e.g., equipment, overheads and materials.

Let $F(x) = \alpha L(x) + \beta K(x)$, where α is the rate of wages and β is the average input unit price. After applying the new operation and when the situation becomes steady, we may assume that the amount saved from labour and capital is also stable, i.e., $A = \alpha L(x) + \beta K(x)$.

Case (i): The values of α and β will not change.

Thus $$PV = \int_0^u F(x)e^{-rx}dx = \int_0^u [\alpha L(x) + \beta k(x)]e^{-rx}dx = A\int_0^u e^{-rx}dx \tag{B4}$$

and since the improvement is permanent, it is assumed that $u \to \infty$, then

$$PV = \lim_{u\to\infty} A\int_0^u e^{-rx}dx = \frac{A}{r}. \tag{B5}$$

Hence investment in improvement is viable if $\frac{A}{r} \geq I$. This means that the investment should be less than the saving of one period divided by the rate of return.

Case (ii): The values of α and β are exponentially increasing at rates g_L and g_K, respectively. Thus

$$PV = \int_0^u F(x)e^{-rx}dx = \int_0^u \left[Le^{g_L x}\right]e^{-rx}dx + \int_0^u \left[Ke^{g_K x}\right]e^{-rx}dx. \tag{B6}$$

Assuming $r > g_L$ and $r > g_K$, both g_L and g_K are growth rates and determined by the investor/management. These growth rates might take into account the future inflation rate. If $u \to \infty$, then using (B5) gives

$$PV = \lim_{u \to \infty} \int_0^u F(x) e^{-rx} dx = \frac{L}{r - g_L} + \frac{K}{r - g_K}. \tag{B7}$$

Thus a similar result to case (i) is obtained, namely that $\frac{L}{r - g_L} + \frac{K}{r - g_K} \geq I$ is the condition for process improvement.

If $g_L = g_K = g$, then $\frac{A}{r - g} \geq I$. □

The above results demonstrate that if the investment in improvement is the same, then

1. The benefit to a company in higher labour cost countries is larger;
2. A low profit margin company will be more eager to improve its productivity;
3. In countries where the inflation rate is expected to be higher, there will be greater motivation to pursue improvement.

These results also easily explain why Ohno and Shingo emphasize the importance of non-physical waste elimination because the investment made in eliminating wasteful and meaningless jobs is much less than the cost of continuing with those wasteful practices. What they suggest is to pay attention to the non-value-added work in a process and to try and identify the hindrance in them. Any latent waste would be revealed once the hindrance is removed in the course of setting up a new process. In fact, most of the hindrance is related to mindsets and culture issues (Suri 1998), which require no capital outlay when it comes to the implementation of the JIT concept (Cheng 1988). This is the reason why Shingo remarks that waste elimination could squeeze water from a dry towel. Obviously, this does not mean that all the hindrance can be eliminated quickly because some hindrance may have existed in a production line due to external restrictions.

Performance improvements do not only come from non-value-added work. They can also be gained from accelerating the value-added work by technology growth, although the investment would be much larger for this type of improvement. The high-tech approach to cope with labour shortage at Toyota's Tahara plant in the early 1990's is an example of value-added work improvement. Unfortunately, Toyota was forced to abandon this strategy due to poor cost efficiency and worker resistance (Benders and Morita 2004). However, discussion of this aspect is beyond the scope of this study.

B.5 An Example

Tana Netting Co., Ltd[4] is a leading manufacturer and distributor of mosquito netting and other products for public health, travel and personal protection, as well as for home and garden applications. Its clients include malaria control projects, UN agencies, international NGOs, businesses, the military, travellers, leading hotels and resorts, and homeowners around the world. Their production facility is located in Chonburi, Thailand. This factory learned about the key concept of Section 6.2 and this appendix in January 2009 and has applied it to its production line with a view to improving productivity.

The production flow is quite simple in the netting business in comparison with other sewing industries. First of all, the fabric is sent to the cutting room after inspection of the incoming material, and then the sewing room makes the product. If the product does not require any further treatment, then it will be sent directly to the packing room. If it needs further treatment, e.g., addition of a specific chemical, the product will go through a further process before being sent for packing (Figure B.1).

