



**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**OPTIMAL DECISIONS FOR CAPACITATED SUPPLY
CHAINS WITH DEMAND FORECAST UPDATING**

ZHENG MEIMEI

SCHOOL OF MECHANICAL AND AEROSPACE ENGINEERING

2015/2016

**Optimal Decisions for Capacitated Supply Chains with
Demand Forecast Updating**

ZHENG MEIMEI

School of Mechanical and Aerospace Engineering

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

DOCTOR OF PHILOSOPHY

2015

Summary

This thesis investigates how capacitated supply chains make decisions in face of high demand uncertainty. Delaying production or ordering decisions would provide demand responsiveness but incur additional costs and operational complexity. The tradeoffs between early and delayed decisions are investigated from the retailer's, manufacturer's and supply chain's perspectives in three scenarios. In each scenario, optimal decisions are derived by mathematical analysis and be demonstrated to improve the performance. Meaningful managerial insights are provided by analytical and numerical results.

Scenario 1 investigates an extension of the newsvendor with multiple ordering opportunities. To acquire more accurate demand forecast, a retailer prefers delaying order placement, which requires a shorter supply lead time. Due to limited capacity, the manufacturer would not only charge a higher cost for a shorter lead time but also set restrictions on the ordering times (or quantity) in supply mode A (or B). For supply mode A, it is proven under justifiable assumptions that the retailer should order either as early or as late as possible. For supply mode B, an algorithm is proposed to simplify the ordering policy by appropriately relaxing the ordering quantity restrictions. The value of demand forecast updating is illustrated by numerical analysis.

Scenario 2 presents a two-stage newsvendor model, where an expensive emergency ordering opportunity with improved demand forecast and limited quantity is provided, besides the regular ordering opportunity. The optimal regular and emergency ordering quantities are derived by dynamic programming. The effects of the emergency order on the

ordering decisions and expected profit are shown by numerical results.

Scenario 3 studies how a supply chain utilizes a fast reactive production with improved demand information, whose quantity is limited by the preparation in advance. The condition under which adding the reactive production is valuable is derived. The condition is related to the demand forecast updating process, in addition to cost parameters. Furthermore, the benefit of this two-mode production system is illustrated by comparing it with two single production systems. In addition, an efficient pricing contract with a return policy is proposed and optimized to coordinate the supply chain. The coordination contract allows maximizing and arbitrarily allocating the supply chain profit, and remains the same when demand information is unknown to the manufacturer. The benefits of the two-mode production and values of coordination are demonstrated by numerical examples.

Acknowledgement

I would like to express my gratitude to my supervisor, Professor Wu Kan. I feel very lucky to have him as my supervisor. Without his support and guidance, I cannot finish my Ph.D. study. He taught me how to do research and gave me invaluable suggestions.

Meanwhile, I would like to thank Professor Srinagesh Gavirneni and Yuan Xue Ming for their helps and encouragements in my research.

I also would like to thank my parents, my uncles, and my brother for their endless support and love.

Finally, I would like to thank to my friends in Singapore and Ithaca. They are always kind and supportive to me.

Contents

CHAPTER 1. INTRODUCTION	1
1.1 Backgrounds.....	1
1.1.1 Demand Information.....	3
1.1.2 Higher Cost.....	4
1.1.3 Timing Restriction	5
1.1.4 Quantity Restriction.....	6
1.2 Research Objective and Motivation	8
1.3 Organization of the Report	15
CHAPTER 2. LITERATURE REVIEW.....	17
2.1 Demand Forecast Updating Models	17
2.1.1 Time Series Models	17
2.1.2 Bayesian Analysis Models.....	18
2.1.3 Markovian Process Models.....	19
2.2 Inventory Management with Demand Forecast Updating	20
2.2.1 Demand Forecast Updating into the Production Plan.....	20
2.2.2 Demand Forecast Updating into the Purchase Plan	24
2.2.2.1 Two Ordering Opportunities	24

2.2.2.2	Multiple Ordering Opportunities	27
2.2.3	Demand Forecast Updating into the Supply Chain Plan.....	31
2.3	Research Differences	34

CHAPTER 3. NEWSVENDOR PROBLEMS WITH DEMAND FORECAST UPDATING AND SUPPLY CONSTRAINTS 37

3.1	Introduction.....	37
3.2	Model	39
3.2.1	Demand Forecast Evolution	41
3.2.2	Supply Lead Time.....	43
3.2.2.1	Purchasing Cost Function	44
3.2.2.2	Supply Restrictions on Ordering Time and Quantity.....	46
3.2.3	Supply Mode A.....	48
3.2.4	Supply Mode B.....	54
3.3	Numerical Results.....	59
3.3.1	Sensitivity Analysis of the Value of Demand Forecast Updating	60
3.3.2	Comparative Results.....	66
3.3.2.1	Comparison among Different Supply Scenarios.....	66
3.3.2.2	Comparison between Proposed Algorithm and Other Ordering Polices	70

3.4	Conclusion	72
3.5	Appendix.....	74

**CHAPTER 4. ON OPTIMAL EMERGENCY ORDERS WITH
UPDATED DEMAND FORECAST AND LIMITED SUPPLY 80**

4.1	Introduction.....	80
4.2	Model	82
4.2.1	Ordering Quantity at Stage 2	83
4.2.2	Ordering Quantity at Stage 1	86
4.2.3	Study of Two Important Demand Distributions.....	92
4.3	Numerical Analysis	95
4.4	Conclusion	111
4.5	Appendix.....	112

**CHAPTER 5. ANALYSIS OF REACTIVE PRODUCTION WITH
PREPROCESSING RESTRICTION IN SUPPLY CHAINS WITH
FORECAST UPDATES117**

5.1	Introduction.....	117
5.2	Model	119

5.2.1	Centralized Supply Chain	121
5.2.2	Comparison with Single-Mode Production	128
5.2.3	Numerical Examples.....	132
5.3	Supply Chain Coordination	139
5.3.1	Decentralized Decisions	140
5.3.2	Buyer	141
5.3.3	Manufacturer	142
5.3.4	Design of Contract Prices	144
5.3.5	Profit Allocation	147
5.3.6	Numerical Illustration.....	148
5.4	Conclusion	151
5.5	Appendix.....	152
 CHAPTER 6. CONCLUSIONS AND FUTURE RESEARCH.....		168
6.1	Summary of the Research	168
6.2	Limitations and Future Work.....	171
6.2.1	Manufacturer	171
6.2.2	Two Retailers and Types of Products.....	172
6.2.3	Supply Uncertainty	172

REFERENCES173

List of Tables

<i>Table 3.1 Value of Demand Forecast Updating in Supply Modes A and B</i>	<i>62</i>
<i>Table 3.2 Summary of Parameter Sensitivities.....</i>	<i>64</i>
<i>Table 3.3 Different Supply Scenarios</i>	<i>67</i>
<i>Table 3.4 Expected Profits among Different Supply Scenarios and Ordering Policies.....</i>	<i>68</i>
<i>Table 3.5 Profit Comparisons among Different Supply Scenarios and Ordering Policies..</i>	<i>68</i>
<i>Table 4.1 The Impact of r on q_1^* and I_f for Different M under ND and LND ...</i>	<i>97</i>
<i>Table 4.2 The Impact of c_1 on q_1^* and I_f for Different M under ND and LND ...</i>	<i>100</i>
<i>Table 4.3 The Impact of c_2 on q_1^* and I_f for Different M under ND and LND ...</i>	<i>101</i>
<i>Table 4.4 The Impact of s on q_1^* and I_f for Different M under ND and LND</i>	<i>103</i>
<i>Table 4.5 The Impact of x_1 on q_1^* and I_f for Different M under ND and LND ...</i>	<i>104</i>
<i>Table 4.6 The Impact of σ_1 on q_1^* and I_f for Different M under ND and LND...</i>	<i>107</i>
<i>Table 4.7 The impact of σ_2 on q_1^* and I_f for different M under ND and LND...</i>	<i>109</i>
<i>Table 5.1 Summary-Production Strategies.....</i>	<i>139</i>

List of Figures

<i>Figure 1.1 Forecast Accuracy Increases after Observing a Portion of Demand (M. Fisher & Raman, 1996)</i>	<i>4</i>
<i>Figure 1.2 Total Cost of the Martingale Demand Model vs. Lead Time for Four Ordering Flexibility Cases (Milner & Kouvelis, 2005)</i>	<i>6</i>
<i>Figure 1.3 Plot of % Benefit Versus Capacity Between Model 0 and Model 1(Gavirneni, et al., 1999).....</i>	<i>7</i>
<i>Figure 1.4 Plot of % Benefit Versus Capacity Between Model 1 and Model 2 (Gavirneni, et al., 1999).....</i>	<i>8</i>
<i>Figure 1.5 GSK Clinical Trials Supply Chain.....</i>	<i>13</i>
<i>Figure 2.1 Observed and Unobserved Part of the Demand (Özer and Wei 2004)</i>	<i>22</i>
<i>Figure 2.2 Timeline of the Model with Two Procurement Opportunities (J. Li, et al., 2008)</i>	<i>26</i>
<i>Figure 2.3 Basic Model Structure of the Single Ordering Problem (T.-M. Choi, et al., 2004)</i>	<i>28</i>
<i>Figure 2.4 Time line of Buyer-Supplier Decisions (Barnes-Schuster, et al., 2002)</i>	<i>33</i>
<i>Figure 3.1 Timeline of Model</i>	<i>40</i>
<i>Figure 3.2 Expected Profit versus Demand Uncertainty under Four Policies in Supply Mode B</i>	<i>71</i>
<i>Figure 4.1 The Two-Stage Model Structure.....</i>	<i>82</i>

Figure 4.2 q_1^* and I_f versus r under ND.....	99
Figure 4.3 q_1^* and I_f versus r under LND.....	99
Figure 4.4 q_1^* and I_f versus c_1 under ND.....	100
Figure 4.5 q_1^* and I_f versus c_1 under LND.....	101
Figure 4.6 q_1^* and I_f versus c_2 under ND.....	102
Figure 4.7 q_1^* and I_f versus c_2 under LND.....	102
Figure 4.8 q_1^* and I_f versus s under ND.....	103
Figure 4.9 q_1^* and I_f versus s under LND.....	104
Figure 4.10 q_1^* and I_f versus x_1 under ND.....	106
Figure 4.11 q_1^* and I_f versus x_1 under LND.....	107
Figure 4.12 q_1^* and I_f versus σ_1 under ND.....	108
Figure 4.13 q_1^* and I_f versus σ_1 under LND.....	108
Figure 4.14 q_1^* and I_f versus σ_2 under ND.....	110
Figure 4.15 q_1^* and I_f versus σ_2 under LND.....	111
Figure 5.1 The Sequence of Events.....	120
Figure 5.2 The break-Even Line in (σ_e, σ_u) Plane When $r = 10$, $c_1 = 3$, $c_2 = 3$, $c_m = 1$, $s_m = 0.05$, $s = 0.1$, $\pi = 0$, and $x_1 = 80$	126
Figure 5.3 Profits and Quantities versus Production Delay Premium Ratio $(c_2 + c_m - c_1)/c_1$ for Three Models.....	131
Figure 5.4 Effects of Upfront Investment Ratio c_m/c_2 , Sales Price r , Shortage Penalty Cost Ratio π/r and Relative Material Salvage Value V on Benefits P_s and P_r ..	133

*Figure 5.5 Effects of Resolved and Unresolved Demand Uncertainties σ_e and σ_u on
Quantities $p_1^{c^*}$ and M^{c^*} and Benefits P_s and P_r 137*

Figure 5.6 Manufacturer and Buyer's Decisions under the Contract.....140

Figure 5.7 Behaviors of Buyer, Manufacturer and Channel with Reservation Fee c_R ...149

Chapter 1. Introduction

Chapter 1 firstly presents high demand uncertainty in current fast-changing market (Boyacı & Özer, 2010; M. L. Fisher, Hammond, Obermeyer, & Raman, 1994). Due to high demand uncertainty, accurate demand information becomes increasingly important for companies' decision-making. Thus, we discuss the usage and value of demand information updating. On the other hand, restrictions (including higher cost, and time and quantity constraints), which may hamper the utilization of demand information updating, are introduced. Then, research problems and motivations are described. Finally, the organization of this thesis is shown.

1.1 Backgrounds

In a contemporary supply chain, an increasing number of companies participate in markets populated by innovative products (M. L. Fisher, 1997) with short life cycles and high demand uncertainty, such as seasonal products, fashions, new electronic devices and trial drugs. To satisfy quickly changing customer needs, companies must speed up their creation of new products. For example, Apple launches new electronic devices within short timeframes to lure customers and retain its competitive edge. Fashion companies must respond to customer preferences quickly and update clothing styles frequently. The acceleration of new product development leads to short product life cycles. For new products, historical data is limited and customer tastes is unpredicted, and thus demand forecasting presents a high level of uncertainty.

Due to high demand uncertainty, there is a great challenge in matching supply and

demand. This challenge leads to increased costs of inventory and stock-outs. For instance, Cotherix Inc. spent millions of dollars on drug overages (Cotherix Inc., 2006). Similarly, the \$1.7 billion US apple industry loses several hundred million dollars annually due to deterioration (Webb A, 2006). Some niche high-tech manufactures suffered sharp demand shortfalls and declared bankruptcy in the early 2000's (Jin & Wu, 2001). To reduce overstock and stock-outs, one efficient strategy is to take advantage of improved demand information. It has been shown that accurate demand information leads to significant reduction of the overage and underage costs in various settings, such as the apparel industry (Iyer & Bergen, 1997), toy industry (Barnes-Schuster, Bassok, & Anupindi, 2002), electronics industry (A. Brown & Lee, 1997; Tsay, 1999; Yan, Liu, & Hsu, 2003) and clinical trials (Fleischhacker & Zhao, 2013).

It is well known that demand information becomes more and more accurate as time evolves (Simchi-Levi, Simchi-Levi, & Kaminsky, 1999). To take advantage of more accurate demand information, companies would like to postpone their all or part of production (or orders). For example, the success of Dell in 1990s shows that it is profitable for companies to make production after customers' orders are placed. For another example, companies break one production (or order) into two production runs (or orders) to take advantage of demand information occurring between the first and second ones (T.-M. J. Choi, Li, & Yan, 2006; M. Fisher & Raman, 1996; Iyer & Bergen, 1997; Doğan A Serel, 2009).

Although later production (or ordering) decisions allow companies to take advantage of improved demand information, it is not always beneficial to companies. Companies need to

pay for a higher cost for decision postponement. Further, the extent of the postponement is limited, due to limited supply chain capacity. For example, the postponed ordering time cannot be later than the due day specified by a manufacturer (Das & Abdel-Malek, 2003). For another example, a later order may be not completely fulfilled due to a short response time (H.-W. Chen, Gupta, & Gurnani, 2013; Jianli Li & Liu, 2008). Overall, all of these factors (i.e., a high cost, and restrictions on the time and quantity) may exist in practice and have negative impacts on the benefit of demand forecast updating.

1.1.1 Demand Information

Demand information accuracy can be improved by observing early demand information. Figure 1.1 shows a dramatic improvement of forecasting accuracy after observing 20% of demand (M. Fisher & Raman, 1996). Companies can have better access to early demand information with the help of modern technologies, such as electronic-data-exchange and point-of-sale data technologies. Moreover, demand information can be updated by observing more market information when it is closer to the selling season. Hence, companies would like to delay their production or orders, which must be completed in a shorter lead time. Given advanced manufacturing technologies, logistics services and globalization, supply lead times can be shortened in various settings. For example, besides ocean shipment, the fast air shipment is used by Hewlett-Packard, who assembles MOD0 in Singapore and delivers them to distribution centers in a variety of areas (e.g., Roseville (California), Grenoble, Guadalajara and Singapore). Due to technology advancing (which enables collecting demand

information), more and more academics and practitioners recognize the importance of demand information accuracy. The description and utilization of demand information updating have been investigated in a variety of works (Refer to books of Suresh Prakash Sethi, Yan, and Zhang (2006) and Thomopoulos (2015)).

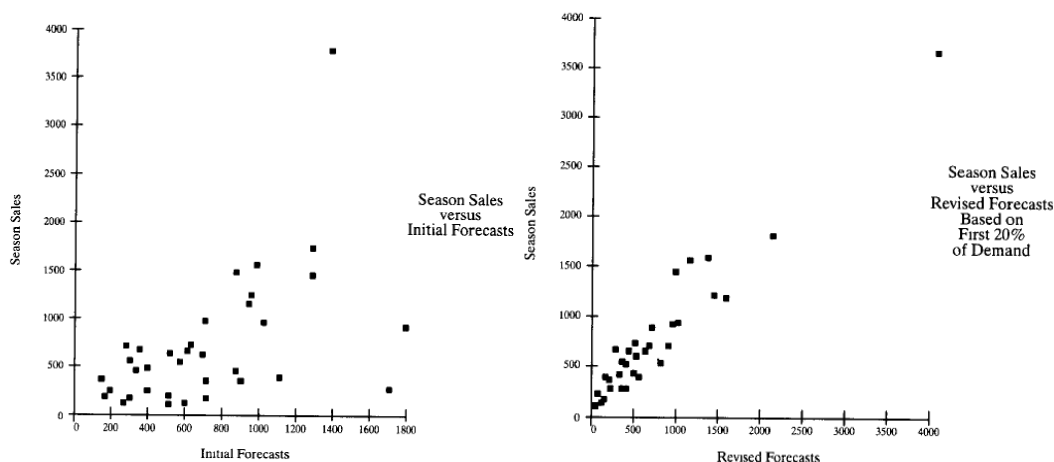


Figure 1.1 Forecast Accuracy Increases after Observing a Portion of Demand (M. Fisher & Raman, 1996)

1.1.2 Higher Cost

In practice, it could be found that a later production (or ordering) cost is higher for the following reasons. First, manufacturers would suffer from later production (or orders), leading to higher production (or ordering) cost. Manufacturers must pay higher costs for high technologies or high run modes to complete later production in a short time. In addition, if retailers' orders are placed too late, it is difficult for manufacturers to do capacity planning under high demand uncertainties. In some cases, manufacturers would offer discounts for

early orders (Tang, Rajaram, Alptekinoglu, & Ou, 2004). Second, additional costs are incurred if products are required to be delivered in a short time. For example, Amazon charges a higher cost for faster deliveries. Third, companies have to source from responsive suppliers at a higher price for later orders. The global sourcing options with a high-cost onshore source and a low-cost offshore source are mentioned in many studies (Allon & Van Mieghem, 2010; Boute & Van Mieghem, 2011). For example, it was reported in apparel industry that the cost is higher from a domestic contractor with a shorter lead time (3-5 days), compared with a Far East contractor with a longer lead time (90 days).

1.1.3 Timing Restriction

For innovative productions with short-life cycles, supply lead times could be very long. Generally, supply lead times usually consist of administration times, transportation times and manufacturing cycle times. Although lead times could be reduced, it cannot be zero due to limited technologies and manufacturing capacity. There exists the shortest lead time (or latest ordering time) in practical contexts. For instance, Das and Abdel-Malek (2003) pointed out that the quickest delivery is restricted by a minimum lead time, and Chan and Chan (2006) showed a range of delivery due dates in a coordination mechanism. Furthermore, the shortest lead time plays a critical role in taking advantage of demand information. Figure 1.2 shows that the cost reduction between the static and timing flexibility, quantity flexibility or full dynamic cases, in which updated demand information could be used to determine the second ordering time, quantity or both, decreases as the supply lead time becomes longer (Milner &

Kouvelis, 2005). Thus, if the shortest supply lead time is long, demand information brings smaller benefits to companies. In some cases, it would be worthless and even harmful for companies to wait for improved demand information if the supply lead time for later production or orders is too long. For instance, compared with China, the additional production cost paid may outweigh the benefit of demand information obtained in Hong Kong, if the supply lead time is not short enough in Hong Kong (Treville, Schürhoff, Trigeorgis, & Avanzi, 2014).