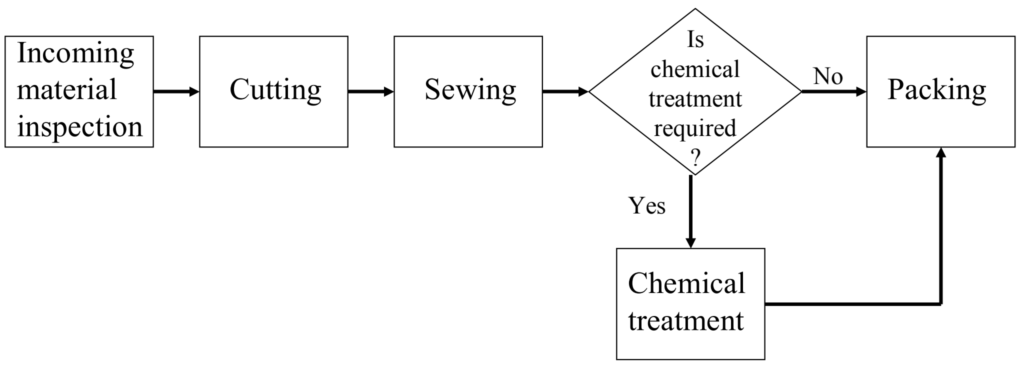

Figure B.1: Production Process in Tana

[4] The website of Tana Netting Co., Ltd. is http://www.tananetting.com

The management team analyzed the cost structure of their production line and identified that the labour cost of the Packing Department was as high as 40.48% of the total direct labour cost. The second highest labour cost was in the Sewing Department. These two departments were therefore selected as the first ones to be dealt with in the process improvement exercise at the beginning of 2009. After a couple of months of operation of the new concept, both departments achieved substantial improvement. For commercial reason, Tana is not willing to share the entire experience, and does not want the results of more than one department be quoted after the introduction of an improved process, so only the results of the Packing Department are shown to illustrate the nature and extent of its achievement. All the figures in this example, therefore, are the results concerning just one production line, not the total output of Tana. It is also assumed for the purpose of this exercise that the production line deals only with one kind of product at all the times. However, the whole Packing Department indeed benefited from the process improvement exercise that resulted in considerable resources saving.

The labour productivity growth process in the Packing Department of Tana can be divided into two periods. The first period was from January to August 2009. It can be seen from Table B.1 that Tana obtained a significant result by reducing the cycle time for each net from 7.56 seconds down to 3.5 seconds in the packing process during this period. The second period was from September 2009 to August 2010 during which the cycle time was reduced even further, down to 3.25 seconds. The daily capacity of the production line therefore increased from 3,333 pieces to 7,200 pieces of net in the first period, and then rose to 7,759 pieces in the second period. Hence the labour productivity increased by 132.79% overall. Even though these figures reflect the results over the whole 20-month period, it should be noted that Tana

Table B.1: Performance of the labour productivity growth in Packing Department

Items	**Before Improvement**	**August 2009**	**August 2010**
Capacity per day	3,333 nets	7,200 nets	7,759 nets
Cycle time per net	7.56 seconds	3.5 seconds	3.25 seconds
Labour productivity change	–	+116.02%	+132.79%

achieved a substantial proportion of the benefit of the process improvement directly after the new packing method was first introduced. The reason for adopting a longer period for this study is to show that the results were maintained over time and the process of improvement continued. The main improvement resulted from a change in the sitting position of the packing workers. Before the change, the workers adopted the traditional practice of sitting on the floor to process the packing (Exhibition B.1). Tana classified the packing procedure into value-added and non-value-added work. In reviewing some of the non-value-added work, they identified that the root

Exhibition B.1. Workers sat on the floor during packing before process improvement

Exhibition B.2. Workers are packing nets on the table after improvement implemented

Table B.2: Saving of the direct labour cost in Packing Department

Items	Before Improvement	August 2009	August 2010
Packing labour cost per net	2.077 baht	0.958 baht	0.889 baht
Total labour cost per net	5.131 baht	4.012 baht	3.943 baht
Portion of packing labour cost	40.48%	23.88%	22.55%
Saving of packing labour cost per net	–	1.119 baht	1.188 baht
Percentage of packing labour cost saving	–	53.88%	57.20%

cause of the hindrance was the operating position of the workers. As a result of their study, they found that packing on a table would shorten the cycle time by more than half (Exhibition B.2), since the new process eliminated some of the movements that were needed in the former position. After a few months' practice, the cycle time became quite stable. In the second period of improvement, they focused on raising the workers' skills by adding a "trainer" to the line and using the continuous-flow process concept. However, they were facing a lot of labour turnover during this period and this required some effort to train up the new comers and for this reason it took a bit longer for the cycle time to become stable.

From the cost point of view, the percentage of saving in the direct labour cost was a little bit better than the saving in cycle time and the increase in daily capacity due to the pay for each worker was not the same. Table B.2 shows that the direct labour cost for packing was 2.077 Thai baht per net before the improvement. In the first period, the cost fell to 0.958 baht, a saving of 53.88%. In the second period, the direct labour cost for packing fell further, down to 0.889 baht per net, i.e., the saving rose to as high as 57.20%. The packing labour cost, therefore, dropped from 40.48% to 22.55% of the total direct labour cost, assuming that all the other departments had maintained their original labour costs[5]. Obviously, the saving in the labour cost illustrates how the new packing process had removed a high proportion of the non-value-added work of the original process. This approach to

[5] The assumption is required for discussion in this thesis since Tana does not disclose the labour cost changes in other departments.

efficiency improvement does not require the participants to search for any waste in the process. Nevertheless, the factory obtained outstanding results in waste elimination simply by changing the physical position of the workers.

The daily output of the Packing Department of Tana is now 7,759 pieces of net and each net represents a saving of 1.188 baht, hence the labour cost saving in the Packing Department is as high as 9,218 baht per day.

Assuming the production capacity is kept constant and that the minimal attractive rate of return is 0.05, according to Eq. (B5) in the proof of Proposition B.1, the present value of the packing labour saving is 55,308,000 baht if the factory runs 50 6-day weeks annually. However, if we assume that there will have to be a certain percentage increase in wages, e.g., a growth rate of 1%, then the present value of the packing labour saving will rise to 69,135,000 baht.

If the amounts saved from labour and capital are combined, it is possible to decide on the validity of the change in the process by comparing the total saving with the investment needed to bring about the improvement. In this case the decision is easy even if the capital saving data are not available because the investment was only for a few working tables and some chairs costing a mere 67,700 baht. This is equivalent to the saving of only a few days of labour cost.

This example shows that people need not follow the recommendation of Ohno to shorten cycle time by searching out and eliminating waste. Tana did not spend a single minute on searching for waste on the production floor but a hidden waste could nevertheless be identified in the Packing Department that could save more than half of the wage costs in the packing process. The concept in Section 6.2 and this appendix is distinct from the traditional TPS since it does not ask practitioners to find waste from the surface phenomena in their processes but to look for hidden issues that could lead to improvement. Obviously, this method is not bound by the argument for the need for waste identification, and makes it easier, therefore, for people to obtain positive results.

B.6 Conclusions

The proposition in this study that waste does not exist in any process explains why waste is so hard to observe. It might also be the reason why so many people do not agree with Shingo (1989, p.80) that waste exists because they could not imagine that any improvement could occur in their systems. A good example that demonstrates this point is a story in Imai's book (1997, p.81) about a press line foreman in a German automotive factory who reacted with disbelief and anger to his consultant setting a 50% reduction target in the setup time without any accompanying technological changes. After the improvement was made, the foreman admitted he had not seen all the "muda" before. In fact, the consultant did not see "muda" either; he only imagined how the process could be changed to allow improvement to happen. The concept that waste exists in any process is a blind spot in JIT and the Toyota Way. This study recommends that what is more important is to find the hindrance behind each piece of non-value-added work and try to eliminate all the hindrance to improve production efficiency.

Changes in any process would probably result in reduced cycle time/lead time. Consequently, the cost of the process might be reduced and productivity improvement would be obtained. Hence, this study explains why "waste elimination" is considered to be the heart of TPS as it enhances Toyota's productivity.

This study may be of benefit to TPS and the Toyota Way because it provides an appropriate explanation for the concept of waste elimination. People who want to learn the main concepts of TPS and the Toyota Way will find them much easier to understand if they use the concept in this study rather than concentrate purely on waste elimination. The propositions we put forward in this study overcome the difficulty of finding waste as lamented by Ohno and Shingo, while still keeping the spirit of the "absolute elimination of waste". Moreover, practitioners can focus on identifying and removing the hindrance in their non-value-added work, which has great potential for process efficiency improvement. Future research may explore the nature and role of hindrance in inhabiting production efficiency improvement.

References

Adachi, T., Enkawa, T., Shih, L.C., 1995. A concurrent engineering methodology using analogies to Just-in-time concepts. International Journal of Production Research, 33(3), 587-609.

Agrawal, V., Seshadri, S., 2000. Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. Manufacturing & Service Operations Management, 2(4), 410-423.

Angelus, A., Porteus, E.L., 2002. Simultaneous capacity and production management of short-life-cycle, produce-to-stock goods under stochastic demand. Management Science, 48(3), 399-413.

Ansari, A., Modarress, B., 1986. Just-in-time purchasing: Problems and solutions. Journal of Purchasing and Materials Management, 22(2), 11-15.

Anupindi, R., Bassok, Y., 1999. Supply contracts with quantity commitments and stochastic demand. In: Tayur, S., Ganeshan, R., Magazine, M. (Eds.), Quantitative Models for Supply Chain Management. Kluwer Academic Publishers, Boston, 197-232.