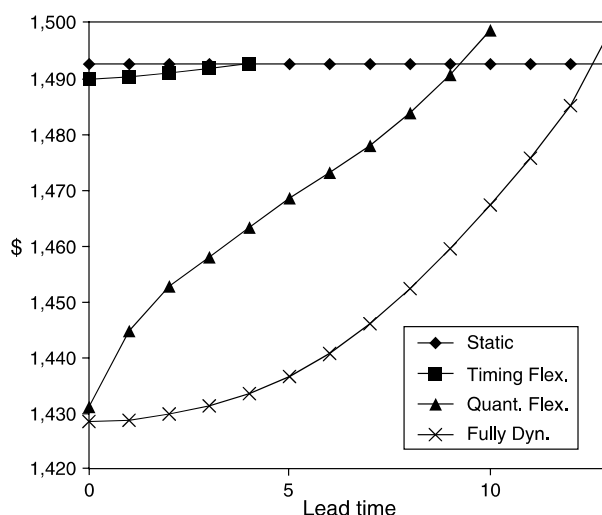


Figure 1.2 Total Cost of the Martingale Demand Model vs. Lead Time for Four Ordering Flexibility Cases (Milner & Kouvelis, 2005)

1.1.4 Quantity Restriction

At the beginning of the planning horizon, companies can produce the required quantity of products with effective capacity and production planning. Due to the fixed capacity and short lead time, the quantity of products that can be produced decreases as time goes. For example,

M. Fisher and Raman (1996) considered a capacity constraint at the second period in a two-stage model with demand forecast updating. Li and Liu (2008) and H.-W. Chen, et al. (2013) showed that a supplier fills a retailer’s regular order in full but place a restriction on the fast-ship order size. Similar to the latest ordering time, the available quantity has a significant impact on the value of demand information updating. Figure 1.3 and Figure 1.4 show that the benefits of some information and increased information increase with the available quantity, respectively (Gavirneni, Kapuscinski, & Tayur, 1999). If the available quantity is very small, more demand information may not bring any value to companies. For instance, in Figure 1.4, the benefit of increased demand information is almost zero when the available quantity is smaller than 25. To increase the benefit, some companies would make quantity or capacity reservation for later replenishments (Erkoc & Wu, 2005).

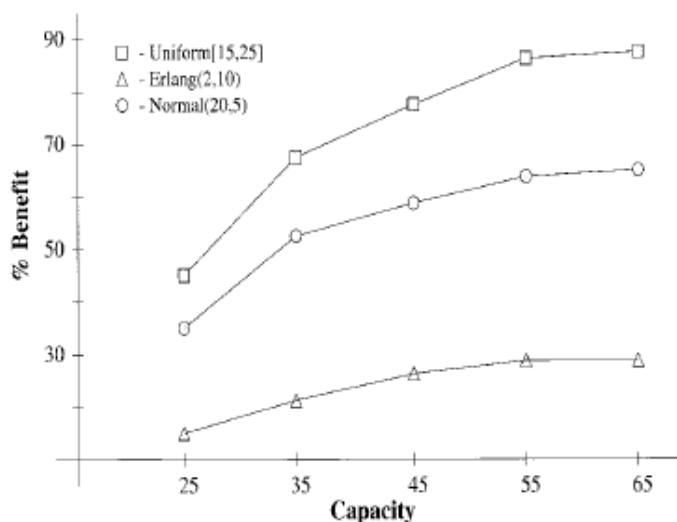


Figure 1.3 Plot of % Benefit Versus Capacity Between Model 0 and Model 1(Gavirneni, et al., 1999)

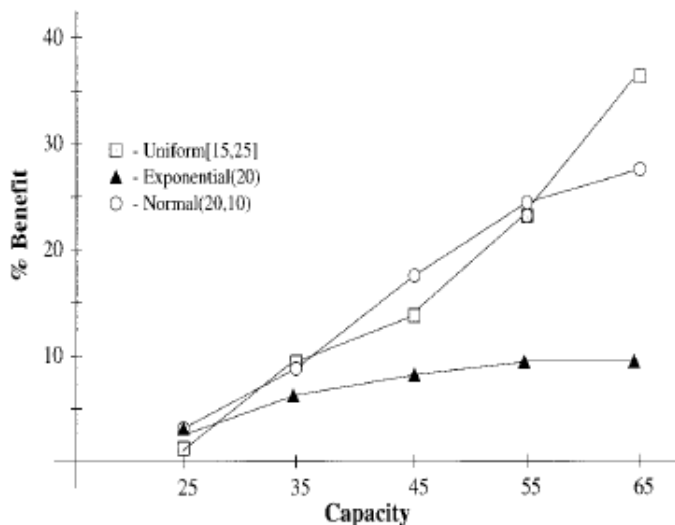


Figure 1.4 Plot of % Benefit Versus Capacity Between Model 1 and Model 2 (Gavirneni, et al., 1999)

1.2 Research Objective and Motivation

Our research focuses on optimal ordering and production decisions for innovative products with demand forecast updating under supply constraints. Three scenarios are investigated from different perspectives in this thesis. We analyze the ordering decisions from a retailer’s perspective in the first and second scenarios. A single-ordering model with multiple ordering opportunities is investigated in the first scenario, while a two-ordering model with two ordering opportunities is studied in the second scenario. Finally, we extend the analysis to a supply chain context, where the retailer’s, manufacturer’s and supply chain’s decisions are investigated. Single-order models with time and quantity constraints are studied in Scenario 1, while two-order models with quantity constraints are investigated in Scenarios 2 and 3.

Research Objective 1: Investigate newsvendor problems with demand forecast updating and supply constraints

For short-life cycle products, one of the major challenges in management is how to match supply with demand to minimize overage and underage costs. Although the conventional newsvendor model is a useful way to solve this problem, there are some shortcomings. The conventional newsvendor model is based on the following assumptions:

1. It is a single period inventory problem.
2. Demand is uncertain and follows a probability distribution known at the beginning of the period.
3. Deliveries are made in advance of demand without considering replenishment lead-time.
4. Due to demand uncertainty, there are overage and underage costs, which are proportional to the amounts of the overage and underage respectively.

The impact of replenishment lead-time known as supply lead-time is not considered in the conventional newsvendor model. However, replenishment lead-time can play a critical role in practical supply chain management. The demand distribution is directly related to the replenishment lead-time. For a shorter lead-time, the demand forecast can be updated closer to the selling season, and thus the forecast accuracy is improved. In other words, when the lead-time is short, the variance of the demand distribution in Assumption 2 becomes smaller.

Moreover, due to limited supplier capacity, supply lead times should be related to order

quantities. Glock (2012) pointed out that lead times may be shortened by decreasing lot sizes. Karmarkar (1987) and Kim and Benton (1995) approximated the relationship between lot size and lead time using a positive linear function. Both of these studies suggested that supply lead times increase along with ordering quantities in many manufacturing systems. However, due to demand uncertainty, a retailer's ordering quantity is unknown to a supplier (or manufacturer) in advance. In this setting, the manufacturer can adopt two possible supply modes. In supply mode A, the manufacturer specifies the available ordering timespan based on the largest estimated ordering quantity. In supply mode B, the manufacturer places a maximum limit on the ordering quantity. This maximum limit decreases as the supply lead time becomes shorter.

Therefore, the manufacturer may impose restrictions on the ordering time or quantity, in addition to charging a higher premium for a shorter lead time. These supply constraints may hamper the benefits of updated demand information. In this study, we are going to analyze the following research questions:

1. What are the relationships among the supply constraints (i.e., the higher purchasing cost, restrictions on the ordering time and quantity)?
2. What are the optimal ordering policies for retailers in the two supply modes? What's the value of demand forecast updating?
3. Which supply mode is more beneficial to retailers? How do the supply constraints differ in the impact on the value of demand forecast updating?

To answer these questions, a single-ordering problem among multiple ordering opportunities

is investigated under the two supply modes from a retailer's perspective.

Research Objective 2: Investigate the optimal emergency orders with updated demand forecast and limited supply.

With more accurate demand information available closer to the selling season (Kaminsky & Swaminathan, 2001), QR has been shown to efficiently reduce overstock and stock-outs and improve customer service levels in various settings (T.-M. J. Choi, et al., 2006; M. Fisher & Raman, 1996; Iyer & Bergen, 1997; Doğan A Serel, 2009). Using the strategy of QR, retailers would place an emergency order at a later stage to adjust the initial order and match supply with demand better based on the improved demand forecast. In practice, an emergency order can be sourced from a secondary supplier (Yan, et al., 2003), market (H. Lee & Whang, 2002) or supply mode (Donohue, 2000).

Compared with the regular opportunity, the emergency opportunity is charged at a higher cost, and the quantity is limited by a maximum limit. As time approaches the start of the selling season, it becomes increasingly difficult for suppliers to either make short-term reallocations of production resources (Donohue, 2000) or deliver fast orders to retailers (T.-M. Choi, Li, & Yan, 2004). Therefore, suppliers will charge a higher cost to provide the emergency ordering opportunity. Furthermore, due to suppliers' limited capacity, the availability of products becomes harder to guarantee with a short lead time. Thus, suppliers would impose a quantity constraint on a later order to make effective production plan and

avoid excessive supply capacity (H. Chen, Y. Chen, C.-H. Chiu, T.-M. Choi, & S. Sethi, 2010a). Moreover, suppliers are usually willing to increase the capacity in advance if retailers make a reservation for future capacity to ensure a large supply commitment (Erkoc & Wu, 2005). Therefore, the emergency ordering quantity is considered to be constrained by a maximum limit, which is predetermined at the beginning.

Scenario 2 analyzes a retailer's ordering problem in a two-stage newsvendor setting. The retailer can place a regular order and a fast but expensive emergency order with updated demand forecast and limited ordering quantity. Examples of such a model can be found in a variety of practical situations. For example, the delayed product differentiation (e.g., the HP and IBM cases in H. L. Lee and Tang (1997)) allows manufactures to postpone some production processes. Retailers can order later at a higher cost, but no more than the amount of the common part that has been processed. For another example, in China, to deal with the changing demands for products in the automobile and PC industries, retailers would place an urgent order by paying a higher wholesale price than the regular order (Jianli Li & Liu, 2008). However, the urgent ordering quantity cannot exceed the manufacturer's reserved capacity. In addition, North and European companies outsource products from lower-cost countries (e.g., Zara in Ghemawat, Nueno, and Dailey (2003)). Although the international supply usually takes a long lead time, the available amount is large in contrast to the limited local supply.

In this context, a retailers faces a trade-off: order earlier to enjoy a lower cost and a reliable supply but with less accurate demand information, or order later to obtain more accurate demand information but suffering a higher cost and a risk of insufficient supply. The

trade-off is investigated in a two-ordering model from a retailer’s perspective.

Research Objective 3: Analyze production strategies and coordination for a capacitated supply chain.

The clinical trials supply chain for GSK is shown in Figure 1.5. Before the clinical trial begins, the Primary Pilot Plant (PPP) begins to order raw materials from suppliers to prepare for the production based on the master production schedule provided by the supply chain manager. Once raw materials are available, the PPP begins to produce the Active Pharmaceutical Ingredient (API). Then, the API is transformed into different forms of drugs at the Secondary Pilot Plant (SPP). After that, drugs are packaged and labeled at the Global Distribution Centers (GDC) and delivered to distribution centers or clinical sites throughout the world. In the last stage, patients get the drugs from the clinical sites.

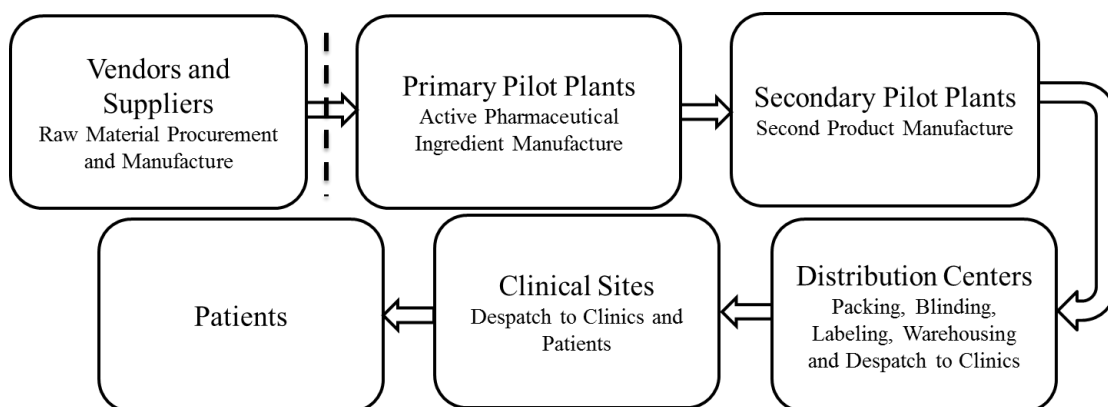


Figure 1.5 GSK Clinical Trials Supply Chain

In this process, there are demand fluctuations caused by patient recruitment uncertainty, flexible dosage during trials, test results (successful or failed), lead-time for regulations during the shipment, and etc. In the presence of these demand fluctuations, it is critical for supply chain managers to take advantage of updated demand information and decide the optimal quantity of production at the optimal time.

At the beginning, the supply chain manager has limited information about the demand. As time goes by, better demand forecast can be acquired since the variables affecting the clinical supplies become less uncertain. For instance, more information about patient population and treatment length becomes available, and the range of possible dosages allocated to the patient narrows down as the trial continues. Given the fact that better demand forecast can reduce costs effectively, it may not be the best strategy to produce all drugs at the beginning. Supply chain managers may break the production into two production runs. Due to more accurate demand information, the second production run allows the adjustment of the initial decision to better match supply with demand. Many applications of models with two production runs have been reported in the literature (Donohue, 2000; Weng, 2004; Zhou & Wang, 2009). They showed that adding a second production is beneficial to supply chains.

However, the lead time for procuring customized materials is long. These materials for the second production should be prepared in advance. According to GSK supply chain managers in clinical trials, the average ordering lead time for customized raw materials from suppliers is about 4.5 months, constituting a considerable part (e.g., around 40%) of the total production cycle time. In contrast, other common materials can be produced by themselves in

around 1.5 months. Moreover, compared with the cycle times at other stages of the production process, the lead time for the customized materials is more difficult to reduce as it depends on the external suppliers. Hence, the amount of the customized materials for the second production (i.e., the first stage in Figure 1.5) needs to be determined before more demand information is acquired.

Inspired by this setting, Scenario 3 investigates a generalized supply chain model with a two-mode production. In this model, based on the early demand information, the supply chain starts the speculative production and prepares the raw materials for the reactive production. Based on the improved demand information, the supply chain makes the reactive production. By investigating this model, we could make the following contributions:

1. Check whether adding a second production improves the supply chain profit if materials must be prepared in advance.
2. Explore how coordinating the supply chain members to maximize the supply chain profit.

1.3 Organization of the Report

In chapter 2, we distinguish our studies from the previous ones with demand information evolution in inventory management. First, we review different traditional models describing the demand forecast evolution. Then, we introduce various inventory models that incorporate demand forecast updating. Finally, we point out the differences of our models from others.

In chapter 3, we investigate an extension of the newsvendor model with demand forecast updating under supply constraints. The optimal decisions on the ordering time and quantity

are studied in supply mode A and B.

Chapter 4 introduces a two-stage newsvendor system with a regular and an emergency order. The emergency order can allow companies to adjust the initial order based on the improved demand information, but the available ordering quantity is limited and the cost is high. By considering the demand forecast updating, ordering cost and quantity constraint, both regular and emergency ordering quantities are determined.

Chapter 5 presents a two-mode production supply chain, where a fast reactive production is utilized to take advantage of improved demand information but the production quantity is limited by the preprocessing in advance. We characterize optimal decisions in the two-mode production system and compare it with other single-mode production systems. In addition, the coordination is investigated if there are two members in the supply chain.

Chapter 6 summarizes the findings and provides directions for future research.

Note that the scope of notations in each chapter is limited to the corresponding chapter.

Chapter 2. Literature Review

In this research, we are particularly interested in demand information evolution in inventory management. Different models describing demand forecast evolution are firstly introduced. Then, the literature on the inventory management incorporating demand information updating is reviewed. Finally, the differences of our studies from the literature are discussed.

2.1 Demand Forecast Updating Models

Previous studies associated with demand information updating have adopted a variety of different models to describe demand forecast evolution (Das & Abdel-Malek, 2003; Hausman, 1969; H. L. Lee, So, & Tang, 2000). These models can be mainly classified into three categories: Time Series, Bayesian analysis and Markovian Process.

2.1.1 Time Series Models

The first category to update demand forecast is by time series. This approach focuses on the correlation between consecutive demand realizations. Many studies adopted time-series as demand forecast models. H. L. Lee, et al. (2000) suggested that the demand process faced by a retailer is a simple auto correlated AR (1) process, i.e.,

$$D_t = d + \rho D_{t-1} + \varepsilon_t, \quad (2.1)$$

where D_t denotes the demand in period t , d and ρ are known constants, and ε_t is i.i.d. normally distributed with parameters $(0, \sigma)$. Through this demand forecast process, they quantified the benefits of initiatives about demand information sharing in a two-level supply

chain. Graves (1999) and Miyaoka and Hausman (2004) used an integrated moving average process (IMA) as follows,

$$D_1 = \mu + \varepsilon_1,$$
$$D_t = D_{t-1} - (1 - \alpha)\varepsilon_{t-1} + \varepsilon_t,$$

where μ and α are known constants. Gilbert (2005) adopted a general class of autoregressive integrated moving average (ARIMA) time-series models for the demand in supply chains by assuming the current demand is a linear function of the past demand and past “shocks” plus the shock of the current period. They estimated forecast errors as the shocks. Aviv (2003) described the underlying demand process in a linear state space form by expressing the demand realization during each period as a linear function of a state vector evolving as a vector autoregressive time series. They further utilized the Kalman filter technique to calculate the minimum mean square error forecast of future demands and derived the forecast adaptive inventory replenishment policy.

2.1.2 Bayesian Analysis Models

Bayesian analysis, which was first introduced by Dvoretzky, Kiefer, and Wolfowitz (1952), is the second category to update demand forecast. In Bayesian models, a demand distribution is chosen from a family of distributions with uncertain parameters. New information is used to update the uncertain parameters of the demand distribution. Bayesian analysis framework has two approaches. The first approach treats the demand process as a normal distribution with unknown mean and known variance. The demand uncertainty has two levels. The first level is attributed to the inherent demand uncertainty and the second level captures the uncertainty

from the mean at time 0. The second approach describes the demand process as a normal distribution with both unknown mean and unknown variance. The first approach of Bayesian analysis has been commonly used in literature (T.-M. Choi, et al., 2004; Iyer & Bergen, 1997). This approach assumes that the demand variance is known and decreases with the demand information updating.

2.1.3 Markovian Process Models

The last category is Markovian process. Markovian forecast revision has two approaches: additive and multiplicative. The additive approach assumes that the new information is additive and that the forecast revisions are independent and normally distributed. In the multiplicative approach, it is assumed that the new information is multiplicative and that the forecast revisions are independent and lognormally distributed. Hausman (1969) firstly introduced a demand forecast process model for inventory management and showed that the data-generating process of the forecast changes has the quasi-Markovain property. Heath and Jackson (1994) proposed the Martingale Model of Forecast Evolution (MMFE) to describe the evolution of demand forecast. They demonstrated the significance of good forecast by combining the MMFE with the linear programming model for the scheduling of a production and distribution system in a rolling horizon. Moreover, MMFE was validated by the empirical sales and forecast data by Graves, Kletter, and Hetzel (1998). Aviv (2001); (2002) and (2007) modeled the process of the available market signals and the forecast explanatory power using the MMFE framework.

In this thesis, we adopt the MMFE to model the demand forecast evolution for the following reasons. First, according to Chen and Lee (2012), most demand models using time series can be transformed into a special case of the generalized MMFE. Secondly, Milner and Kouvelis (2005) showed that MMFE models are more capable than Bayesian models to capture the demand process when the demand is determined by exogenous forces. Our research considers the evolving demand information based on the exogenous and continuous shocks from the market trend reports, expert estimates, trade shows and fluctuations in business environment. Finally, compared with other models, the MMFE model is a simpler and more direct way to describe the demand evolution when the demand forecast becomes more accurate as time evolves.