Anupindi, R., Bassok, Y., Zemel, E., 2001. A general framework for the study of decentralized distribution systems. Manufacturing & Service Operations Management, 3(4), 349-368.

Arrow, K., Harris, T., Marschak, J., 1951. Optimal inventory policy. Econometrica, 19, 250-272.

Ashburn, A., 1977. Toyota's "famous Ohno system", (a reprinted version of the article in the 'American Machinist, 7:120-123'). In: Monden, Y. (Ed.), Applying Just In Time: the American/Japanese Experience (1986).

Industrial Engineering and Management Press, IIE, 9-12.

Aviv, Y., Federgruen, A., 1999. The benefits of design for postponement. In Tayur, S., Ganeshan, R., Magazine, M. (Eds.), Quantitative Models for Supply Chain Management. Kluwer Academic Publishers, Boston, 553-584.

Barnes-Schuster, D., Bassok, Y., Anupindi, R., 2002. Coordination and flexibility in supply contracts with options. Manufacturing & Service Operations Management, 4(3), 171-207.

Benders, J., Morita, M., 2004. Changes in Toyota Motors' operations management. International Journal of Production Research, 42(3), 433-444.

Berman, O., Krass, D., Taibakhsh, M.M., 2011. On the benefits of risk pooling in inventory management. Production and Operations Management, 20(1), 57-71.

Black, F., 1976. The pricing of commodity contracts. Journal of Financial Economics, 3, 167-179.

Black, J.T., 2007. Design rules for implementing the Toyota Production System. International Journal of Production Research, 45(16), 3639-3664.

Blackstone, Jr. J.H., 1989. Capacity Management. South-Western, Cincinnati, Ohio.

Bolwijn, P.T., Brinkman, S., 1987. Japanese manufacturing: Strategy and practice. Long Range Planning 20, 25-34.

Bradford, J.W., Sugrue, P.K., 1990. A Bayesian approach to the two-period style-goods inventory problem with single replenishment and heterogeneous Poisson demands. Journal of Operational Research Society, 41(3), 211-218.

Bradley, J.R., Arntzen, B.C., 1999. The simultaneous planning of production, capacity, and inventory in seasonal demand environments. Operations Research, 47(6), 795-806.

Brandenburger, A.M., Nalebuff, B.J., 1996. Co-opetition. Doubleday, New York.

Brandenburger, A., Stuart, H., 2007. Biform game. Management Science, 53(4), 537-549.

Burnetas, A., Ritchken, P., 2005. Option pricing with downward-sloping demand curves: The case of supply chain options. Management Science, 51(4), 566-580.

Cachon, G.P., Harker, P.T., 2002. Competition and outsourcing with scale economies. Management Science, 48(10), 1314-1333.

Cachon, G.P., Lariviere, M.A., 2001. Contracting to assure supply: How to share demand forecasts in a supply chain. Management Science, 47(5), 629-646.

Cachon, G.P., Netessine, S., 2003. Game theory in supply chain analysis. In: Simchi-Levi, D., Wu, S.D., Shen, Z.M. (Eds.), Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era. Kluwer Academic Publishers, New York, 13-65.

Cachon, G.P., Terwiesch, C., 2009. Matching Supply with Demand – An Introduction to Operations Management. McGraw-Hill Irwin, Boston.

Caldentey, R., Haugh, M., 2006. Optimal control and hedging of operations in the presence of financial markets. Mathematics of Operations Research, 31(2), 285-304.

Cattani, K.D., Dahan, E., Schmidt, G.M., 2008. Tailored capacity: speculative and reactive fabrication of fashion goods. International Journal of Production Economics, 114(2), 416-430.

Chamberlin, E., 1948. The Theory of Monopolistic Competition: A Re-orientation of the Theory of Value, 6th ed. Harvard University Press, Cambridge.

Chavas, J.P., 2004. Risk Analysis in Theory and Practice. Elsevier Academic Press, London.

Cheng, T.C.E., 1988. The Just-in-time production: A survey of its development and perception in the Hong Kong electronics industry. OMEGA-International Journal of Management Science, 16, 25-32.

Chod, J., Rudi, N., Van Mieghem, J.A., 2010. Operational flexibility and financial hedging: complements or substitutes? Management Science, 56(6), 1030-1045.

Christopher, M., 2005. Logistics and Supply Chain Management – Creating Value-adding Networks. Prentice Hall, Harlow, England.

Chung, C., Flynn, J., 2001. A newsboy problem with reactive production. Computers & Operations Research, 28(8), 751-765.

Church, A.H., 1917. Manufacturing Costs and Accounts. McGraw Hill, New York.