2.2 Inventory Management with Demand Forecast Updating

There is an extensive literature on studying demand forecast updating in a variety of different environments, which can be broadly classified into three categories: the production plan, purchase plan and supply chain plan. The supply chain plan is the integration of the production plan and purchase plan.

2.2.1 Demand Forecast Updating into the Production Plan

A number of papers have discussed the production plan with demand information evolution. These studies focused on utilizing the advance demand information to improve the accuracy of the demand forecast and make better decisions on the time and quantity to produce, or the capacity to build.

Kaminsky and Swaminathan (2001) proposed a model to capture the demand forecast evolution and utilized it to make production decisions for a single product with terminal demand in a capacitated environment. When there is time t left, the forecast for the terminal demand is described as the band with the lower bound a_t and the width w_t . They assumed that the progression of w_t is described as

$$w_{t-1} = w_t - \alpha_t,$$

where α_t is a positive and known parameter, which captures the fact that the forecast becomes more accurate as it is closer to the terminal demand. It is also assumed that a_{t-1} is the random variable in the range $[a_t, a_t + w_t - w_{t-1}]$ to make sure that

$$[a_{t-1}, a_{t-1} + w_{t-1}] \subseteq [a_t, a_t + w_t].$$

As a result, the range of future forecasts is contained in the current range $[a_t, a_t + w_t]$. Under this demand forecast model, they proved the optimality of a threshold policy that determines whether to produce in each period. In addition, they provided four simple heuristics for the holding cost case and demonstrated that two of them are very accurate in terms of the optimal solution. Finally, they investigated the effects of the information update patterns on the costs and provided the conditions for the most useful information update.

Özer and Wei (2004) analyzed a periodic-review, stochastic, capacitated, finite and infinite horizon production system with the advance demand information for both zero and positive fixed production cost, where unsatisfied demand is backordered. The advance demand information is collected through the advance orders placed by customers since customers can place an order for a future delivery. For each period t , the demand is assumed

as a vector D_t denoted by $(D_{t,t}, \dots, D_{t,t+N})$. $D_{t,s}$ means that the orders are placed at time t and delivered at time s , and N represents the length of the information horizon. At the beginning of the current period t , the demand at the future period s can be decomposed into two parts: the observed part $\sum_{r=s-N}^{t-1} D_{r,s}$ and the unobserved part $\sum_{r=t}^s D_{r,s}$ as shown in Figure 2.1. Under this demand information model, the state of the period t is defined by the modified inventory position and the observations of the demand beyond the lead-time. As a result, for the case with zero fixed cost, they established the optimal policy which is a state dependent modified base stock policy. For the case with positive fixed cost, they proposed a state dependent threshold policy if the production decision is restricted to either full capacity or not at all. In addition, they indicated that the advance information is beneficial for companies to reduce inventories, make good use of the available capacity, and be responsive to the changing demand.

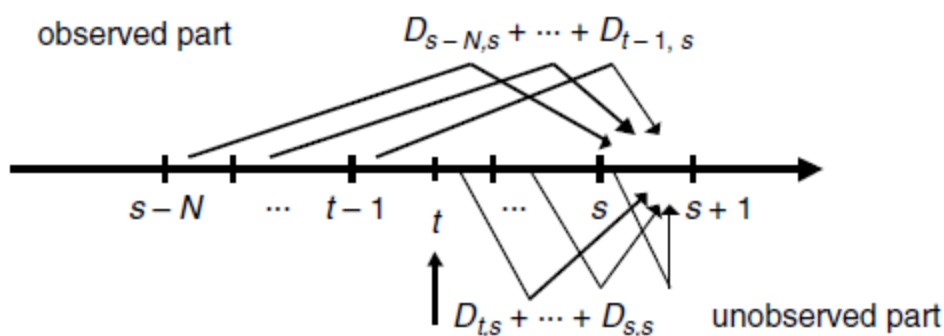


Figure 2.1 Observed and Unobserved Part of the Demand (Özer and Wei 2004)

Boyacı and Özer (2010) investigated a capacity planning strategy in a setting where the

advance demand information can be acquired by the advance selling prior to building the capacity. Market demand is modeled by an iso-elastic, price-sensitive aggregated model, which has a multiplicative form of uncertainty. For each period t , the actual demand is dependent on the price p_t , the uncertain market size ζ_t , and the advance sales information available at the beginning of the period defined by the cumulative commitments q_t and the relevant historical information μ_t . Then, the actual demand is described by

$$D_t(p_t|q_t, \mu_t) = f_t(q_t, \mu_t)\zeta_t p_t^{-b},$$

where $f_t(q_t, \mu_t)$ indicates the future demand predicted by the advance sales information (q_t, μ_t) , ζ_t is the independent random variable with increasing failure rates, and b is the price elasticity with the same value across periods. The form of the demand implies that the information from the advance selling can provide the predictive value for the current and the future demand. Hence, the manufacturer can improve the demand forecast accuracy based on the periodic advance commitments before the regular sales season, and then makes a better capacity decision. However, the manufacturer may pay additional costs and have the risks of selling capacity at a lower price for delaying the capacity decision. In order to balance these multiple trade-offs, they proposed the control-band policy to determine the optimal time to stop the advance sales and the optimal capacity level for the exogenous and optimal pricing policies. Finally, they showed that the strategy of the advance selling improves the profits and found the conditions under which the strategy is the most valuable.

All the above studies have presented how manufacturers take good advantages of demand information in their short-run and long-run production plans. According to this

literature, except Kaminsky and Swaminathan (2001), manufacturers mainly use the records about preorders placed by customers to improve demand forecast accuracy. However, our research concerns the inventory control for demand information sourcing from market trend reports, expert estimates, and trade shows. This demand information usually becomes more accurate as time goes on. In other words, the current demand information is the most accurate until now, capturing all the past information. Therefore, we adopt the Martingale Model of Forecast evolution (MMFE) approach (Graves, et al., 1998; Hausman, 1969; Heath & Jackson, 1994) to model the demand updating progress.

2.2.2 Demand Forecast Updating into the Purchase Plan

Many researchers have investigated demand forecast updating in the purchase plan (Gurnani & Tang, 1999; J. Li, Chand, Dada, & Mehta, 2008; Y. Wang & Tomlin, 2009). The studies showed that the buyer has to evaluate the trade-off between the benefit of the updated information and the additional cost for the information. As a result, they have established the optimality of the ordering strategy for the buyer to obtain the maximum profit. Along with this trend, ordering decisions for the models with two ordering opportunities or multiple ordering opportunities have been studied. These models are introduced in the next two sections.

2.2.2.1 Two Ordering Opportunities

In this section, we consider a case that a buyer can order twice (denoted as two instants) from a supplier, where demand information can be updated at the second instant.

Gurnani and Tang (1999) addressed the trade-off between improving the demand information and potentially increasing the unit cost at the second instant prior to the single selling season. They investigated the impact of forecast updating by assuming a bivariate normal joint distribution of the two related variables (i.e., the information variable $I \sim N(m, s)$) and the demand variable ($D \sim N(\mu, \sigma)$) with the correlation coefficient ρ . Given the demand information between the first instant and the second instant, they concluded that the demand follows the normal distribution with the parameters μ' and σ' as follows,

$$\begin{aligned}\mu' &= \mu + \rho(I - m)\sigma/s, \\ \sigma' &= \sigma\sqrt{1 - \rho^2}.\end{aligned}\tag{2.2}$$

Equation (2.2) shows that the demand forecast becomes more accurate when the information between two instants is available since σ' is smaller than σ . Under this demand information model, they proposed a nested newsvendor model for the retailer to determine the optimal order quantities for the following two cases: worthless information ($\rho = 0$) and perfect information ($\rho = 1$). Moreover, they identified the conditions for the retailer to wait for the demand information and order at the second instant.

Huang, Sethi, and Yan (2005) investigated the ordering policy of a buyer, who has a second ordering opportunity provided by a purchasing contract to adjust the initial order with fixed and variable costs based on the improved demand information. They denoted the demand in stage 2 (second opportunity) as a random variable D with the p.d.f. $f_D(\cdot)$, and assumed the observed signal ψ in stage 1 (first opportunity) with the p.d.f. $f_\psi(\cdot)$. The conditional distribution $F_{D|\psi}(\cdot)$ represents the updated demand forecast. Under general

demand and signal distributions, they showed that the optimal tuning policy in stage 2 is a two-sided (s, S) policy. In addition, they obtained the closed form for the optimal order quantity when the demand distribution conditioned on the signal can be parameterized by the signal as its location parameter and the signal distribution observed in Stage 1 is PF2 (polya frequency function of order 2). Finally, they assumed that ψ and D follow the uniform distributions. Under this demand update model, they provided a critical value of the contract exercise cost to decide whether to invest in improving the forecast or not. They also proved that the contract policy has advantages over other risk hedging methods, such as product substitution, only if the contract exercise cost is less than the critical value.

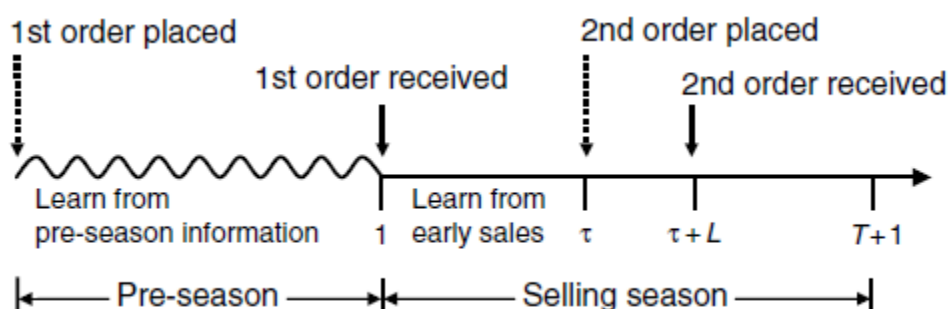


Figure 2.2 Timeline of the Model with Two Procurement Opportunities (J. Li, et al., 2008)

J. Li, et al. (2008) analyzed a framework of the model with two procurement opportunities, in which the second order is placed during the selling season to adjust the preseason stocking decision, depicted in Figure 2.2. The second order opportunity enables the decision maker to not only have accurate information over the remaining demand, but also

take advantage of the current inventory level by adjusting the initial order quantity. The demand at each period t during the selling season is assumed as a random variable with the probability density function $f_t(\cdot)$. Given the time of the second order, they outlined the optimal policy structure for the order quantity at each instant under some general conditions. They showed that this optimal policy structure can be extended to incorporate the forecast revision during the preseason and the demand update based on the observed demand in the selling season. When the time of the second order is determined dynamically, they provided the conditions where (s, S) type is the optimal policy for the second order.

In Gurnani and Tang (1999) and Huang, et al. (2005), the demand forecast is updated and used to place the second order before the selling season. In J. Li, et al. (2008), the second ordering decision is made based on both the preseason information and early sales during the selling season. In all these studies, it has been showed that the optimal second ordering quantity depends on the demand forecast revision observed at the second stage.

2.2.2.2 Multiple Ordering Opportunities

In the previous section, we discussed two ordering opportunities modes with different costs and demand forecast accuracies. Some studies have extended this analysis to the case with multiple ordering or delivery opportunities.

T.-M. Choi, et al. (2004) studied a newsvendor-type retailer's optimal ordering policy with multiple deliveries mode and updated market information. In their paper, the demand forecast at the current stage is updated by the market information gathered at the earlier stage

using the Bayesian method. The method describes the demand forecast process using a normal distribution with unknown mean and known variance. Under this forecast evolution structure, they found that the demand uncertainty decreases after the update. Figure 2.3 shows that the cost of the faster delivery mode is higher, but the demand forecast is improved due to obtaining more market observations. In order to strike the balance between delivery costs and demand information uncertainty, they formulated a multi-stage dynamic optimization model to obtain the optimality of the threshold policy and provided an algorithm to determine the ordering decisions of the retailer. Moreover, the simulation results show that the benefit of the optimal policy is substantial for the cases with large prior demand variance, relative small inherent demand variance, and low profit margin.

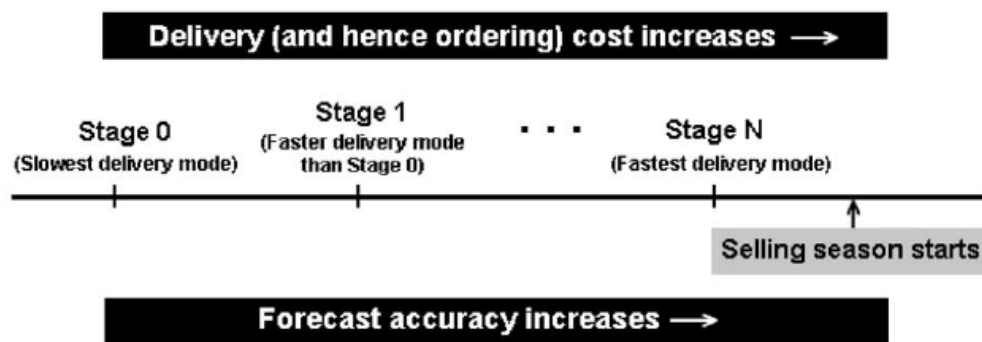


Figure 2.3 Basic Model Structure of the Single Ordering Problem (T.-M. Choi, et al., 2004)

Y. Wang and Tomlin (2009) investigated a trade-off between lead-time risk and demand risk faced by the companies sourcing products from multiple distant suppliers. They formulated a demand forecast process using the discrete time Multiplicative Martingale

model of Forecast Evolution (MMFE) with residual uncertainty. Under this model, the ratios of the successive forecasts before the selling season T are independent and identically distributed random variables, following the lognormal distribution with parameters (μ, σ) . Assuming that the forecast is x_t at the current time t , the demand forecast X_T at time T is lognormally distributed with parameters $((T-t)\mu + \ln x_t, \sqrt{(T-t)\sigma^2})$. They supposed that the ratio of the actual demand X_D and X_T follows the lognormal distribution with parameters $(\hat{\mu}, \hat{\sigma})$, representing the residual uncertainty. Given the current forecast x_t , X_D is lognormally distributed with parameters $((T-t)\mu + \hat{\mu} + \ln x_t, \sqrt{(T-t)\sigma^2 + \hat{\sigma}^2})$. Under this demand forecast updating model, they found that the optimal order quantity is dependent on the forecast evolution, but the optimal procurement time does not depend on the forecast evolution when the supply lead-time is uncertain. Moreover, they proved that the firm with forecast should purchase earlier than the firm without forecast under the deterministic lead-time setting, but should not necessarily buy earlier in the random lead-time setting. They also observed that more efficient forecast updating weakens the impact of lead-time uncertainty.

T. Wang, Atasu, and Kurtulus (2012) considered the ordering policy for the newsvendor-type problem with demand forecast updating and multiple ordering opportunities that are provided by a variety of different sources with different lead-times and costs. In their paper, the demand forecast evolution $\{D_n, n = 1, \dots, N + 1\}$ is modeled by the additive and multiplicative MMFE, where D_{N+1} denotes the final demand in the selling season. Under the additive MMFE, the demand forecast at time n is

$$D_n = u + \varepsilon_2 + \dots + \varepsilon_n,$$

where the expected demand u is the forecast D_1 in the first period, and the forecast adjustments ε_i independently follows the normal distribution with parameters $(0, \sigma_i)$. In this case, the cumulative forecast adjustment until period n is denoted as $I_n = \sum_{i=2}^{i=n} \varepsilon_i$. For the multiplicative MMFE case, the demand forecast at time n is given by

$$\log D_n = \log D_1 + \varepsilon_2 + \dots + \varepsilon_n, \quad (2.3)$$

where ε_i independently follows the normal distribution with parameters $(-\sigma_i^2/2, \sigma_i)$.

Assuming that the expectation u of the logarithm of demand in period 1 is

$$u = \log D_1 - \sum_{i=2}^{i=N+1} \sigma_i^2 / 2.$$

Equation (2.3) can be rewritten as

$$\log D_n = u + (\varepsilon_2 + \sigma_2^2) \dots + (\varepsilon_n + \sigma_n^2).$$

Hence, the mean-adjusted cumulative forecast adjustment I_n equals $\sum_{i=2}^{i=n} (\varepsilon_i + \sigma_i^2/2)$. Under the additive or multiplicative MMFE, given the observation of I_n , the actual demand D_{N+1} is normally or lognormally distributed with parameters $(\mu + I_n, \sum_{i=n+1}^{i=N+1} \sigma_i^2 / 2)$. Under this setting, they obtained the optimal multi-ordering policy which is a state-dependent base stock policy with the base-stock level linearly or log-linearly depending on the information state I_n for the additive or multiplicative MMFE. Numerical results demonstrate that the multi-ordering policy is always more valuable than the single-ordering one. In addition, they found that the dynamic ordering policy could significantly improve the profits in the multi-ordering scenario, but not always in the single-ordering scenario.

T.-M. Choi, et al. (2004) and Y. Wang and Tomlin (2009) addressed the optimal ordering

time among multiple ordering opportunities, and T. Wang, et al. (2012) considered the optimal ordering policy with multiple orders without fixed cost. All these studies have shown the effect of the demand forecast updating process on the ordering decisions. In particular, they focused on how the ordering decisions depend on demand forecast states. The optimal ordering policies can boil down to the static policy, threshold policy (base-stock policy) and control band policy, which can be found in Suresh P Sethi, Yan, and Zhang (2001), Iida and Zipkin (2006) and Oh and Özer (2013).

2.2.3 Demand Forecast Updating into the Supply Chain Plan

There have been a number of studies dealing with supply chains incorporating demand forecast updating. The imperfect demand information has impact on not only manufacturer's production plan, but also retailer's ordering strategy. Therefore, it is important for both parties to estimate the potential value of the demand information in decentralized and centralized systems.

Supply chain coordination receives a great deal of attention from researchers and practitioners in various settings. Due to double marginalization, when supply chain members optimize their own profits, total profits in decentralized systems are lower than those in centralized systems. To improve the total supply chain profits, the members can be aligned through contracts, such as return policies (Emmons & Gilbert, 1998; Pasternack, 1985), Quantity flexibility (Tsay, 1999; J. Wu, 2005), capacity reservation (Bonser & Wu, 2001; Dogan A Serel, Dada, & Moskowitz, 2001) and revenue sharing (Gérard P Cachon &

Lariviere, 2005; Yao, Leung, & Lai, 2008). Refer to Gerard P Cachon (2003), Kanda and Deshmukh (2008) and X. Li and Wang (2007) for a broader overview of supply contracts.

Donohue (2000) studied a contract between one manufacturer and multiple distributors with two production modes, one of which is faster but more expensive allowing distributors to update their demand forecasts at a later stage. The demand forecast is updated by using the conditional probability distribution of the total demand given the market information observed at the later stage. In order to coordinate the channel, they analyzed the traditional contract, i.e., the pricing schemes with parameters (w_1, w_2, b) , where w_1 and w_2 denote the prices for the two modes and b is the return price paid by the manufacturer at the end of the season. As a result, they characterized the efficient conditions on these parameters to maximize the profit of the channel and showed that these conditions change with the improvement of the demand forecast accuracy between the two stages and the availability of the information at the manufacturer.

Barnes-Schuster, et al. (2002) investigated how options in supply chain contracts affect a buyer-supplier system for a short life-cycle product using a two-period model with symmetric correlated demand (see Figure 2.4). Demands during the two periods are assumed to be normally distributed and correlated with a correlation coefficient. They demonstrated how the buyer reacts to the demand forecast updating in the second period with flexibility provided by the options. They provided the sufficient conditions with linear transfer prices to ensure channel coordination. However, under these conditions, the supplier cannot get any profit. They proved that the return policy is useful to coordinate the channel with positive profits for

the supplier in a subset of these conditions. Finally, they pointed out that the options improve the joint profits of the channel, especially with a higher demand correlation.

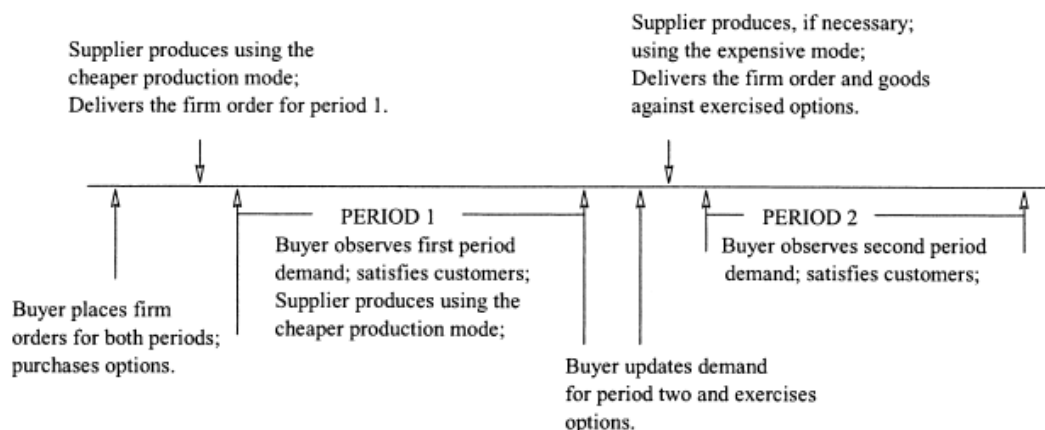


Figure 2.4 Time line of Buyer-Supplier Decisions (Barnes-Schuster, et al., 2002)

Özer, Uncu, and Wei (2007) proposed and analyzed a dual purchase and price-only contract. In the contract, the manufacturer provides a discount to allure the retailer to order before the forecast update. They assumed that demand is the sum of the forecast update X and the residual market uncertainty ε . X and ε are continuous random variables. Before the forecast update, X is assumed to be distributed with the p.d.f. $f(\cdot)$ and is realized after a market research. Under this dual purchase contract, they showed that the retailer's and joint profits are improved compared to the wholesale price contract. Moreover, they derived the conditions to increase both the manufacturer's and retailer's profits, under which the strict Pareto improvement is created.

These studies have investigated the role of demand forecast updating in supply chain

management from the perspectives of the buyer, manufacturer and channel. In different scenarios, they adopted different types of contracts and optimized the contract parameters to ensure the channel coordination.

2.3 Research Differences

Among the forecast-updating literature, most of them assumed that the supply capacity is unlimited. However, the assumption is not always justifiable in practice, especially when the supply lead time is short. In our three scenarios, the limited capacity is considered in different forms in different environments. The contributions of our models to the forecast-updating literature are illustrated by explicitly differentiating them from the most relevant papers.

Scenario 1 investigates a similar optimal ordering problem with multiple ordering opportunities, but differs from the aforementioned models in Section 2.2.2.2 in several ways. First, the ordering decision is investigated in supply mode B, where the decreasing maximum ordering quantity is specified. To the best of my knowledge, this is the first one to consider quantity restrictions in this stream. Second, the latest ordering time is considered in supply mode A. Although it is a special version of the single-ordering model proposed by Wang, et al. (2012), it is found in our case that whether to use demand forecast updating strongly depends on the latest ordering time. Third, we use queuing theory to analyze the relationships among supply lead times, the procurement cost function, and ordering time and quantity restrictions in detail. The relations among them are rarely examined in the existing literature.

Scenario 2 is similar to inventory models with two ordering opportunities in Vlachos and

Tagaras (2001), Jianli Li and Liu (2008) and H.-W. Chen, et al. (2013), which consider a constraint on the second replenishment quantity for a single product. Vlachos and Tagaras (2001) analyzed a periodic review inventory system with a regular supply mode and an emergency mode, considering a capacity constraint on the emergency channel. Through approximate cost models, two near optimal emergency ordering policies were provided and compared to study the effectiveness of the emergency replenishment. In contrast to Vlachos and Tagaras (2001), we develop an optimal ordering policy for a two-stage newsvendor problem. Jianli Li and Liu (2008) investigated a supply chain coordination policy in an extended newsboy model, where the retailer's second order is limited by the manufacturer's reserved capacity. H.-W. Chen, et al. (2013) studied three different contract types, through which the supplier limits the fast-ship commitment within the short selling season. Both of these two papers dealt with two-stage newsvendor problems similar to ours. However, they considered only the special case, where perfect demand information is observed at Stage 2. This study provides analytical and numerical results for general cases (i.e., imperfect demand information). To the best of our knowledge, this study is the first one to investigate the issue of the emergency ordering opportunity with a quantity constraint and imperfect demand information in a two-stage newsvendor model.

Scenario 3 investigates a two-stage supply chain issue with the second replenishment quantity limited by the amount of the preprocessing material. Barnes-Schuster, et al. (2002) considered purchasing options in advance for the second replenishment. Our study differs from theirs in several dimensions. First, in contrast with their model where demand occurs at

two periods, in our model demand forecast is updated at two stages before the selling season and demand occurs only once during the selling seasons. Thus, the retailer places two orders in their study but only one order in this study at the beginning of the planning horizon. Second, while their analysis is restricted to a bivariate normal distribution, our study analyzes the model with general demand distributions where the future demand forecast is stochastically increasing in the current demand forecast. Third, although they compared the two-production mode with one single-production mode in numerical examples, our study provides both analytical and numerical results for the comparison between the two-production mode and two single-production modes. Fourth, our study shows that whether the second production is beneficial is related to demand forecast updating process, which was not observed before.

Chapter 3. Newsvendor Problems with Demand Forecast

Updating and Supply Constraints

This chapter investigates an extension of the newsvendor model with demand forecast updating under supply constraints. A retailer can postpone order placement to obtain a better demand forecast with a shorter supply lead time. However, the manufacturer would charge the retailer a higher cost for a shorter lead time and set restrictions on the ordering times and quantities. This prevents a retailer from taking full advantage of demand forecast updating to improve profits. In studying the manufacturer-related effects, two supply modes are investigated: supply mode A, which has a limited ordering time scale, and supply mode B, which has a decreasing maximum ordering quantity. For supply mode A, it is proven under justifiable assumptions that a retailer should order either as early or as late as possible. For supply mode B, an algorithm is proposed to simplify the ordering policy by appropriately relaxing the ordering quantity restrictions. Numerical analysis is conducted to investigate the influence of product and demand parameters on the value of demand forecast updating in the two supply modes. A comparison of the different supply scenarios demonstrates the negative effects of increased purchasing cost and ordering time and quantity restrictions when demand forecast updating is implemented.

3.1 Introduction

This study investigates an extension of the newsvendor model with multiple ordering

opportunities under demand forecast updating and supply constraints. The demand forecast updating process is modeled by using the Martingale model of forecast evolution (MMFE) approach (Graves et al. 1998, Hausman 1969, Heath and Jackson 1994). The supply constraints, including the increasing ordering cost, latest ordering time and decreasing maximum ordering quantity, are investigated by modeling the manufacturing system as a series of M/M/1 queues. The latest ordering time is specified in supply mode A, while the decreasing maximum limit on the ordering quantity is considered in supply mode B.

Based on supply constraints and demand forecast updating, retailers' optimal ordering decisions are investigated in the two supply modes. In supply mode A, under justifiable assumptions, the optimal ordering time is proved to be either the earliest or latest time epoch among the feasible ordering times, independent of demand forecast evolution. Hence, under time constraints, demand forecast updating can be either valuable or completely valueless to a retailer. In contrast, in supply mode B, the optimal ordering time may depend on demand forecast evolution. Due to the quantity restrictions, it is difficult to obtain a closed-form solution for the optimal ordering time. An efficient algorithm is proposed to determine the time by relaxing the restrictions with justifiable assumptions.

Numerical results are analyzed to investigate the effects of product and demand characteristics on the value of demand forecast updating in the two supply modes. If a product has a high initial purchasing cost, low additional purchasing cost or small salvage value, demand forecast updating is more beneficial to the retailer. Furthermore, in supply mode A, the value of demand forecast updating can increase or decrease in price, but it

decreases in price in most of the examined cases in supply mode B. In terms of demand, retailers are found to obtain more benefits from demand forecast updating with higher forecast efficiency in most cases.

Through the queuing models, different supply scenarios can be compared and some managerial insights can be obtained. The numerical results show that retailers tend to choose supply mode B rather than A if the manufacturer is risk-averse, especially when the demand is highly uncertain, and vice versa. Moreover, if the demand uncertainty is low, the benefit of demand information decreases largely due to the increase in purchasing cost. Otherwise, the retailer should persuade the manufacturer to relax the ordering time and quantity restrictions to make good use of the updated demand information.

The rest of this chapter is organized as follows. Section 3.2 presents the relevant models and analyzes optimal ordering decisions. Section 3.3 investigates the role of demand forecast updating in a purchasing plan based on numerical results. Section 3.4 summarizes the results.

3.2 Model

The following problem is considered: before the selling season starts, a retailer purchases innovative products from a manufacturer, who can shorten the supply lead time. In this situation, the retailer can order at other discrete time epochs later than the original ordering time. Figure 3.1 shows the discrete timeline of the model. The original ordering time is the starting period, denoted by time 0 . The latest ordering time is T . The time at which the order is received, i.e., time L , is the beginning of the selling season. On this timeline, time t

indicates that the lead time is decreased by t and that the original lead time is L . Throughout this study, the selling price for the retailer is denoted by r , which is an exogenous parameter. The salvage value per unit at the end of the selling season is denoted by s .

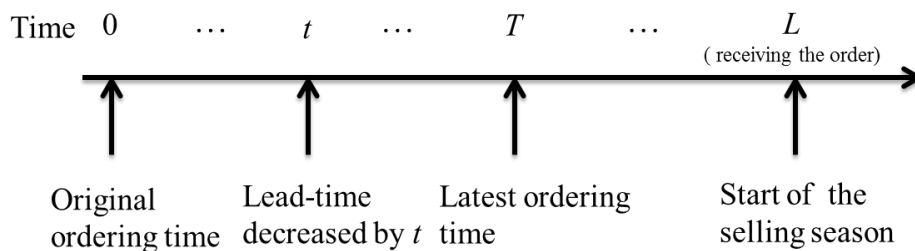


Figure 3.1 Timeline of Model

The retailer is assumed to place an order only once among the multiple ordering opportunities. In practice, the fixed cost of an order comprises booking fees, delivery expenses and administrative costs, among others. The fixed cost could be relatively large. For instance, the delivery cost is high for North American and European companies that source from emerging economies (e.g. China and India). Hence, the retailer can decrease the fixed cost if only an order is placed instead of multiple orders. In addition, the manufacturer prefers the single order from the retailer since he can take advantage of the economies of scale. For example, retailers are encouraged to make a big order, which facilitates manufacturers to make production planning and reduces the setup costs in the semiconductor and fashion apparel industries.

Consequently, the retailer must determine when to order during the planning horizon $[0, T]$ to receive products at the beginning of the selling season. Based on demand information (Section 3.2.1) and supply constraints (Section 3.2.2), optimal ordering policies are proposed for the retailer to achieve the maximum profit.

3.2.1 Demand Forecast Evolution

The demand forecast evolution process can be characterized by a special case of the MMFE proposed by Heath and Jackson (1994). Successive forecasts of future demand are assumed to form a Martingale process. Hence, the conditional expected future forecast and current forecast are equal in value. There are two types of MMFE: the additive (A-MMFE) and multiplicative (M-MMFE) models. It is noted that demand values could be negative with a significant probability in the A-MMFE and be never in the M-MMFE. It is also pointed out that industry forecasts are more likely to be updated in a relative sense (as denoted by the M-MMFE) than an absolute sense (as denoted by the A-MMFE) [42, 43]. Moreover, it is indicated the M-MMFE fits empirical data better than the A-MMFE [32, 33, 38, 44, 45]. Therefore, the M-MMFE is used in this study.

Let X_t denote the forecast made at time t for the actual demand X_D during the selling season. According to the M-MMFE, the successive forecast ratios, $\varepsilon_{t+1} = X_{t+1}/X_t$ form a series of independent and lognormally distributed random variables, where t equals 0 to $T - 1$. Hence, if the demand forecast is x_t at the current time t , X_D can be expressed as follows:

$$\ln X_D | x_t = \ln x_t + \ln \varepsilon_{t+1} + \dots + \ln \varepsilon_T + \ln \varepsilon_D, \quad (3.1)$$

where the final residual uncertainty $\varepsilon_D = X_D / X_T$. Further, the ratios are assumed to be identical random variables. Based on this presentation, ε_t is lognormally distributed with the parameters $(-0.5\sigma_f^2, \sigma_f)$, where σ_f is defined as the forecast efficiency.

Definition 3.1 (*Forecast efficiency*) *The forecast efficiency σ_f denotes the standard deviation of the ratio of successive forecasts and measures how fast the demand uncertainty is resolved by the forecast.*

Note that the demand uncertainty cannot be completely resolved prior to the latest ordering time T . The final residual uncertainty ε_D is assumed to be lognormally distributed with the parameters $(-0.5\sigma_D^2, \sigma_D)$, where $\sigma_D = \sqrt{\alpha}\sigma_f$ and α denotes the residual demand uncertainty factor. Based on this representation, given the demand forecast x_t at time t , X_D is lognormally distributed with the parameters $(\ln x_t - 0.5(T - t + \alpha)\sigma_f^2, \sqrt{T - t + \alpha}\sigma_f)$. Thus, let μ_t and σ_t denote the distribution parameters. The parameters and expected demand $E(X_D | x_t)$ at time t are expressed as follows:

$$\mu_t = \ln x_t - 0.5(T - t + \alpha)\sigma_f^2, \quad (3.2)$$

$$\sigma_t^2 = (T - t + \alpha)\sigma_f^2, \quad (3.3)$$

$$E(X_D | x_t) = e^{\ln x_t - 0.5(T - t + \alpha)\sigma_f^2 + 0.5(T - t + \alpha)\sigma_f^2} = x_t. \quad (3.4)$$

Hence, σ_t decreases with time t , which captures the characteristic of the demand forecast processes, i.e., the forecast becomes more accurate as it approaches the demand

realization. Further, the linear relationship between σ_t^2 and t shows that the demand forecast evolution follows a geometric Brownian motion. In addition, the expected demand always equals the current forecast due to the Martingale property of demand forecast evolution.

3.2.2 Supply Lead Time

Manufacturing cycle time includes the time required for processing (also known as service time) and waiting. Under fixed utilization, the waiting time of a job can be changed by adjusting its priorities. Therefore, a customer's supply lead time can be decreased by increasing the priority and pulling in customer jobs. Consequently, the customer is charged an additional cost to compensate for the delay of the other customers' jobs. Note that although job cycle time can be a random variable, order confirmation is commonly managed by Available-To-Promise (ATP) in practice (Zhao, Ball, & Kotake, 2005). To ensure proper service levels, ATP is created based on historical mean cycle time plus a predetermined safety margin. For ease of representation, without loss of generality, only historical mean cycle time is considered and used in our model to measure the supply lead time (Similar approaches can be found in Su et al. (2005)). Because historical mean cycle time is computed based on prior months' performance in practice, steady-state analysis of queueing theory is used to approximate the historical mean cycle time.

It is assumed that (a) a job must go through a manufacturing system consisting of N single-server workstations in series, (b) the initial arrival process is a Poisson process and (c)

the service time distributions are exponential and mutually independent. According to Burke (1956), all of the workstations in this system see Poisson arrivals. Thus, a system consisting of N workstations can be decomposed into N independent subsystems. Each subsystem is an M/M/1 queue. As a result, the total supply lead time reduction can be obtained by adding up the reductions for each subsystem.

To decrease a retailer's supply lead time, the priorities of the retailer's jobs should be increased. This scenario is formulated as a non-preemptive priority queue with two priority types, where the jobs of all of the other retailers are grouped into one type. In practice, system capacity and monthly throughput are predetermined at the planning level and thus fixed at the time when an order is placed. Hence, system utilization is assumed to be fixed in the current model. Further, the service time B_i is assumed to have a common distribution for the jobs from different retailers at each workstation i . Therefore, an M/M/1 queuing model with two priority types is used to analyze the supply lead times at each workstation. Based on the analysis, the procurement cost function and ordering time and quantity restrictions are obtained in the following sections.

3.2.2.1 Purchasing Cost Function

A retailer's lead time can be decreased by giving higher priority to the order at some workstations. Increasing the priority of some orders can delay other orders, and in such cases additional costs are incurred for the delays. For each workstation i , denote the total utilization by ρ_i ; the utilizations of high and low priority jobs by ρ_{iH} and ρ_{iL} , respectively; and the

mean waiting time without given priority by $E(W_i)$. Let λ_H denote the mean arrival rate of high priority jobs from the underlying retailer, and λ_L that of low priority jobs from the other retailers. In the N workstations system, the system utilization ρ refers to the bottleneck utilization, i.e., $\rho = \max\{\rho_1, \dots, \rho_N\}$.

According to the conservation law (Kleinrock, 1965), the weighted sum of the mean waiting times with non-preemptive priority equals the weighted sum of those without priority. Note that $\rho_{iH} = \lambda_H/\mu_{iH}$, $\rho_{iL} = \lambda_L/\mu_{iL}$, $\rho_i = \rho_{iH} + \rho_{iL}$ and the service rate $\mu_i = \mu_{iH} = \mu_{iL}$. Therefore, the following equation including the mean waiting times $E(W_{iH})$ and $E(W_{iL})$ for high and low priority jobs at workstation i holds:

$$\lambda_H[E(W_i) - E(W_{iH})] = \lambda_L[E(W_{iL}) - E(W_i)]. \quad (3.5)$$

The manufacturer receives orders at the beginning of each period. The total mean arrival rate λ is the average ordering quantity q_{total} from the retailers per period. Furthermore, λ_H is the ordering quantity q from the underlying retailer at some period, and λ_L denotes the quantity of the total orders q_{ot} from the other retailers at that period. From Equation (3.5), the following Equation results for the N workstations system:

$$q\Delta L_H = q_{ot}\Delta L_L, \quad (3.6)$$

where the lead time reduction for the underlying retailer $\Delta L_H = \sum_{i=1}^n (E(W_i) - E(W_{iH}))$ and the delayed time for the other retailers $\Delta L_L = \sum_{i=1}^n (E(W_{iL}) - E(W_i))$, assuming that the priorities of the underlying retailer's jobs are increased at n workstations.

Let c_L be the additional cost for one job unit from the other retailers delayed in one time unit. From Equation (3.6), the total cost of lead time reduction ΔL_H can be written as

follows:

$$C(\Delta L_H) = c_L q_{ot} \Delta L_L = c_L q \Delta L_H. \quad (3.7)$$

Hence, at time t , the total lead time reduction cost is $c_L q t$. This total cost is assumed to ultimately be charged to the underlying retailer. Therefore, for the underlying retailer, the purchasing cost per unit c_t at time t is expressed as follows:

$$c_t = c_0 + c_L t, \quad (3.8)$$

where c_0 is the original purchasing cost per unit at time 0. Rather than deriving from queueing models, the linear purchasing cost in equation (8) is also used by Chen and Chuang [50], Ray and Jewkes [51] and Teng [52].

3.2.2.2 Supply Restrictions on Ordering Time and Quantity

For the N workstations system consisting of a series of M/M/1 queues without priority, the original lead time L and cycle time CT are expressed as follows:

$$L = [CT] = \left\lceil \sum_{i=1}^N \left(\frac{\rho_i}{1 - \rho_i} E(B_i) + E(B_i) \right) \right\rceil, \quad (3.9)$$

where $E(B_i)$ is the mean service time at workstation i and $[.]$ denotes the ceiling function.