Coace, R.H., 1937. The nature of firm. Economica, 4(16), 386-405.

Corchon, L.C., 2001. Theories of Imperfectly Competitive Markets, 2nd ed. Springer, Berlin.

De Leeuw, F.D., 1962. The concept of capacity. Journal of the American Statistical Association, 57, 826-840.

Ding, Q., Dong, L., Kouvelis, P., 2007. On the integration of production and financial hedging decisions in global markets. Operations Research, 55(3), 470-489.

Dong, L., Rudi, N., 2004. Who benefits from transhipment? Exogenous vs endogenous wholesale prices. Management Science, 50(5), 645-657.

Donohue, K.L., 2000. Efficient supply contracts for fashion goods with forecast updating and two production modes. Management Science, 46(11), 1397-1411.

Eppen, G.D., 1979. Effects of centralization on expected costs in a multi-location newsboy problem. Management Science, 25(5), 498-501.

Farrell, M.J., 1957. The measurement of productive efficiency. Journal of the Royal Statistical Society, Series A, 120(3), 253-290.

Ferdows, K., Lewis, M.A., Machuca, J.A.D., 2004. Rapid-fire fulfillment. Harvard Business Review, November, 104-110.

Fisher, M., Rajaram, K., 2000. Accurate retail testing of fashion merchandise: Methodology and application. Marketing Science, 19(3), 266-278.

Fisher, M., Raman, A., 1996. Reducing the cost of demand uncertainty through accurate response to early sales. Operations Research, 44(1), 87-99.

Frazier, R.M., 1986. Quick response in soft lines. The Discount Merchandiser, January, 42-56.

Fujimoto, T., 1999. The Evolution of a Manufacturing System at Toyota. Oxford University Press, New York.

Gabay, D., Moulin, H., 1980. On the uniqueness and stability of Nash-equilibria in noncooperative games. In Bensoussan, A., Kleindoreer, P., Tapiero, C.S. (Eds.), Applied Stochastic Control in Econometrics and Management Science. North-Holland, The Netherlands.

Gan, X., Sethi, S.P., Yan, H., 2005. Coordination of a supply chain with risk-averse agents. Production and Operations Management, 14(1), 80-89.

Gaur, V., Seshadri, S., 2005. Hedging inventory risk through market instruments. Manufacturing & Service Operations Management, 7(2), 103-120.

Gerchak, Y., He, Q.M., 2003. On the relation between the benefits of risk pooling and the variability of demand. IIE Transactions, 35(11), 1027-1031.

Gerchak, Y., Magazine, M.J., Gamble, A.B., 1988. Component commonality with service level requirements. Management Science, 34(6), 753-760.

Gerchak, Y., Mossman, D., 1992. On the effect of demand randomness on inventories and costs. Operations Research, 40(4), 804-807.

Gilreath, T.L., Reeve, J.M., Whalen, Jr. C.E., 1995. Time is Money: Understanding the product velocity advantage. Bobbin, March, 50-55.

Gockenbach, M.S., 2010. Finite-dimensional Linear Algebra. CRC Press, New York.

Gollier, C., 2001. The Economics of Risk and Time. The MIT Press, London.

Granot, D., Sosic, G., 2003. A three-stage model for a decentralized distribution system of retailers. Operations Research, 51(5), 92-105.

Hall, R.V., 1983. Zero Inventories. McGraw-Hill, New York.

Hallihan, A., Sackett, P., Williams, G.M., 1997. JIT manufacturing: The evolution to an implementation model founded in current practice. International Journal of Production Research, 35(4), 901-920.

Handfield, R.B., Nichols, E.L., 2002. Supply Chain Redesign – Transforming Supply Chain into Integrated Value Systems. Financial Times Prentice Hall, Upper Saddle River, N.J.

Heiko, L., 1989. Some relationships between Japanese culture and Just-in-time. Academy of Management Executive, 3(4), 319-321.

Hertenstein, H.J., Polutnik, L., McNair, C.J., 2006. Capacity cost measures and decisions: Two field studies. The Journal of Corporate Accounting & Finance, 17 (3), 63-77.

Hicks, C., Heidrich, O., McGovern, T., Donnelly, T., 2004. A functional model of supply chains and waste. International Journal of Production Economics, 89(2), 165-174.

Hines, T., 2004. Supply Chain Strategies: Customer Driven and Customer Focused. Butterworth-Heinemann, Oxford.

Hinkelman, E.G., 2002. Dictionary of International Trade, 5th ed. World Trade Press, Novato, California.

Holweg, M., 2006. The genealogy of lean production. Journal of Operations Management, 25(2), 420-437.

Horn, R.A., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press, Cambridge.