For the N workstations system with priority, the minimum mean waiting time for the underlying retailer is $\sum_{i=1}^N E(W_{iH})$, where the high priority jobs have priority over the low priority jobs at every workstation. When the jobs only have priority at some of the workstations, the mean waiting time is larger than $\sum_{i=1}^N E(W_{iH})$ but smaller than $\sum_{i=1}^N E(W_i)$. Thus, according to Adan and Resing (2002), the lead time reduction t is limited by

$$\begin{aligned}
 t &\leq \left\lfloor \sum_{i=1}^N (E(W_i) - E(W_{iH})) \right\rfloor \\
 &= \left\lfloor \sum_{i=1}^N \left(\frac{\rho_i E(B_i)}{1 - \rho_i} - \frac{\rho_i E(B_i)}{1 - qE(B_i)} \right) \right\rfloor,
 \end{aligned} \tag{3.10}$$

where $\rho_i = \rho_{iH} + \rho_{iL}$, $qE(B_i) = \rho_{iH}$, and $\lfloor \cdot \rfloor$ denotes the floor function. The last step holds, as $E(W_i)$ is the mean waiting time without priority and $E(W_{iH})$ is the mean waiting time with high priority, for an M/M/1 queue with two non-preemptive priority types. Note that the lead time reduction t refers to the ordering time t (see Figure 3.1).

In supply mode A, the feasible span of the ordering times can be specified at time 0 according to Equation (3.10), if the largest possible ordering quantity can be estimated from the initial demand information. The policy for estimating the ordering quantity depends on the manufacturer's preference. For example, a conservative manufacturer would overestimate the ordering quantity and place tight restrictions on the feasible ordering time, and vice versa. In this context, if the actual ordering quantity cannot be completely satisfied at a feasible ordering time, the manufacturer must use the additional capacity (e.g., overtime working) to fulfill the order.

In supply mode B, the ordering quantity restriction under the given ordering time t can also be obtained based on the inequality in Equation (3.10). Hence, $q \leq k_t$, where the maximum ordering quantity k_t satisfies the following equation:

$$\sum_{i=1}^N \rho_i E(B_i) \left(\frac{1}{1 - \rho_i} - \frac{1}{1 - k_t E(B_i)} \right) = t. \tag{3.11}$$

From Equation (3.11), the maximum quantity k_t that the manufacturer can supply decreases with time, implying that less of a product quantity is available with a shorter lead time. The properties of the two modes are summarized in the following theorem.

Theorem 3.1

(a) In supply mode A, the latest ordering time T^A is

$$T^A = \left\lceil \sum_{i=1}^N \rho_i E(B_i) \left(\frac{1}{1 - \rho_i} - \frac{1}{1 - q_e E(B_i)} \right) \right\rceil, \quad (3.12)$$

where q_e is the largest possible ordering quantity.

(b) In supply mode B, the maximum quantity k_t satisfies

$$\sum_{i=1}^N \frac{\rho_i}{\mu_i - k_t} = \sum_{i=1}^N \frac{\rho_i}{\mu_i - \lambda} - t. \quad (3.13)$$

Hence, k_t decreases with the ordering time t during $[0, T^B]$, where the latest possible ordering time $T^B = \min\{t | k_t \leq 0\} - 1$.

(c) Under the same condition, $T^A \leq T^B$.

For the sake of clarity, all of the proofs are relegated to Appendix. Theorem 3.1 characterizes the two supply modes that a manufacturer may adopt. In what follows, the optimal ordering decisions for retailers in supply modes A and B are analyzed and addressed. Note that only the values of T^A , T^B and k_t are provided by the manufacturer to retailers.

3.2.3 Supply Mode A

In supply mode A, if a retailer places an order at time $t \in \{0, \dots, T^A\}$, given a demand forecast x_t , the expected profit is

$$\pi^A(t, x_t) = \max_{q>0} \{-c_t q + rE[\min(q, X_D | x_t)] + sE[(q - X_D)^+ | x_t]\}. \quad (3.14)$$

Equation (3.14) can be rewritten as follows:

$$\pi^A(t, x_t) = (r - c_t)E[X_D|x_t] - \min_{q>0} \{(c_t - s)E[(q - X_D)^+|x_t] + (r - c_t)E[(X_D - q)^+|x_t]\}. \quad (3.15)$$

Hence, the expected profit belongs to a traditional newsvendor-type model with the purchasing cost per unit c_t and the conditional demand $X_D|x_t$. Note that $X_D|x_t$ is lognormally distributed with the parameters that satisfy Equations (3.2) and (3.3).

Let $G_t(x_t)$ denotes the profit-to-go function at time t . At any time t , the retailer updates the demand forecast and determines whether to order or wait. If

$$\pi^A(t, x_t) \geq E[G_{t+1}(X_{t+1})], \quad (3.16)$$

then the retailer places an order. Otherwise, the retailer moves to the next period without ordering and repeats it until time $T^A - 1$. If the retailer decides to wait at time $T^A - 1$, the order must be placed at time T^A . Thus, the equations are

$$G_{T^A}(x_{T^A}) = \pi^A(T^A, x_{T^A}), \quad (3.17)$$

$$G_t(x_t) = \max\{\pi^A(t, x_t), E[G_{t+1}(X_{t+1})]\}, \quad (3.18)$$

where $X_{t+1} = x_t \cdot \varepsilon_{t+1}$ and the ratio of successive forecasts ε_{t+1} is lognormally distributed with the parameters $(-0.5\sigma_f^2, \sigma_f)$. Then, the following lemma for the optimal ordering policy is obtained.

Lemma 3.1

(a) *If the retailer is determined to place an order at time t , given demand forecast x_t , the optimal ordering quantity $q_t^{A*}(x_t)$ and the corresponding expected profit $\pi^A(t, x_t)$ are*

given by

$$q_t^{A*}(x_t) = I_t x_t, \quad (3.19)$$

$$\pi^A(t, x_t) = (r - s) \bar{\Phi}(\sigma_t - z_{\beta_t}) x_t, \quad (3.20)$$

where $z_{\beta_t} = \Phi^{-1}\left(\frac{r-c_t}{r-s}\right)$, $I_t = e^{\sigma_t z_{\beta_t} - 0.5\sigma_t^2}$, $\bar{\Phi}(\sigma_t - z_{\beta_t}) = [1 - \Phi(\sigma_t - z_{\beta_t})]$ and

$\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

(b) The optimal ordering time t^{A*} is determined by

$$t^{A*} = \arg \max\{(z_{\beta_0} - \sigma_0), (z_{\beta_1} - \sigma_1), \dots, (z_{\beta_{T^A}} - \sigma_{T^A})\}, \quad (3.21)$$

which is independent of the demand forecast x_t .

Definition 3.2 (Profitable Factor) The profitable factor I_t is defined as

$$I_t = e^{\sigma_t z_{\beta_t} - 0.5\sigma_t^2}, \quad (3.22)$$

which gauges the profitability of a product based on its price, cost, salvage value and demand variance.

Lemma 3.1 (a) shows that given the ordering time t , both the optimal ordering quantity and the corresponding expected profit are positively and directly proportional to the demand forecast x_t . If demand is predicted to be higher during the selling season, the retailer should place an order with a larger quantity and expect a higher profit. According to Equation (3.19), the optimal ordering quantity also increases with the profitable factor I_t , as the retailer should order more product when it is more profitable to do so. Lemma 3.1(b) indicates that the retailer can decide the optimal ordering time at the beginning of the planning horizon. In

other words, the dynamic ordering policy can be simplified to a static policy. This result is consistent with those of studies by Y. Wang and Tomlin (2009) and T. Wang, et al. (2012).

In Equation (3.21), the optimal ordering time depends on the difference between σ_t and z_{β_t} . The two discrete variables σ_t and z_{β_t} represent how the deterministic demand and supply parameters evolve with time, respectively. To study the properties of σ_t and z_{β_t} , the following continuous functions are considered based on Equations (3.3) and (3.8):

$$\sigma(t) = \sigma_t = \sqrt{T^A - t + \alpha \sigma_f}, \quad (3.23)$$

$$z_{\beta}(t) = z_{\beta_t} = \Phi^{-1}\left(\frac{r - c_0 - c_L t}{r - s}\right). \quad (3.24)$$

The following lemma characterizes the monotonicities and convexities of $\sigma(t)$ and $z_{\beta}(t)$.

Lemma 3.2

- (a) *The standard deviation of demand $\sigma(t)$ is a decreasing and concave function in relation to t .*
- (b) *If $z_{\beta}(t) > 0$, then $z_{\beta}(t)$ is a decreasing and convex function in relation to t .*

In the model, $z_{\beta}(t)$ is assumed to be positive, implying that the critical fractile $(r - c_t)/(r - s)$ (i.e., the service level) is larger than 50%. For innovative products with high profit margins, the stock-out cost is higher than the savings obtained from the inventory reduction. That is, the shortage cost outweighs the overstock cost, and the critical fractile should be above 50% (Raz, Druehl, & Blass, 2013). This assumption is also supported by the need for companies to provide high service levels at retailers for innovative products (M. L.

Fisher, 1997).

Hence, both $\sigma(t)$ and $z_\beta(t)$ decrease as time passes. In Equation (3.21), the evolutions of the demand and supply parameters over time have opposing effects on the optimal ordering time. Based on Lemma 3.2, the following theorem can be stated.

Theorem 3.2 *In supply mode A, the optimal ordering time during the feasible timespan $[0, T^A]$ is $\operatorname{argmax}[F(0), F(T^A)]$, where*

$$F(t) = z_\beta(t) - \sigma(t) = \Phi^{-1}\left(\frac{r-c_0}{r-s} - \frac{c_L}{r-s}t\right) - \sigma_f\sqrt{T^A - t + \alpha}.$$

$F(t)$ is a convex function of t . Hence, the optimal ordering time is either the earliest or latest time epoch of the feasible timespan provided by the manufacturer; independent of the demand forecast x_t .

Theorem 3.2 shows that the retailer can determine whether to use demand forecast updating at the beginning of the planning horizon, based on a comparison between $F(0)$ and $F(T^A)$. The demand accuracy tolerance $T_v(T^A)$ is defined to measure the effect of demand accuracy based on the difference between $F(0)$ and $F(T^A)$.

Definition 3.3 *(Demand Accuracy Tolerance) Demand accuracy tolerance $T_v(T^A)$ describes the effect of demand accuracy on a retailer, and is defined as follows:*

$$T_v(T^A) = F(0) - F(T^A). \quad (3.25)$$

A retailer has high demand accuracy tolerance if $T_v(T^A) \geq 0$; otherwise, it is deemed to have low demand accuracy tolerance.

A retailer with high demand accuracy tolerance pays attention to the supply cost, which leads to making orders at the earliest time and at the least cost. In contrast, a retailer with low demand accuracy tolerance cares about demand accuracy and takes full advantage of the demand information by placing orders at the latest time available. Theorem 3.2 also indicates that T^A plays a critical role in determining the optimal ordering time. Because $F(t)$ is convex in t , $T_v(T^A)$ in Equation (3.25) is concave in T^A under fixed initial demand forecast volatility σ_0 . Thus, if T^A is large enough, $T_v(T^A)$ eventually declines to a value less than zero, resulting in the latest order for the largest value of demand forecast updating. However, if T^A is small, then $T_v(T^A)$ may be larger than zero. In this situation, the retailer is concerned more about supply cost rather than demand accuracy, limiting the use of demand forecast updating. Therefore, the effectiveness of demand forecast updating is degraded under the ordering time restriction.

Under the optimal ordering policy in Theorem 3.2, the effects of the parameters on the optimal expected profit and demand accuracy tolerance are investigated as follows.

Theorem 3.3 *The following statements hold for supply mode A:*

(a) *the optimal expected profit $\pi^A(t^{A*}, x_{t^{A*}})$ increases in the salvage value per unit s , but*

decreases in the original purchasing cost per unit c_0 , additional cost per unit c_L , forecast efficiency σ_f and residual demand uncertainty factor α ; and

(b) the demand accuracy tolerance T_v increases in the additional cost per unit c_L and residual demand uncertainty factor α , but decreases in the forecast efficiency σ_f and original purchasing cost per unit c_0 .

The statement of Theorem 3.3(a) is obvious, as costs and demand uncertainty decrease profits and salvage value improves them. The result of Theorem 3.3(b) is also intuitive. As previously noted, σ_f measures how fast demand uncertainty is resolved by the demand forecast, and α measures the ratio of demand uncertainty that cannot be resolved by the forecast. Thus, when σ_f is higher and α is lower, demand forecast updating is more important to the retailer, leading to a decrease in the demand accuracy tolerance. For products with a lower original purchasing cost and a higher additional cost, retailers must pay relatively higher prices for the improved demand forecast. The higher cost of updating demand information decreases the benefit of accurate demand and thus increases the demand accuracy tolerance.

3.2.4 Supply Mode B

In supply mode B, a retailer places an order equal to or less than k_t at time $t \in \{0, \dots, T^B\}$.

Given the demand forecast x_t , the expected profit at time t is

$$\pi^B(t, x_t) = (r - c_t)E[X_D|x_t]$$

$$- \min_{0 < q \leq k_t} \{(c_t - s)E[(q - X_D)^+ | x_t] + (r - c_t)E[(X_D - q)^+ | x_t]\}. \quad (3.26)$$

Hence, the expected profit is of newsvendor type with an additional parameter k_t limiting the ordering quantity. Under the limit of k_t , the optimal ordering quantity at time t is given in Lemma 3.3.

Lemma 3.3 *If the retailer orders no more than k_t at time t , given the current demand forecast x_t , the optimal ordering quantity and corresponding expected profit are as follows:*

$$(a) \quad q_t^{B*} = \min\{q_t^*, k_t\}, \quad (3.27)$$

where the optimal newsvendor-type quantity q_t^* is expressed as

$$q_t^* = e^{\sigma_t \Phi^{-1}\left(\frac{r-c_t}{r-s}\right) + \ln x_t - 0.5\sigma_t^2}; \text{ and} \quad (3.28)$$

(b) $\pi^B(t, x_t)$ is an increasing and concave function in relation to x_t .

Lemma 3.3(a) shows that the retailer should order the optimal newsvendor-type quantity q_t^* if it is smaller than k_t ; otherwise, the retailer should order k_t . Lemma 3.3(b) states that the expected profit $\pi^B(t, x_t)$ increases at a decreasing rate. A larger demand forecast x_t implies a higher demand, which results in a higher expected profit. However, due to the ordering quantity limit, the retailer cannot order more to achieve a higher profit as x_t increases, slowing down the increase rate of $\pi^B(t, x_t)$. Hence, the ordering quantity restrictions may decrease the benefits of demand forecast updating. Because $\pi^B(t, x_t)$ is not directly proportional to x_t , the ordering policy cannot be the same as that in supply mode A.

Based on Lemma 3.3, the dynamic equations are analysed via backward induction at time

t , i.e.,

$$G_{T^B}(x_{T^B}) = \pi^B(T^B, x_{T^B}), \quad (3.29)$$

$$G_t(x_t) = \max\{\pi^B(t, x_t), E[G_{t+1}(x_t, \varepsilon_{t+1})]\}. \quad (3.30)$$

For time $T^B - 1$,

$$G_{T^B}(x_{T^B}) = \pi^B(T^B, x_{T^B}), \quad (3.31)$$

$$G_{T^B-1}(x_{T^B-1}) = \max\{\pi^B(T^B - 1, x_{T^B-1}), E[G_{T^B}(x_{T^B-1}, \varepsilon_{T^B})]\}. \quad (3.32)$$

In Lemma 3.3(b), $\pi^B(T^B, x_{T^B})$ and $\pi^B(T^B - 1, x_{T^B-1})$ are concave up in x_{T^B} and x_{T^B-1} , respectively. In Equation (3.31), $E[G_{T^B}(x_{T^B-1}, \varepsilon_{T^B})]$ is also concave up in x_{T^B-1} . As a result, in Equation (3.32), the ordering decision at time $T^B - 1$ is dependent on the number of cross points of the two concave functions $\pi^B(T^B - 1, x_{T^B-1})$ and $E[G_{T^B}(x_{T^B-1}, \varepsilon_{T^B})]$ in relation to x_{T^B-1} . If there is no cross point, the decision whether to order at time $T^B - 1$ is independent of x_{T^B-1} ; otherwise, it depends on x_{T^B-1} . However, it is difficult to determine the number of cross points of the two concave functions. In addition, the concavity of $G_{T^B-1}(x_{T^B-1})$ in Equation (3.32) cannot be determined, as it is the maximum of the two concave functions. Using induction, it is difficult to determine whether $E[G_t(x_t, \varepsilon_{t+1})]$ in Equation (3.30) is concave in x_t before time $T^B - 1$. This observation is concluded in Property 3.1.

Property 3.1 *At any time $t \in \{0, \dots, T^B - 2\}$, although $\pi^B(t, x_t)$ in Equation (3.30) is concave in x_t , the concavity of $E[G_t(x_t, \varepsilon_{t+1})]$ in x_t cannot be determined. Hence, the optimal ordering time is difficult to obtain by solving the dynamic equations, and the decision*

related to optimal ordering time may depend on x_t .

Comparing supply modes A and B, the ordering quantity restriction makes determining the optimal ordering time much more complicated despite its prevalence in practice. Therefore, the restriction must be relaxed properly. Lemma 4 shows how the optimal newsvendor-type quantity in Equation (3.28) changes over time during $[t, T_t]$ given x_t .

Lemma 3.4 *Under the M-MMFE forecast evolution, if the current demand forecast x_t is observed, the optimal newsvendor-type quantities q_t^* at time t and $q_{t'}^*$ at time $t' \in \{t + 1, \dots, T_t\}$ have the following properties:*

- (a) $P(q_{t'}^* \leq q_t^*) > 0.5$; and
- (b) $P(q_{t'}^* \leq q_t^*) > P_{t'} = \Phi \left(\Phi^{-1} \left(\frac{r - ct}{r - s} \right) \sqrt{\frac{\sigma_t - \sigma_{t'}}{\sigma_t + \sigma_{t'}}} \right)$ and the lower bound $P_{t'}$ is increasing in time t' .

Hence, $q_{t'}^* \leq q_t^*$ holds with high probability for time $t' \in \{t + 1, \dots, T_t\}$.

Lemma 3.4 indicates that the optimal newsvendor-type quantity $q_{t'}^*$ in the future is smaller than q_t^* at the current time with high probability ($P > 0.5$). Furthermore, as time passes, the lower bound of the probability $P_{t'}$ becomes larger. By using Lemma 3.4 and the decreasing trend of k_t with t , the ordering quantity restriction can be relaxed appropriately as follows.

If the retailer has not ordered at time t , given x_t , the optimal newsvendor-type ordering quantity q_t^* can be determined by Equation (3.28). If q_t^* outweighs k_t , the retailer should place an order of quantity k_t immediately due to the decreasing maximum ordering quantity in Theorem 3.1(b). Otherwise, q_t^* and the maximum order quantities during $[t, T^B]$ are compared. As a result, the retailer can determine the latest time T_t that the ordering quantity q_t^* can be completely satisfied by the manufacturer, i.e.,

$$q_t^* \leq k_{t'}, \text{ for } t \leq t' \leq T_t, \text{ and } q_t^* > k_{T_t+1}.$$

Due to Lemma 3.4, it is assumed that

$$q_{t'}^* \leq q_t^*, \text{ for } t \leq t' \leq T_t.$$

Hence,

$$q_{t'}^* \leq k_{t'}, \text{ for } t \leq t' \leq T_t.$$

These inequalities mean that the optimal newsvendor-type quantity can be completely satisfied during $[t, T_t]$. Hence, the ordering quantity restrictions during $[t, T_t]$ can be relaxed, and the decision whether to wait at time t in supply mode B is the same as that in supply mode A with the ordering timespan $[t, T_t]$. According to Theorem 3.2, the retailer should order at time t if $F(t)$ is larger than $F(T_t)$; otherwise, the retailer should wait.

Based on the preceding analysis, an algorithm is proposed to determine the optimal ordering decisions in supply mode B as follows.

Step 1. Initialization: set $t = 0$.

Step 2. Provide the values of k_t and T^B according to Theorem 3.1(b).

Step 3. Update x_t and calculate q_t^* according to Equation (3.28). If $q_t^* \geq k_t$, go to step

4 (Termination 1); otherwise, determine T_t by comparing q_t^* and $k_t, k_{t+1}, \dots, k_{T^B}$.

Furthermore, determine whether to order or wait at time t . If $F(t)$ is larger than $F(T_t)$, go to step 4 (Termination 1); otherwise, the retailer should wait and make the ordering decision at time $t + 1$.

If $1 \leq t + 1 \leq T^B - 1$, repeat step 3; otherwise, go to step 4 (Termination 2).

Step 4. Termination 1: place the order of the quantity $\min\{q_t^*, k_t\}$ at time t .

Termination 2: place the order of the quantity $\min\{q_{T^B}^*, k_{T^B}\}$ at time T^B .

3.3 Numerical Results

For innovative products such as Smart phones and Tablets, manufactures may charge a higher premium or impose restrictions on urgent orders. Retailers therefore face a trade-off between a lower ordering cost (with a longer production leadtime) and more accurate demand information (with a shorter production leadtime). Inspired by the dilemma, numerical tests are conducted to investigate the role of demand forecast updating in the purchasing plans in supply modes A and B. How the value of demand forecast updating changes with the product and demand parameters $(c_0, r, s, c_L, \sigma_f)$ is explored. In addition, to obtain a deeper understanding of demand forecast updating, the expected profits are compared under different supply scenarios and ordering policies.

Because the basic property of the supply lead time reduction is the same for each M/M/1 workstation according to Equation (3.10), whether N workstations are identical does not change the basic property of the total supply lead time reduction and the results. To simplify

the numerical analysis, N workstations are assumed to be identical. Let the total utilization ρ_i and the average service time $E(B_i)$ of each workstation be 0.8 and 0.027, respectively, and the number of workstations N be 50. According to Equation (3.9), the original lead time L is 7. Hence, the timeline starts at time 0 and terminates at time 7 (see Figure 3.1).

The initial demand forecast x_0 is fixed at 7. In supply mode A, it is assumed that the manufacturer sets the largest possible ordering quantity q_e at two times x_0 . In Equation (3.12), the latest ordering time T^A is 3. For the sake of comparison, the cases where q_e is the same as or three times x_0 are examined in Section 3.3.2. In supply mode B, according to Theorem 3.1(b), the values of the latest possible ordering time T^B and the maximum order quantities k_t are obtained as follows: $T^B = 4$, $k_0 = 29.63$, $k_1 = 27.95$, $k_2 = 25.27$, $k_3 = 20.37$ and $k_4 = 8.47$. Let the residual uncertainty factor $\alpha = L - T^A$ (or $L - T^B$). Its value is 4 (or 3) in supply mode A (or B). All of these parameter values remain constant unless otherwise specified. Note that $c_t > s$ and $\frac{r-c_t}{r-s} > 0.5$ are assumed in the numerical studies.

3.3.1 Sensitivity Analysis of the Value of Demand Forecast Updating

The effects of the product and demand parameters on demand forecast updating in the two supply modes are examined in this subsection. The value of demand forecast updating is quantified by comparing the cases with demand forecast updating with the cases without it. Without demand forecast updating, an order is always placed at time 0, referred to as the original policy. The value of demand forecast updating VD^i is defined as the percentage

improvement of the profit, i.e., $VD^i = (\pi^i(t^{i*}, x_{t^{i*}}) - \pi^i(0, x_0)) / \pi^i(0, x_0) \times 100\%$, where i denotes supply mode A or B. Table 3.1 presents the product and demand parameters (c_0, r, s, c_L, σ_f) and the corresponding VD^i . There are 324 different settings in each supply mode. Under each setting, VD^A is calculated based on the results seen in Section 3.2.3, and VD^B is obtained by running the simulation 10 million times based on the algorithm seen in Section 3.2.4. Table 3.2 summarizes the directional effects of these parameters on the value of demand forecast updating VD^i , from which Property 3.2 follows.

Property 3.2 (*Property of Demand Forecast Updating*)

- (1) VD^i is increasing in the original purchasing cost per unit c_0 in supply modes A and B.
- (2) VD^i is decreasing in the salvage value per unit s in the two supply modes.
- (3) VD^A is increasing, decreasing, or increasing and then decreasing in sales price r , and VD^B is decreasing in r in most cases.
- (4) VD^A is increasing in forecast efficiency σ_f , and VD^B is increasing in σ_f in most cases.
- (5) VD^i is decreasing in additional cost per unit c_L in the two supply modes.

Table 3.1 Value of Demand Forecast Updating in Supply Modes A and B

r	c_0	s	Supply Mode A									Supply Mode B									
			c_L	0.05	0.05	0.05	0.1	0.1	0.1	0.15	0.15	0.15	0.1	0.05	0.05	0.1	0.1	0.1	0.15	0.15	0.15
			σ_f	0.3	0.65	1	0.3	0.65	1	0.3	0.65	1	0.3	0.65	1	0.3	0.65	1	0.3	0.65	1
4	0.5	0.2	0	19.56	90.41	0	1.08	46.68	0	0	15.28	0	8.80	49.05	0	*0	24.92	0	0	4.22	
4	0.5	0.3	0	11.02	64.68	0	0	23.63	0	0	0	0	0.40	23.80	0	0	3.72	0	0	0	
4	0.5	0.4	0	0.00	29.79	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	0.8	0.2	2.77	37.97	149.42	0	18.14	98.39	0	1.36	58.93	2.88	30.97	122.80	0	14.67	85.34	0	*0	53.61	
4	0.8	0.3	1.42	32.74	132.58	0	12.95	82.79	0	0	45.09	1.35	24.58	100.77	0	8.88	66.74	0	0	37.51	
4	0.8	0.4	0	26.76	113.70	0	7.03	65.43	0	0	29.86	0	17.32	77.46	0	2.71	47.17	0	0	20.77	
4	1.1	0.2	6.03	52.05	197.34	0	30.46	139.08	0	11.79	92.45	6.84	51.17	198.19	0	30.87	145.60	0	12.15	100.58	
4	1.1	0.3	4.98	47.99	183.67	0	26.48	126.42	0	8.01	81.15	5.57	45.42	177.08	0	25.70	127.83	0	7.67	85.48	
4	1.1	0.4	3.84	43.58	169.00	0	22.14	112.86	0	3.91	69.12	4.19	39.32	155.03	0	20.35	109.47	0	2.98	69.88	
8	0.5	0.2	0	17.04	76.19	0	5.89	48.75	0	0	27.79	0	0	17.40	0	0	8.38	0	0	*0	
8	0.5	0.3	0	10.77	56.88	0	0	30.01	0	0	10.37	0	0	0	0	0	0	0	0	0	
8	0.5	0.4	0	1.92	29.99	0	0	4.83	0	0	0	0	0	0	0	0	0	0	0	0	
8	0.8	0.2	3.10	30.43	118.47	0	19.51	89.93	0	9.91	66.54	2.59	14.56	60.48	0	8.44	47.63	0	2.46	35.58	
8	0.8	0.3	2.17	26.64	106.44	0	15.56	77.94	0	5.95	55.02	1.61	10.28	47.34	0	4.50	35.53	0	*0	24.62	
8	0.8	0.4	1.13	22.32	92.85	0	11.06	64.46	0	1.44	42.16	0.60	6	33.57	0	0	22.95	0	0	12.86	
8	1.1	0.2	5.48	40.38	150.82	0.86	29.47	121.34	0	19.64	96.28	5.26	27.03	100.72	0.70	19.77	83.95	0	12.75	69.25	
8	1.1	0.3	4.75	37.47	141.34	0.09	26.45	111.87	0	16.61	87.10	4.42	23.33	88.66	*0	16.38	73.21	0	9.50	58.48	
8	1.1	0.4	3.97	34.31	131.14	0	23.18	101.70	0	13.32	77.25	3.54	19.44	76.12	0.00	12.73	61.50	0	6.17	48.06	
12	0.5	0.2	0.32	15.14	67.97	0	6.70	46.70	0	0	30.08	0	0	0	0	0	0	0	0	0	
12	0.5	0.3	0	9.93	51.63	0	1.22	30.42	0	0	14.49	0	0	0	0	0	0	0	0	0	
12	0.5	0.4	0	2.56	28.61	0	0	8.15	0	0	0	0	0	0	0	0	0	0	0	0	
12	0.8	0.2	2.78	26.20	102.97	0	18.22	81.85	0	11.11	64.18	1.99	7	35.42	0	3	28.49	0	0	21.72	
12	0.8	0.3	2.05	23.06	93.01	0	14.91	71.71	0	7.73	54.19	1.26	0	25.71	0	0	19.22	0	0	13.08	
12	0.8	0.4	1.24	19.48	81.72	0	11.13	60.24	0	3.88	42.97	0.51	0	15	0	0	10	0	0	4	
12	1.1	0.2	4.69	34.31	129.05	1.53	26.56	108.01	0	19.49	89.78	4.03	16.39	64.15	1.13	12.31	55.32	0	8.37	47.34	
12	1.1	0.3	4.11	31.91	121.31	0.91	24.04	100.13	0	16.92	81.96	3.38	13.52	55.47	0.51	9.62	47.25	0	5.73	39.25	
12	1.1	0.4	3.48	29.30	112.97	0.24	21.31	91.63	0	14.14	73.54	2.72	10.47	46.26	*0	6.76	38.51	0	3.02	31.13	
16	0.5	0.2	0.43	13.79	62.46	0	6.83	44.61	0	0.85	30.46	0	0	0	0	0	0	0	0	0	
16	0.5	0.3	0	9.22	47.93	0	1.98	29.91	0	0	16.19	0	0	0	0	0	0	0	0	0	
16	0.5	0.4	0	2.77	27.32	0	0	9.62	0	0	0	0	0	0	0	0	0	0	0	0	
16	0.8	0.2	2.51	23.45	93.20	0	16.99	75.93	0	11.19	61.31	1.54	0	21.43	0	0	16.89	0	0	12.45	
16	0.8	0.3	1.89	20.70	84.46	0	14.08	66.93	0	8.20	52.34	0.96	0	0	0	0	0	0	0	0	
16	0.8	0.4	1.21	17.57	74.52	0	10.75	56.72	0	4.79	42.22	0.34	0	0	0	0	0	0	0	0	
16	1.1	0.2	4.12	30.48	115.75	1.68	24.31	98.89	0	18.64	84.13	3.21	10.15	44.36	1.12	7.52	38.90	0	4.80	33.42	
16	1.1	0.3	3.62	28.38	109.02	1.15	22.10	91.97	0	16.37	77.18	2.68	8	37.28	0.60	5	32.12	0	3	27.04	
16	1.1	0.4	3.10	26.11	101.75	0.59	19.70	84.49	0	13.91	69.69	2.14	0	30.03	0.08	0	24.90	0	0	20.17	

* VD^B is presented as 0 when it is negative. Among the 324 settings, only 6 have negative values, and all of these negative values are no less than -4%. Under these six settings, the retailer cannot benefit from demand forecast updating and even suffers a small loss in profit. Hence, the order should be placed at time 0, and the negative value of VD^B is replaced by 0.

Property 3.2-(1) states that demand forecast updating is more valuable when the original purchasing cost c_0 is higher. When c_0 is high, the additional cost per unit c_L , i.e., the cost

of updating demand information, becomes less important. In other words, the retailer can take advantage of improved demand forecast accuracy from late orders at relatively small expense if c_0 is high. Therefore, the benefit of demand forecast updating is substantial if the product has a high original purchasing cost.

Property 3.2-(2) indicates that the retailer benefits more from demand forecast updating when the salvage value is lower. A lower salvage value leads to a higher overage cost, which makes demand forecasting more important. Therefore, if overstock products have a lower salvage value, the retailer should try to obtain as much demand information as possible to improve the accuracy of the demand forecast.

Property 3.2-(3) says that a higher price r may improve, deteriorate, or improve and then deteriorate the value of demand forecast updating VD^A due to the joint effects of price changes. First, as r increases, c_L becomes smaller, and it is thus beneficial for a retailer to update the demand forecast. Second, if r is high, implying that products are profitable, a retailer tends to purchase more products to avoid stock-out. As a result, demand information becomes less important. In addition, the second effect plays a more important role as r increases. Therefore, the ultimate change in VD^A with r depends on the combination of these two effects: if r is very small (or large), the first (or second) effect dominates, resulting in an increasing (decreasing) trend; if r is moderate, the second effect increases from weak to large enough to counteract the first effect, leading to an initial rise and a subsequent decline in VD^A .

There is another negative effect in supply mode B. As r increases, the ordering quantity

becomes larger, resulting in tighter ordering quantity restrictions. Due to the tighter restrictions, it is more difficult for a retailer to order the corresponding quantity based on the improved demand forecast. In addition, the tighter restrictions impede the retailer's overstock for the higher profit when r is larger, thus counteracting the second effect. Due to the three effects, VD^B may increase, decrease, or increase and then decrease with r , as shown in Table 3.1 and Table 3.2. Furthermore, Table 3.1 reveals that VD^B decreases with r in 74 out of 81 cases. Therefore, due to the quantity restrictions, the value of demand forecast updating becomes smaller as r increases in most cases in supply mode B.

Table 3.2 Summary of Parameter Sensitivities

	c_0	s	r	σ_f	c_L
VD^A	↑	↓	↑↓	↑	↓
VD^B	↑	↓	↑↓	↓↑	↓

Property 3.2-(4) states that an increase in forecast efficiency σ_f increases the value of the demand forecast updating in supply mode A. A retailer benefits more from demand forecast updating with higher forecast efficiency. Hence, VD^A increases with σ_f . However, VD^B does not always increase with σ_f , as shown in Table 3.1 and Table 3.2. Higher forecast efficiency also means that demand fluctuates more, which makes the ordering quantity restrictions tighter and impedes the use of demand information. Due to the quantity restrictions, VD^B may decrease with σ_f . For example, the decreasing trend is more

observable when r is higher (12 or 16) in Table 3.1, as the quantity restrictions become tighter as r increases. Furthermore, the VD^B of 5 out of 91 cases slightly decreases with σ_f , and the VD^B of the other 86 cases increases with σ_f . Therefore, the major effect of forecast efficiency on the value of demand forecast updating in supply mode B is positive.

Property 3.2-(5) indicates that the updated demand forecast is more valuable for a retailer in paying a smaller additional cost c_L . An increase in c_L increases the purchasing cost in Equation (3.8) and thus decreases the expected profit from the demand forecast updating. However, without the demand forecast updating, the profit is not affected by c_L because the order is always placed at time 0 without additional costs being paid. According to the definition of VD^i , it decreases with c_L . Therefore, it is beneficial for the retailer to update the demand forecast at a low additional cost. Note that manufacturers set a high additional cost c_L to adjust the priority during peak seasons. During peak seasons, demand forecast updating is less valuable. Hence, the probability of optimizing the ordering decision is smaller during peak seasons, which gives rise to stochastic double booking in the following property.

Property 3.3 (*Stochastic Double Booking*)

During peak seasons, due to the higher additional cost per unit c_L , a retailer is inclined to order more than it would during off-peak seasons.

Property 3.3 states that during peak seasons, an increase in c_L increases the ordering

quantity, resulting in the stochastic double booking effect. A higher c_L decreases the value of demand forecast updating, resulting in earlier orders. When retailers order earlier, they encounter higher demand uncertainty according to Equation (3.8). This leads to higher ordering quantities based on Lemma 3.4. Hence, in contrast with the conventional understanding of double booking according to the bullwhip effect (H. L. Lee, Padmanabhan, & Whang, 1997), retailers may still book more than the real demand due to the higher additional costs during peak seasons, even if the supply chain is synchronized.

3.3.2 Comparative Results

The benefits of demand forecast updating are further examined by comparing expected profits under different supply scenarios and ordering policies. Each expected profit is obtained by simulating 10 million sample paths of the demand forecast processes. The following parameter values are used: $s = 0.2$, $c_0 = 0.8$, $r = 4$ and $c_L = 0.05$. Table 3.1 shows that under this setting, VD^i is always larger than zero for different σ_f in the two supply modes. VD^i may be zero in other settings, and it would not be necessary to study demand forecast updating any further in those settings.

3.3.2.1 Comparison among Different Supply Scenarios

Several supply scenarios are studied to investigate the relationships between demand forecast updating and supply constraints. They are detailed in Table 3.3. Scenario 0 is a baseline scenario in which an order is placed at time 0 without demand forecast updating. Scenario F is the other extreme, in which an order is always placed at time 4 (i.e., the latest possible time)

with full demand forecast updating and without additional costs or order quantity restrictions. Scenario A0 corresponds to the extreme case of supply mode A, where T^A equals 4 by setting $q_e = x_0$. Scenario A refers to supply mode A with $T^A = 3$, and Scenario B denotes supply mode B, where k_t is as given previously. According to Equation (3.3), the demand uncertainty effect can be quantified by varying the forecast efficiency σ_f . Table 3.4 depicts the expected profits as σ_f changes under the five scenarios. As expected, the profits decrease with the demand uncertainty in every scenario.

Table 3.3 Different Supply Scenarios

Scenario	Demand Forecast Updating	Supply Restrictions			Ordering Policy	
		Additional Cost	Latest Ordering Time	Maximum Ordering Quantity	Time	Quantity
0	no	-	-	-	Always at time 0	Lemma 1
F	yes	no	4	no	Always at time 4	Lemma 1
A0	yes	c_t	4	no	Theorem 1	Lemma 1
A	yes	c_t	3	no	Theorem 1	Lemma 1
B	yes	c_t	4	k_t	Proposed Algorithm	

The expected profits for different scenarios in Table 3.3 are compared. The entire value of updated demand information V_f is defined as the percentage of profit improvement from Scenario 0 to Scenario F, i.e., $V_f = (\pi^F - \pi^0)/\pi^0 \times 100\%$. The loss of demand information value L_c due to the increased purchasing cost is measured by the profit difference between Scenario F and Scenario A1 over the profit improvement from Scenario 0 to Scenario F, i.e., $L_c = (\pi^F - \pi^{A0})/(\pi^F - \pi^0) \times 100\%$. In a similar manner, the losses of demand information value due to the ordering time and quantity restrictions are calculated as $L_t = (\pi^{A0} -$

$\pi^A)/(\pi^F - \pi^0) \times 100\%$ and $L_q = (\pi^{A0} - \pi^B)/(\pi^F - \pi^0) \times 100\%$. The values of V_f , L_c ,

L_t and L_q are summarized in Table 3.5, leading to Property 3.4.

Table 3.4 Expected Profits among Different Supply Scenarios and Ordering Policies

σ_f	Scenari 0 (Traditional Policy)	Scenario F	Scenario A0	Scenario A	Scenario B (Proposed Algorithm)	Scenario A1	Static Policy	Theoretical Policy
0.1	20.48	21.19	20.48	20.48	20.48	20.48	20.48	20.76
0.2	18.14	19.79	18.14	18.14	18.14	18.14	18.14	19.32
0.3	15.50	18.23	16.27	15.93	15.94	15.70	15.83	17.74
0.4	12.71	16.54	14.48	13.84	13.76	13.35	13.51	15.87
0.5	9.97	14.75	12.65	11.73	11.49	11.00	11.14	13.77
0.6	7.44	12.91	10.83	9.67	9.28	8.76	8.89	11.61
0.7	5.27	11.08	9.08	7.75	7.28	6.72	6.89	9.52
0.8	3.53	9.28	7.42	6.03	5.55	4.97	5.18	7.59
0.9	2.24	7.70	5.99	4.55	4.12	3.54	3.79	5.91
1	1.34	6.22	4.71	3.33	2.98	2.41	2.68	4.49

Table 3.5 Profit Comparisons among Different Supply Scenarios and Ordering Policies

σ_f	V_f (%)	L_c (%)	L_t (%)	L_q (%)	D_{AB} (%)	D_{AIB} (%)	Profit Drop (%)	Contribution (%)
0.1	3.49	100.00	0.00	0.00	0.00	0.00	1.37	0.00
0.2	9.09	100.00	0.00	0.00	0.00	0.00	6.10	0.00
0.3	17.60	71.65	12.65	12.06	-0.10	-1.54	10.12	19.83
0.4	30.09	53.75	16.84	18.80	0.55	-2.99	13.27	33.27
0.5	48.05	43.90	19.19	24.31	2.13	-4.22	16.58	40.00
0.6	73.65	38.03	21.11	28.29	4.23	-5.63	20.05	44.21
0.7	110.38	34.43	22.81	30.96	6.50	-7.62	23.52	47.34
0.8	163.20	32.34	24.20	32.59	8.71	-10.46	26.92	49.70
0.9	243.69	31.25	26.42	34.28	10.41	-14.17	30.25	51.28
1	364.68	30.89	28.21	35.38	11.73	-19.25	33.62	52.13

Property 3.4 *The value of demand information V_f increases with the demand uncertainty.*

As the demand uncertainty increases, the loss of demand information value L_c decreases due to the increased purchasing cost, and L_t and L_q increase due to the ordering time and quantity restrictions.