Hu, X., Duenyas, I., Kapuscinski, R., 2007. Existence of coordinating transshipment prices in a two-location inventory model. Management Science, 53(8), 1289-1302.

Huang, X., Sosic, G., 2010. Transshipment of inventories: Dual allocations vs. transshipment prices. Manufacturing & Service Operations Management, 12(2), 299-318.

Hunter, A., 1990. Quick Response in Apparel Manufacturing. The Textile

Institute, Manchester.

Hunter, N.A., King, R.E., Nuttle, H.L.W., 1996. Evaluation of traditional and quick-response retailing procedures by using a stochastic simulation model. The Journal of the Textile Institutes, Part 2, 87, 42-55.

Hunter, N.A., Valentino, P., 1995. Quick response - ten years later. International Journal of Clothing Science and Technology, 7(4), 30-40.

Imai, M., 1997. Gemba Kaizen: A Common Sense Low-cost Approach to Management. McGraw-Hill, New York.

Iyer, A.V., Bergen, M.E., 1997. Quick response in manufacturer-retailer channels. Management Science, 43(4), 559-570.

Jain, K., Silver, E.A., 1995. The single period procurement problem where dedicated supplier capacity can be reserved. Naval Research Logistics, 42(6), 915-934.

Jin, M., Wu, S.D., 2007. Capacity reservation contracts for high-tech industry. European Journal of Operational Research, 176(3), 1659-1677.

Klammer, T.P., 1996. Capacity Measurement & Improvement: A Manager's Guide to Evaluating and Optimizing Capacity Productivity. Irwin, Chicago.

Klein, L.R., 1960. Some theoretical issues in the measurement of capacity. Econometrica, 28(2), 272-286.

Klein, L.R., Preston, R.S., 1967. Some new results in the measurement of capacity utilization. The American Economic Review, 57(1), 34-58.

Konopka, J.M., 1995. Capacity utilization bottleneck efficiency system – CUBES. IEEE Transactions on Components, Packing, and Manufacturing Technology – Part A, 18(3), 484-491.

Kotler, P., Brown, L., Adam, S., Burton, S., Armstrong, G., 2007. Marketing, 7th ed. Pearson Education Australia/Prentice Hall.

Krajewski, L.J., Ritzman, L.P., 2007. Operations Management – Strategy and Analysis. Pearson Prentice Hall, New York.

Krishna, K.M., Tan, L.H., 1998. Rags and Riches – Implementing Apparel

Quotas under the Multi-Fibre Arrangement. The University of Michigan Press, Ann-Arbor.

Krishnan, K.S., Rao, V.R.K., 1965. Inventory control in n warehouses. The Journal of Industrial Engineering, 16, May-June, 212-215.

Kumar, A., Schwarz, L.B., Ward, J.E., 1995. Risk-pooling along a fixed delivery route using a dynamic inventory-allocation policy. Management Science, 41(2), 344-361.

Kurata, H., Liu, J.J., Yao, D.Q., 2007. Analysis of supply chain's strategic resolutions to order variability – risk pooling and bundling – under vector autoregressive demand. International Journal of Operations and Quantitative Management, 13(3), 173-198.

Lander, E., Liker, J.K., 2007. The Toyota Production System and art: making highly customized and creative products the Toyota way. International Journal of Production Research, 45(16), 3681-3698.

Lau, H.S., 1980. The newsboy problem under alternative optimization objectives. The Journal of the Operational Research Society, 31(6), 525-535.

Lau, H.S., Lau, A.H.L., 1999a. A Note on Bradford-Sugrue's `A Bayesian approach to the two-period style-goods inventory-problem with single replenishment and heterogeneous Poisson demands'. The Journal of the Operational Research Society, 50(2), 187-188.

Lau, H.S., Lau, A.H.L., 1999b. Manufacturer's pricing strategy and return policy for a single-period commodity. European Journal of Operational Research, 116(2), 291-304.

Lee, B.H., Jo, H.J., 2007. The mutation of the Toyota Production System: Adapting the TPS at Hyundai Motor Company. International Journal of Production Research, 45(16), 3665-3679.

Li, J., Liu, L., 2008. Supply chain coordination with manufacturer's limited reserve capacity: An extended newsboy problem. International Journal of Production Economics, 112(2), 860-868.

Lieberman, M.B., Lau, L.J., Williams, M.D., 1990. Firm-level productivity and management influence: A comparison of U.S. and Japanese automobile producers. Management Science, 36(10), 1193-1215.

Liker, J.K., 2004. The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer. McGraw-Hill, New York.

Liker, J., Meier, D., 2006. The Toyota Way Fieldbook - A Practical Guide for Implementing Toyota's 4Ps. McGraw-Hill, New York.