Property 3.4 states that an increase in demand uncertainty increases the value of demand information V_f . Demand information updating further improves the profit when there is a larger variation in demand. Furthermore, as the demand uncertainty increases, the increased purchasing cost plays a less critical role in decreasing the value of demand information, and the ordering time and quantity restrictions become more important. When the demand uncertainty is low, having a low purchasing cost is more important than improving demand accuracy, leading to a considerable reduction in the value of demand information. When the demand uncertainty is higher, less demand information can be acquired due to the ordering time restriction, resulting in a higher loss of demand information value. Furthermore, higher demand uncertainty implies the additional fluctuation of the demand forecast. The ordering quantity restrictions become tighter, resulting in the loss of additional demand information value. Therefore, the increased purchasing cost prevents a retailer from taking advantage of demand forecast updating, especially when the demand uncertainty is low. As the demand uncertainty increases, the ordering time and quantity restrictions prevent the retailer from using demand forecast updating. Hence, it is important for retailers to negotiate with manufacturers to relax restrictions when facing high demand uncertainty.

Supply modes A and B are also compared to determine which mode retailers prefer. In terms of supply mode A, a more conservative (or risk-averse) manufacturer may set $T^A = 2$ when $q_e = 3x_0$, rather than $T^A = 3$. This scenario is denoted as Scenario A1. The percentage of profit difference D_{AB} (or D_{A1B}) between Scenarios A (or A1) and B is shown

in Table 3.5. When $\sigma_f > 0.3$, D_{AB} remains positive and D_{A1B} remains negative. However, both absolute values increase as the demand uncertainty increases. This leads to the following property.

Property 3.5 *When the demand uncertainty is high, a retailer prefers Scenario A to Scenario B to Scenario A1.*

Property 3.5 shows that a retailer's supply mode preference depends on the manufacturer's risk tolerance. A retailer may prefer supply mode B (e.g., Scenario B over A1) if the manufacturer is risk-averse. A risk-averse manufacturer would provide a shorter ordering timespan to decrease the risk of satisfying an order using the additional capacity at a higher cost. However, a retailer may prefer supply mode A (e.g., Scenario A over B) because it provides greater flexibility than supply mode B. This preference becomes stronger as the demand variance becomes larger, suggesting that the supply mode selection is important when the retailer faces highly uncertain demand.

3.3.2.2 Comparison between Proposed Algorithm and Other Ordering Policies

Note that the optimal ordering time is approximated by the proposed algorithm in supply mode B. The theoretical optimal policy and static policy are compared to quantify the effectiveness of the algorithm. The theoretical optimal policy is a deterministic policy that declares whether a retailer should order by comparing the current profit with the average of

future profits, based on a given sample path of the demand forecast process. The static policy determines the optimal ordering time statically at time 0 by comparing the expected profits at the feasible ordering times.

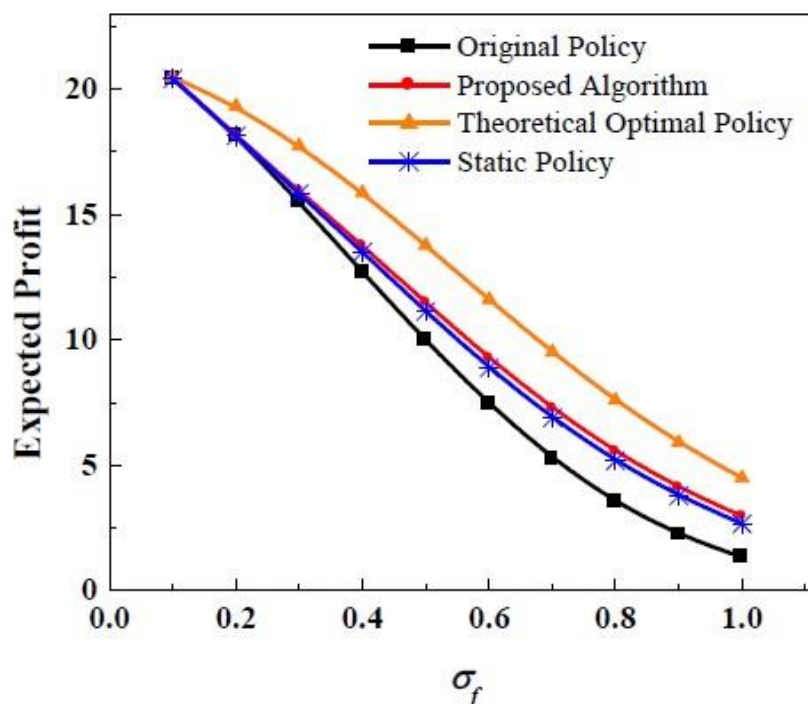


Figure 3.2 Expected Profit versus Demand Uncertainty under Four Policies in Supply Mode

B

Table 3.5 and Figure 3.2 demonstrate the expected profits from the proposed algorithm and theoretical optimal, static and original policies. Table 3.5 shows that as the demand uncertainty increases, the percentage of profit drop from the theoretical optimal policy to the proposed algorithm increases. It is more difficult to determine the optimal ordering time when the demand is highly uncertain. In contrast, compared with the original policy, the ratio

of profit improvement using the proposed algorithm to that using the theoretical optimal policy (i.e., contribution) increases as the demand uncertainty increases and ultimately exceeds 50%, as shown in Table 3.5. Moreover, Figure 3.2 shows that the expected profit from the proposed algorithm outweighs that from the static policy when the value of demand information is positive ($VD^B \neq 0$). Therefore, the proposed algorithm is effective for determining the optimal ordering time in supply mode B.

3.4 Conclusion

This study investigates an extension of the conventional newsvendor model by considering demand forecast updating and supply restrictions. The demand forecast becomes more accurate as time passes, which is described by the M-MMFE. Hence, retailers prefer to postpone orders to improve demand accuracy. However, this postponement is associated with supplier pricing and is therefore limited. In a manufacturing system consisting of a series of M/M/1 queues, the cost charged to the retailer increases with time, and the ordering time and quantity are restricted. These supply features negate the benefits of demand forecast updating. Therefore, the optimal ordering decisions are investigated from the viewpoint of a retailer seeking to strike a balance between demand forecast accuracy and supply constraints in supply modes A and B.

In supply mode A, the optimal ordering time is independent of the demand forecast evolution, and the optimal ordering quantity is dependent on the current demand forecast and profitable factor. Based on justifiable assumptions, the optimal ordering time is one of two

endpoints of the feasible timespan. Hence, demand forecast updating is valuable if the retailer has low demand accuracy tolerance. Otherwise, it is completely valueless. In supply mode B, due to the ordering quantity restrictions, the optimal ordering time may depend on the demand forecast evolution and becomes difficult to obtain. Thus, an algorithm is proposed to approximate the optimal ordering time by relaxing the ordering quantity restrictions appropriately.

The effects of product and demand attributes on the value of demand forecast updating in supply modes A and B are shown through numerical experiments. Demand forecast updating is found to be more beneficial to retailers in the two supply modes, particularly when (a) the original purchasing cost is more expensive, (b) overstock products have smaller value, (c) the additional cost for lead time reduction is lower and (d) the demand forecast efficiency is higher. Furthermore, although in most cases the value of demand forecast updating increases when the price is lower in supply mode B, the relationship is inconstant in supply mode A.

The expected profits are compared under different supply scenarios. The benefit of demand forecast updating decreases largely due to the increased purchasing cost when the demand uncertainty is low. In a market with highly uncertain demand, a retailer should try to relax the ordering time and quantity restrictions to take advantage of the demand forecast updating. This suggests that a proactive retailer should try to find multiple suppliers and choose one with a shorter supply lead time when necessary. Moreover, the supply mode comparison indicates that it is more beneficial to retailers in supply mode B than supply mode A if the manufacturer is risk-averse, especially in a market with highly uncertain

demand, and vice versa.

3.5 Appendix

Proof of Theorem 3.1. (a). T^A is obtained from Equation (3.10). (b). Note that $E(B_i) = 1/\mu_i$, $\rho_i = \lambda/\mu_i$ and $\mu_i > \lambda = q_{total} > k_t$. Thus, Equation (3.11) can be simplified to Equation (3.13), from which it is easy to determine that k_t decreases with t . If k_t declines to zero or a negative value, there is no ordering opportunity at time t . Hence, $T^B = \min\{t|k_t \leq 0\} - 1$. (c). This shows that $T^A = \left\lceil \sum_{i=1}^N \rho_i E(B_i) \left(\frac{1}{1-\rho_i} - \frac{1}{1-qE(B_i)} \right) \right\rceil \leq \left\lceil \sum_{i=1}^N \rho_i E(B_i) \left(\frac{1}{1-\rho_i} - 1 \right) \right\rceil$. According to Equation (3.11), $T^B = \min\{t|k_t \leq 0\} - 1 = \left\lceil \sum_{i=1}^N \rho_i E(B_i) \left(\frac{1}{1-\rho_i} - 1 \right) \right\rceil - 1 = \left\lfloor \sum_{i=1}^N \rho_i E(B_i) \left(\frac{1}{1-\rho_i} - 1 \right) \right\rfloor$. As a result, $T^A \leq T^B$. The theorem then follows. Q.E.D.

Proof of Lemma 3.1. (a) Let $TP(q) = (c_t - s)E[(q - X_D)^+ | x_t] + (r - c_t)E[(X_D - q)^+ | x_t]$. As this is a newsvendor-type expression, it is easy to determine that $TP(q)$ is convex in q , and $\frac{dTP(q)}{dq} = (r - s)Pr(X_D < q | x_t) - (r - c_t)$. Note that $X_D | x_t$ is lognormally distributed with the parameters μ_t and σ_t in Equations (3.2) and (3.3). Solving $\frac{dTP(q)}{dq} = 0$, $q_t^{A*} = e^{\sigma_t \Phi^{-1}\left(\frac{r-c_t}{r-s}\right) + \ln x_t - 0.5\sigma_t^2} = x_t e^{\sigma_t \Phi^{-1}\left(\frac{r-c_t}{r-s}\right) - 0.5\sigma_t^2} > 0$. From Equation (3.15), it is determined that q_t^{A*} is the optimal ordering quantity. Substituting q_t^{A*} and Equation (3.4) into Equation (3.15), the corresponding expected profit is

$$\begin{aligned} \pi^A(t, x_t) &= (r - c_t)E[X_D | x_t] - [(r - s)E[X_D | x_t]\Phi(\sigma_t - z_{\beta_t}) - (c_t - s)E[X_D | x_t]] \\ &= (r - s)x_t[1 - \Phi(\sigma_t - z_{\beta_t})] \end{aligned}$$

$$= (r - s)x_t \bar{\Phi}(\sigma_t - z_{\beta_t}).$$

(b) For the last decision epoch $T^A - 1$, using Equations (3.20), (3.17) and (3.18),

$$\begin{aligned} G_{T^A-1}(x_{T^A-1}) &= \max\{\pi^A(T^A - 1, x_{T^A-1}), E[G_{T^A}(x_{T^A-1} \cdot \varepsilon_{T^A})]\} \\ &= \max\{\pi^A(T^A - 1, x_{T^A-1}), E[\pi^A(T^A, x_{T^A-1} \cdot \varepsilon_{T^A})]\} \\ &= \max\{(r - s)x_{T^A-1} \bar{\Phi}(\sigma_{T^A-1} - z_{\beta_{T^A-1}}), E[\varepsilon_{T^A}](r - s)x_{T^A-1} \bar{\Phi}(\sigma_{T^A} \\ &\quad - z_{\beta_{T^A}})\} \\ &= \max\{(r - s)x_{T^A-1} \bar{\Phi}(\sigma_{T^A-1} - z_{\beta_{T^A-1}}), (r - s)x_{T^A-1} \bar{\Phi}(\sigma_{T^A} - z_{\beta_{T^A}})\} \\ &= (r - s)x_{T^A-1} \max\{\bar{\Phi}(\sigma_{T^A-1} - z_{\beta_{T^A-1}}), \bar{\Phi}(\sigma_{T^A} - z_{\beta_{T^A}})\}, \end{aligned} \quad (3.33)$$

where the fourth equality holds because $E[\varepsilon_{T^A}] = 1$ under the Martingale process. Hence,

$G_{T^A-1}(x_{T^A-1})$ at time $T^A - 1$ is directly proportional to the demand forecast x_{T^A-1} . By

induction, it must be proved that the profit-to-go function can be written as follows:

$$G_t(x_t) = (r - s)x_t \max\{\bar{\Phi}(\sigma_t - z_{\beta_t}), \bar{\Phi}(\sigma_{t+1} - z_{\beta_{t+1}}), \dots, \bar{\Phi}(\sigma_T - z_{\beta_T})\}. \quad (3.34)$$

Equations (3.20) and (3.33) show that this is true for times T^A and $T^A - 1$. Assuming that it

is true for time t , it can be proved that it is also true at time $t - 1$. $G_{t-1}(x_{t-1})$ can be

derived as follows:

$$\begin{aligned} G_{t-1}(x_{t-1}) &= \max\{\pi^A(t - 1, x_{t-1}), E[G_t(X_t)]\} \\ &= \max\{\pi^A(t - 1, x_{t-1}), E[G_t(x_{t-1} \cdot \varepsilon_t)]\} \\ &= \max\{(r - s)x_{t-1} \bar{\Phi}(\sigma_{t-1} - z_{\beta_{t-1}}), E[\varepsilon_t](r - s)x_{t-1} \\ &\quad \max\{\bar{\Phi}(\sigma_t - z_{\beta_t}), \bar{\Phi}(\sigma_{t+1} - z_{\beta_{t+1}}), \dots, \bar{\Phi}(\sigma_T - z_{\beta_T})\}\} \\ &= (r - s)x_{t-1} \max\{\bar{\Phi}(\sigma_{t-1} - z_{\beta_{t-1}}), \max\{\bar{\Phi}(\sigma_t - z_{\beta_t}), \end{aligned}$$

$$\begin{aligned} & \bar{\Phi}(\sigma_{t+1} - z_{\beta_{t+1}}), \dots, \bar{\Phi}(\sigma_T - z_{\beta_T}) \} \\ & = (r - s)x_{t-1} \max\{\bar{\Phi}(\sigma_{t-1} - z_{\beta_{t-1}}), \bar{\Phi}(\sigma_t - z_{\beta_t}), \dots, \bar{\Phi}(\sigma_T - z_{\beta_T})\}, \end{aligned}$$

where the third equality is established because $E[\varepsilon_t] = 1$ under the Martingale forecast process. Hence, Equation (3.34) is also true at time $t - 1$. The desired result follows directly.

Based on Equation (3.34), the decision whether to order at any time t depends not on x_t but on the comparative results of $\bar{\Phi}(\sigma_t - z_{\beta_t})$ during $[t, T^A]$. As a result, the results of Equation (3.21) can be easily seen, as $\Phi(\cdot)$ is an increasing function. This completes the proof. Q.E.D.

Proof of Lemma 3.2. According to Equation (3.23), taking the first and second derivatives of σ_t with t , $\frac{d\sigma(t)}{dt} = -\frac{\sigma_f}{2} \frac{1}{\sqrt{T-t+a}} < 0$ and $\frac{d^2\sigma(t)}{d^2t} = -\frac{\sigma_f}{4} (T-t+a)^{-\frac{3}{2}} < 0$. According to Equation (3.24), taking the first and second derivatives of $z_{\beta}(t)$ in relation to t yields $\frac{dz_{\beta}(t)}{dt} = -\frac{c_L}{r-s} [\phi(\Phi^{-1}(\frac{r-c_0}{r-s} - \frac{c_L}{r-s}t))]^{-1} < 0$, and $\frac{d^2z_{\beta}(t)}{d^2t} = \Phi^{-1}(\frac{r-c_t}{r-s}) (\frac{c_L}{r-s})^2 [\phi(\Phi^{-1}(\frac{r-c_t}{r-s}))]^{-2} = z_{\beta}(t) (\frac{c_L}{r-s})^2 [\phi(\Phi^{-1}(\frac{r-c_t}{r-s}))]^{-2} > 0$, for $z_{\beta}(t) > 0$. The statement in the lemma follows. Q.E.D.

Proof of Theorem 3.2. According to Lemma 3.2, $F(t) = z_{\beta}(t) - \sigma(t)$ is convex. Theorem 3.2 follows from Theorem 3.1 and Lemma 3.2. Q.E.D.

Proof of Theorem 3.3. (a). According to Equation (3.14), for any fixed $t \in \{0, \dots, T^A\}$, the expected profit $\pi^A(t, x_t)$ is non-decreasing in s . According to Equations (3.20), (3.23) and

(3.24),

$$\begin{aligned}\pi^A(t, x_t) &= (r - s)x_t \overline{\Phi(\sigma_t - z_{\beta_t})} \\ &= (r - s)x_t \left[1 - \Phi \left(\sigma_f \sqrt{T^A - t + \alpha} - \Phi^{-1} \left(\frac{r - c_0}{r - s} - \frac{c_L}{r - s} t \right) \right) \right].\end{aligned}$$

Thus, for any fixed $t \in \{0, \dots, T^A\}$, $\pi^A(t, x_t)$ is non-increasing in c_0 , c_L , σ_f and α . Note that the optimal expected profit $\pi^A(t^{A*}, x_{t^{A*}})$ is the maximum of the profits at the feasible times during $[0, T^A]$. Thus, $\pi^A(t^{A*}, x_{t^{A*}})$ is non-decreasing in s but non-increasing in c_0 , c_L , σ_f and α . (b). According to Equation (3.25), $T_v = \Phi^{-1} \left(\frac{r - c_0}{r - s} \right) - \sigma_f \sqrt{T^A + \alpha} - \Phi^{-1} \left(\frac{r - c_0}{r - s} - \frac{c_L}{r - s} T^A \right) + \sigma_f \sqrt{\alpha}$. Taking the first derivatives of T_v in relation to σ_f , c_L , α c_0 yields the following:

$$\begin{aligned}\frac{\partial T_v}{\partial \sigma_f} &= \sqrt{\alpha} - \sqrt{T^A + \alpha} < 0; \\ \frac{\partial T_v}{\partial c_L} &= \frac{T^A}{r - s} \left[\phi \left(\Phi^{-1} \left(\frac{r - c_0}{r - s} - \frac{c_L}{r - s} T^A \right) \right) \right]^{-1} > 0; \\ \frac{\partial T_v}{\partial \alpha} &= 0.5 \sigma_f \left(\frac{1}{\sqrt{\alpha}} - \frac{1}{\sqrt{T^A + \alpha}} \right) > 0; \\ \frac{\partial T_v}{\partial c_0} &= \frac{1}{r - s} \left[\left(\phi \left(\Phi^{-1} \left(\frac{r - c_0}{r - s} - \frac{c_L}{r - s} T^A \right) \right) \right)^{-1} - \left(\phi \left(\Phi^{-1} \left(\frac{r - c_0}{r - s} \right) \right) \right)^{-1} \right] < 0.\end{aligned}$$

The last inequality holds because $z_{\beta}(t) = \Phi^{-1} \left(\frac{r - c_0}{r - s} - \frac{c_L}{r - s} t \right) > 0$ for $t \in \{0, \dots, T^A\}$. The statement follows directly. Q.E.D.