Liker, J.K., Morgan, J.M., 2006. The Toyota Way in Services: The Case of Lean Product Development. Academy of Management Perspectives, May, 5-20.

Lowson, B., King, R., Hunter, N.A., 1999. Quick Response: Managing the Supply Chain to Meet Consumer Demand. John Wiley & Sons, New York.

Luenberger, D.G., 1992. Benefit functions and duality. Jouranl of Mathematical Economics, 21(5), 461-481.

Magretta, J., 1998. Fast, global, and entrepreneurial: Supply chain management, Hong Kong style, an interview with Victor Fung. Harvard Business Review, 76(5), 102-114.

Mattila, H., King, R., Ojala, N., 2002. Retail performance measures for seasonal fashion. Journal of Fashion Marketing and Management, 6(4), 340-351.

McNair, C.J., Vangermeersh, R., 1998. Total Capacity Management: Optimizing at the Operational, Tactical and Strategic Levels. St. Lucie Press, Boca Raton, Florida.

Milner, J.M., Rosenblatt, M.J., 2002. Flexible supply contracts for short life-cycle goods: The buyer's perspective. Naval Research Logistics, 49(1), 25-45.

Monden, Y., 1981a. Adaptable kanban system helps Toyota maintain just-in-time production. Industrial Engineering, 13(5), 29-46.

Monden, Y., 1981b. Smoothed production lets Toyota adapt to demand changes and reduce inventory. Industrial Engineering, 13(8), 42-51.

Monden, Y., 1981c. How Toyota shortened supply lot production time, waiting and conveyance time. Industrial Engineering, 13(9), 22-31.

Monden, Y., 1983. Toyota Production System: Practical Approach to Production Management. Industrial Engineering and Management Press.

Mostard, J., De Koster, R., Teunter, R., 2005. The distribution-free newsboy problem with resalable returns. International Journal of Production Economics, 97(3), 329-342.

Mostard, J., Teunter, R., 2006. The newsboy problem with resalable returns: A single period model and case study. European Journal of Operational Research, 169(1), 81-96.

Netessine, S., Rudi, N., 2003. Centralized and competitive inventory models with demand substitution. Operation Research, 51(2), 329-335.

Nicholas, J.M., 1998. Competitive Manufacturing Management. Irwin McGraw-Hill, New York.

Ohno, T., 1988. Toyota Production System: Beyond Large Scale Production. Productivity Press, Cambridge.

Ozen, U., Fransoo, J., Norde, H., Slikker, M., 2008. Cooperation between multiple newsvendors with warehouses. Manufacturing & Service Operations Management, 10(2), 311-324.

Panzar, J.C., 1976. A neoclassical approach to peak load pricing. The Bell Journal of Economics, 7(2), 521-530.

Phillips, A., 1963. Industrial capacity. The American Economic Review, 53(2), 275-290.

Pinnow, A.D., King, R.E., 1997. Break-even costs for traditional versus quick response apparel suppliers. NCSU-IE Technical Report #97-4, Industrial Engineering Department, North Carolina State University.

Plambeck, E.L., Taylor, T.A., 2005. Sell the plant? The impact of contract manufacturing on innovation, capacity, and profitability. Management Science, 51(1), 133-150.

Plambeck, E.L., Taylor, T.A., 2007. Implications of renegotiation for optimal contract flexibility and investment. Management Science, 53(12), 1872-1886.

Porteus, E.L., 2002. Foundations of Stochastic Inventory Theory. Stanford Business Books, California.

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

Russell, R.S., Taylor, B.W., 1999. Operations Management, 2nd ed. Prentice-Hall, New Jersey.

Sakong, Y., Hayes, D.J., Hallam, A., 1993. Hedging production risk with options. American Journal of Agricultural Economics, 75(2), 408-415.

Schonberger, R.J., 1982. Japanese Manufacturing Techniques. The Free Press, New York.

Schonberger, R.J., 2007. Japanese production management: An evolution - with mixed success. Journal of Operations Management, 25, 403-419.

Schwarz, L.B., 1989. A model for assessing the value of warehouse risk-pooling: Risk-pooling over outside-supplier lead-times. Management Science 35(7), 828-842.

Sethi, S.P., Yan H., Zhang, H., 2005. Inventory and Supply Chain Management with Forecast Updates. Springer, New York.

Shapley, L.S., 1967. On balanced sets and cores. Naval Research Logistics Quarterly, 14, 453-460.

Shingo, S., 1989. A Study of the Toyota Production System from an Industrial Engineering Viewpoint. Productivity Press, Portland, Oregon.

Silver, E., Jain, K., 1994. Some ideas regarding reserving supplier capacity and selecting replenishment quantities in a project context. International Journal of Production Economics, 35(1-3), 177-182.

Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E., 2008. Designing and Managing the Supply Chain: Concepts, Strategies, and Case Studies, 3rd ed. McGraw-Hill/Irwin, Boston.

Slikker, M., Fransoo, J., Wouters, M., 2005. Cooperation between multiple news-vendors with transshipments. European Journal of Operational Research, 167(2), 370-380.

Smith, C.E., Gilbert, S.M., Burnetas, A.N., 2002. Partial quick response policies in a supply chain. In: Geunes, J., Pardalos, P. M., Romeijn, H. E., (Eds.), Supply Chain Management: Models, Applications, and Research Directions. Kluwer Academic Publishers, Boston, 97-115.

Sosic, G., 2006. Transshipment of inventories among retailers: Myopic vs. farsighted stability. Management Science, 52(10), 1493-1508.

Spear, S., Bowen, H.K., 1999. Decoding the DNA of the Toyota production system. Harvard Business Review, September - October, 96-106.

Stewart, T.A., Raman, A.P., 2007. Lessons from Toyota's long drive. Harvard Business Review, July-August, 74-83.

Stuart, A.W., 2005. Biform analysis of inventory competition. Manufacturing & Service Operations Management, 7(4), 347-359.

Sugimori, Y., Kusunoki, K., Cho, F., Uchikawa, S., 1977. Toyota production system and kanban system: Materialization of just-in-time and respect-for-human system. International Journal of Production Research, 15(6), 553-564.

Suri, R., 1998. Quick Response Manufacturing: A Companywide Approach to Reducing Lead Times. Productivity Press, Portland, Oregon.

Thomas, D.J., Tyworth, J.E., 2007. Is pooling lead-time risk by splitting orders simultaneously worthwhile? Journal of Business Logistics. 28(1), 169-193.

Towill, D.R., 2007. Exploiting the DNA of the Toyota production system. International Journal of Production Research, 45(16), 3619-3637.

Van Loggerenberg, B.J., Cucchiaro, S.J., 1981. Productivity measurement and the bottom line. National Productivity Review, Winter, 87-99.

Van Mieghem, J.A., 2003. Capacity management, investment, and hedging: Review and recent developments. Manufacturing & Service Operations Management, 5(4), 269-302.

Van Mieghem, J.A., 2008. Operations Strategy: Principles and Practice. Dynamic Ideas, Belmont, Massachusetts.

Verespej, M.A., 1996. Wainwright Industries. Industry Week, 245, 72-74.

Vlachos, D., Dekker, R., 2003. Return handling options and order quantities for single period products. European Journal of Operational Research, 151(1), 38-52.

Walley, K., 2007. Coopetition. International Studies of Management and Organization, 37(2), 11-31.

Wang, X., Liu, L., 2007. Coordination in a retailer-led supply chain through option contract. International Journal of Production Economics, 110(1-2), 115-127.

Watts, T., McNair, C.J., Baard, V., Polutnik, L., 2009. Structural limits of capacity and implications for visibility. Journal of Accounting & Organization Change, 5(2), 294-312.

Womack, J.P., Jones, D.T., 1996. Lean Thinking: Banish Waste and Create Wealth in Your Corporation. Simon & Schuster, New York.

Womack, J.P., Jones, D.T. and Roos, D., 1990. The Machine that Changed the World. Rawson Associates, New York.

Wu, J., 2005. Quantity flexibility contracts under Bayesian updating. Computers & Operations Research, 32(5), 1267-1288.

Wu, J., Li, J., Wang, S., Cheng, T. C. E., 2009. Mean-variance analysis of the newsvendor model with stockout cost. Omega, 37(3), 724-730.

Yang, H., Schrage, L., 2009. Conditions that cause risk pooling to increase inventory. European Journal of Operational Research, 192(3), 837-851.

Yu-Lee, R.T., 2003. Why capacity trumps costs. Journal of Corporate Accounting and Finance, 14(3), 17-22.

Zabel, E., 1955. On the meaning and use of a capacity concept. Naval Research Logistics Quarterly, 2(4), 237-249.

Zabel, E., 1956. Measures of industry capacity. Naval Research Logistics Quarterly, 3(4), 229-244.

Zhang, J., 2005. Transshipment and its impact on supply chain members' performance. Management Science, 51(10), 1534-1539.

Zhao, X., Atkins, D., 2009. Transshipment between competing retailers. IIE Transactions, 41(8), 665-676.

Zou, L., Dresner, M., Windle, R., 2010. A two-location inventory model with transhipments in a competitive environment. International Journal of Production Economics, 125(2), 235-250.