Proof of Lemma 3.3. (a). From the proof of Lemma 3.1, it is determined that $TP(q)$ is convex in q and that $TP(q)$ is the lowest when $q^* = e^{\sigma_t \Phi^{-1} \left(\frac{r - c_t}{r - s} \right) + \ln x_t - 0.5 \sigma_t^2} > 0$. Thus, for $0 \leq q \leq k_t$, the ordering quantity is equal to $\min\{q^*, k_t\}$ to minimize $TP(q)$. From Equation (3.26), it is determined that $q_t^{B*} = \min\{e^{\sigma_t \Phi^{-1} \left(\frac{r - c_t}{r - s} \right) + \ln x_t - 0.5 \sigma_t^2}, k_t\}$. (b). Substituting q_t^{B*} into Equation (3.26) and taking the first and second derivatives of

$\pi^B(t, x_t)$ with x_t yield the following:

$$\frac{\partial \pi^B(t, x_t)}{\partial x_t} = \begin{cases} (r-s) \left[1 - \Phi \left(\sigma_t - \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \right) \right] > 0, & \ln x_t \leq \ln k_t + \frac{\sigma_t^2}{2} - \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \sigma_t \\ (r-s) \left[1 - \Phi \left(\sigma_t - \left(\frac{\ln k_t - \ln x_t + \frac{\sigma_t^2}{2}}{\sigma_t} \right) \right) \right] > 0, & \ln x_t > \ln k_t + \frac{\sigma_t^2}{2} - \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \sigma_t \end{cases},$$

$$\frac{\partial^2 \pi^B(t, x_t)}{\partial^2 x_t} = \begin{cases} 0, & \ln x_t \leq \ln k_t + \frac{\sigma_t^2}{2} - \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \sigma_t \\ -\frac{(r-s)}{\sigma_t x_t} \phi \left(\sigma_t - \left(\frac{\ln k_t - \ln x_t + \frac{\sigma_t^2}{2}}{\sigma_t} \right) \right) < 0, & \ln x_t > \ln k_t + \frac{\sigma_t^2}{2} - \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \sigma_t \end{cases}.$$

Therefore, $\pi^B(t, x_t)$ is concave up in relation to x_t . The desired statement is thus made.

Q.E.D.

Proof of Lemma 3.4. Based on the demand forecast process of the M-MMFE, if the demand forecast is x_t at current time t , then at time $t' > t$,

$$\begin{aligned} \ln x_{t'} &= \ln x_t + \ln \varepsilon_{t+1} + \dots + \ln \varepsilon_{t'} \\ &= \ln x_t + e_{t,t'} - (0.5\sigma_t^2 - 0.5\sigma_{t'}^2), \end{aligned}$$

where σ_t and $\sigma_{t'}$ follow Equation (3.3) and $e_{t,t'} = \ln \varepsilon_{t+1} + \dots + \ln \varepsilon_{t'} + 0.5\sigma_t^2 - 0.5\sigma_{t'}^2$.

Note that $e_{t,t'}$ is normally distributed with the parameters $(0, \sqrt{\sigma_t^2 - \sigma_{t'}^2})$ because

$\ln \varepsilon_{t+1} \sim N(-0.5\sigma_f^2, \sigma_f)$ for $t \in \{0, \dots, T^B - 1\}$. The optimal news vendor-type quantity $q_{t'}^*$

in (3.28) can then be derived as follows:

$$q_{t'}^* = e^{\sigma_{t'} \Phi^{-1} \left(\frac{r-c_{t'}}{r-s} \right) + \ln x_{t'} - 0.5\sigma_{t'}^2}$$

$$= \exp(\sigma_{t'} \Phi^{-1} \left(\frac{r-c_{t'}}{r-s} \right) + \ln x_t + e_{t,t'} - 0.5\sigma_{t'}^2). \quad (3.35)$$

Hence, from Equations (3.28) and (3.35),

$$\begin{aligned} P(q_{t'}^* \leq q_t^*) &= P(\exp(\sigma_{t'} \Phi^{-1} \left(\frac{r-c_{t'}}{r-s} \right) + \ln x_t + e_{t,t'} - 0.5\sigma_{t'}^2) \leq \exp(\sigma_t \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \\ &\quad + \ln x_t - 0.5\sigma_t^2)) \\ &= P(\sigma_{t'} \Phi^{-1} \left(\frac{r-c_{t'}}{r-s} \right) + e_{t,t'} \leq \sigma_t \Phi^{-1} \left(\frac{r-c_t}{r-s} \right)) \\ &= P(e_{t,t'} \leq \sigma_t \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) - \sigma_{t'} \Phi^{-1} \left(\frac{r-c_{t'}}{r-s} \right)) \\ &= \Phi \left((\sigma_t \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) - \sigma_{t'} \Phi^{-1} \left(\frac{r-c_{t'}}{r-s} \right)) / \sqrt{\sigma_t^2 - \sigma_{t'}^2} \right) \\ &> \Phi \left((\sigma_t \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) - \sigma_{t'} \Phi^{-1} \left(\frac{r-c_t}{r-s} \right)) / \sqrt{\sigma_t^2 - \sigma_{t'}^2} \right) \\ &= \Phi \left(\Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \sqrt{\sigma_t - \sigma_{t'}} / \sqrt{\sigma_t + \sigma_{t'}} \right) = P_{t'} \\ &> 0.5. \end{aligned}$$

The fourth equality holds because $e_{t,t'}$ is normally distributed with the parameters $(0, \sqrt{\sigma_t^2 - \sigma_{t'}^2})$, and the first inequality follows from $\Phi^{-1} \left(\frac{r-c_{t'}}{r-s} \right) < \Phi^{-1} \left(\frac{r-c_t}{r-s} \right)$, where $t' > t$, because c_t increases with time according to Equation (3.8). Because $\sigma_{t'}$ decreases with time t' (see Equation (3.3)), $\Phi(\cdot)$ is an increasing function, and $z_\beta(t) = \Phi^{-1} \left(\frac{r-c_t}{r-s} \right) > 0$, $P_{t'} = \Phi \left(\Phi^{-1} \left(\frac{r-c_t}{r-s} \right) \sqrt{\sigma_t - \sigma_{t'}} / \sqrt{\sigma_t + \sigma_{t'}} \right)$ increases with time t' during $[t, T_t]$ and is larger than 0.5. The statement in this lemma follows. Q.E.D.

Chapter 4. On Optimal Emergency Orders with Updated Demand Forecast and Limited Supply

This chapter presents a two-stage newsvendor system with a regular and an emergency order. The emergency order can be placed at a later time based on a more accurate demand forecast. However, the unit cost for the emergency order is higher, and the quantity is limited. To maximize the expected profit, a retailer should determine both regular and emergency order quantities by considering the demand forecast updating, ordering cost and quantity constraint. Using dynamic programming, optimal ordering quantities are derived, and properties of the optimal solutions are obtained. Numerical experiments are carried out to illustrate the effect of the emergency order on the ordering decisions and expected profit. Some managerial insights are gained from the numerical results.

4.1 Introduction

This study analyzes a retailer's ordering problem in a two-stage newsvendor setting. The retailer can place a regular order and a fast but expensive emergency order with updated demand forecast and limited ordering quantity. Examples of such a model can be found in a variety of practical situations. For example, the delayed product differentiation (e.g., the HP and IBM cases in H. L. Lee and Tang (1997)) allows manufactures to postpone some production processes. Retailers can order later at a higher cost, but no more than the amount of the common part that has been processed. For another example, in China, to deal with the

changing demands for products in the automobile and PC industries, retailers would place an urgent order by paying a higher wholesale price than the regular order (Jianli Li & Liu, 2008). However, the urgent ordering quantity cannot exceed the manufacturer's reserved capacity. In addition, North American and European companies outsource products from lower-cost countries (e.g., Zara in Ghemawat, et al. (2003)). Although the local supply usually takes a short lead time, the available amount is limited in contrast to the international supply.

In this two-stage newsvendor model, optimal ordering quantities are derived by using dynamic programming. How the regular ordering quantity and impact of the maximum emergency ordering quantity change under different prices, salvage values, unit costs, magnitude of the maximum limits and demand characteristics is also investigated. Based on the investigation, an algorithm is proposed to determine the optimal regular ordering quantity. Numerical experiments are conducted to support and further extend the analytical results. It suggests that the retailer should take advantage of the emergency order especially for products with (a) high initial demand forecast under lognormal distributions, (b) high demand variability before the emergency order under normal distributions, or (c) low demand variability after the emergency order under both normal and lognormal distributions.

The rest of this chapter is organized as follows. Section 4.2 presents models and characterizes optimal solutions. Section 4.3 provides numerical results to support the analysis and obtain more insights. Section 4.4 gives concluding remarks and directions for future research.

4.2 Model

This study considers an ordering problem in a two-stage setting, where a retailer has both regular and emergency ordering opportunities and receives orders in one shipment prior to the selling season as shown in Figure 4.1. At Stage 1, the retailer places a regular order q_1 from the supplier at a unit cost c_1 . After observing the updated demand information at Stage 2, the retailer can pay a higher unit cost $c_2 > c_1$ for an emergency order q_2 under a maximum limit M (i.e., $q_2 \leq M$). During the selling season, the retailer obtains revenue $r > c_1$ for each sold unit and salvage value $s < c_1$ for each unsold unit.

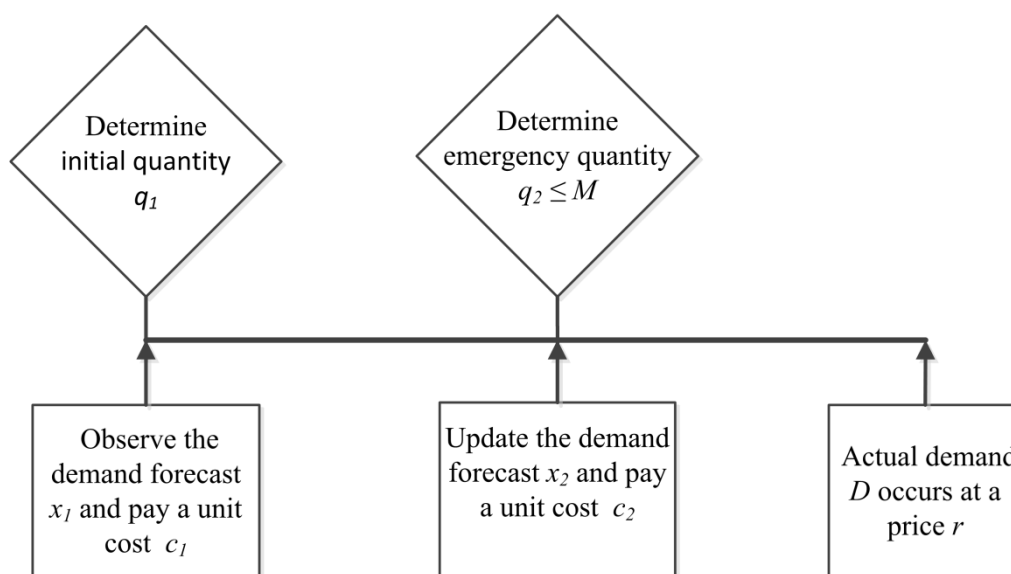


Figure 4.1 The Two-Stage Model Structure

The retailer determines the first and second ordering quantities based on demand information at the two stages. Let X_1 and X_2 denote the demand forecasts made at Stages 1

and 2 for the actual demand X during the selling season. Given the demand forecast x_1 at Stage 1, X_2 is assumed to have a distribution function $F_1(\cdot|x_1)$ and density function $f_1(\cdot|x_1)$. Given the updated demand forecast x_2 at Stage 2, the distribution and density function of X are $F_2(\cdot|x_2)$ and $f_2(\cdot|x_2)$. Demand forecasts are assumed to be non-negative. In addition, it is assumed that the inverse distribution function $F_2^{-1}((r - c_2)/(r - s)|x_2)$ is differentiable, strictly increases and asymptotically diverges to infinity when $x_2 \rightarrow \infty$.

This study solves the retailer's ordering problem using dynamic programming. The optimal ordering quantity q_2^* at Stage 2 is first determined for a given q_1 , and then the optimal ordering quantity q_1^* at Stage 1 is determined. In addition, the effects of the demand characteristics on the optimal solutions are investigated.

4.2.1 Ordering Quantity at Stage 2

Given the current demand forecast x_2 and first-stage ordering quantity q_1 , the expected profit at Stage 2 is

$$w_2(q_1, q_2, x_2) = -c_2q_2 + rE_{F_2(\cdot|x_2)}((q_1 + q_2) \wedge X) + sE_{F_2(\cdot|x_2)}((q_1 + q_2 - X)^+). \quad (4.1)$$

Then, the optimal expected profit at Stage 2 is given by

$$v_2(q_1, x_2) = \max\{w_2(q_1, q_2, x_2), q_2 \in [0, M]\}.$$

Let the total ordering quantity $q = q_1 + q_2$, and

$$SP(q, x_2) = -c_2q + rE_{F_2(\cdot|x_2)}(q \wedge X) + sE_{F_2(\cdot|x_2)}((q - X)^+). \quad (4.2)$$

Note that $SP(q, x_2)$ is a standard newsvendor form with the conditional distribution of X at Stage 2. Equation (4.1) can be rewritten as

$$w_2(q_1, q_2, x_2) = c_2 q_1 + SP(q, x_2). \quad (4.3)$$

Lemma 4.1. *The profit function $w_2(q_1, q_2, x_2)$ at Stage 2 is concave in q_2 .*

Proof.

$$\begin{aligned} \frac{\partial w_2(q_1, q_2, x_2)}{\partial q_2} &= \frac{\partial SP(q, x_2)}{\partial q} \\ &= -\int_0^q (c_2 - s) f_2(x|x_2) dx + \int_q^\infty (r - c_2) f_2(x|x_2) dx \end{aligned} \quad (4.4)$$

$$= (r - c_2) - (r - s)F_2(q|x_2). \quad (4.5)$$

$$\frac{\partial^2 w_2(q_1, q_2, x_2)}{\partial^2 q_2} = \frac{\partial^2 SP(q, x_2)}{\partial^2 q} = -(r - s)f_2(q|x_2) < 0. \quad (4.6)$$

Hence, $w_2(q_1, q_2, x_2)$ is a concave function.

Q.E.D.

Let $q_2^o(q_1, x_2) = F_2^{-1}((r - c_2)/(r - s)|x_2) - q_1$. Given Lemma 4.1, the optimal ordering quantity under a quantity constraint M at Stage 2 can be obtained in the following theorem.

Theorem 4.1. *If the retailer orders no more than M at Stage 2, given q_1 and x_2 , the optimal ordering quantity q_2^* at Stage 2 is given by*

$$q_2^* = \begin{cases} 0 & x_2 < x_L(q_1) \\ q_2^o(q_1, x_2) & x_L(q_1) \leq x_2 \leq x_M(q_1), \\ M & x_2 > x_M(q_1) \end{cases} \quad (4.7)$$

where $x_L(q_1)$ and $x_M(q_1)$ satisfy:

$$F_2(q_1|x_L(q_1)) = \frac{r - c_2}{r - s}, \quad (4.8)$$

$$F_2(M + q_1|x_M(q_1)) = \frac{r - c_2}{r - s}. \quad (4.9)$$

Moreover, $x_L(q_1)$ and $x_M(q_1)$ are differentiable.

Proof. Since $w_2(q_1, q_2, x_2)$ is a concave function, the optimal ordering quantity q_2^* at Stage 2 is given by

$$q_2^* = \begin{cases} 0 & q_2^o < 0 \\ q_2^o(q_1, x_2) & 0 \leq q_2^o \leq M, \\ M & q_2^o > M \end{cases} \quad (4.10)$$

According to Equations (4.5) and (4.6), the optimal total ordering quantity $q^* = \max(F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2\right), q_1)$. Note that $F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2\right)$ strictly increases with x_2 and $x_2 \geq 0$. Hence,

$$q^* \geq F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2 = 0\right). \quad (4.11)$$

In this situation, a rational retailer would order at least $F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2 = 0\right)$ at Stage 1 due to the lower ordering cost. It means that $q_1 \geq q_{min}$, where $q_{min} = F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2 = 0\right)$. In addition, it is assumed that $\lim_{x_2 \rightarrow \infty} F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2\right)$ diverges to infinity. According to the intermediate value theorem, there exists one and only one value of $x_L(q_1)$ (respectively, $x_M(q_1)$) satisfying $F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_L\right) = q_1$ ($F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_M\right) = q_1 + M$). As a result, Equation (4.10) can be transformed into Equation (4.7). Furthermore, because $F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2\right)$ is a strictly increasing and differentiable function with respect to x_2 , $x_L(q_1)$ and $x_M(q_1)$ are differentiable with respect to q_1 . This completes the proof. Q.E.D.

4.2.2 Ordering Quantity at Stage 1

According to Theorem 4.1, given the current demand forecast x_1 , the expected profit function at Stage 1 is given by

$$\begin{aligned}
 w_1(q_1, x_1) = & -c_1 q_1 + \int_0^{x_L(q_1)} w_2(q_1, 0, x_2) f_1(x_2|x_1) dx_2 \\
 & + \int_{x_L(q_1)}^{x_M(q_1)} w_2(q_1, q_2^o(q_1, x_2), x_2) f_1(x_2|x_1) dx_2 \\
 & + \int_{x_M(q_1)}^{\infty} w_2(q_1, M, x_2) f_1(x_2|x_1) dx_2.
 \end{aligned} \tag{4.12}$$

Then, the optimal expected profit can be written as

$$v_1(x_1) = \max\{w_1(q_1, x_1), q_1 \geq 0\}.$$

Lemma 4.2. *The profit function $w_1(q_1, x_1)$ at Stage 1 is concave in q_1 .*

Proof. From Theorem 4.1, $q_2^o(q_1, x_L(q_1)) = 0$, and $q_2^o(q_1, x_M(q_1)) = M$. According to Equation (4.12) and Leibniz's rule,

$$\begin{aligned}
 \frac{\partial w_1(q_1, x_1)}{\partial q_1} = & -c_1 + \int_0^{x_L(q_1)} \frac{\partial w_2(q_1, 0, x_2)}{\partial q_1} f_1(x_2|x_1) dx_2 \\
 & + \int_{x_L(q_1)}^{x_M(q_1)} \frac{\partial w_2(q_1, q_2^o(q_1, x_2), x_2)}{\partial q_1} f_1(x_2|x_1) dx_2 \\
 & + \int_{x_M(q_1)}^{\infty} \frac{\partial w_2(q_1, M, x_2)}{\partial q_1} f_1(x_2|x_1) dx_2.
 \end{aligned}$$

By Equation (4.3), $\partial w_2(q_1, q_2, x_2)/\partial q_1 = c_2 + \partial SP(q, x_2)/\partial q$. Thus, based on Equation

(4.5)

$$\begin{aligned} \frac{\partial w_1(q_1, x_1)}{\partial q_1} = c_2 - c_1 - \left\{ \int_0^{x_L(q_1)} [(r - s)F_2(q_1|x_2) - (r - c_2)] f_1(x_2|x_1) dx_2 \right. \\ \left. - \int_{x_M(q_1)}^{\infty} [(r - c_2) - (r - s)F_2(q_1 + M|x_2)] f_1(x_2|x_1) dx_2 \right\}. \end{aligned} \quad (4.13)$$

By Equations (4.8) and (4.9), $(r - c_2) - (r - s)F_2(q_1|x_L(q_1)) = 0$, and $(r - c_2) - (r - s)F_2(q_1 + M|x_M(q_1)) = 0$. Thus, according to Leibniz's rule,

$$\begin{aligned} \frac{\partial^2 w_1(q_1, x_1)}{\partial^2 q_1} = -(r - s) \left[\int_0^{x_L(q_1)} f_2(q_1|x_2) f_1(x_2|x_1) dx_2 \right. \\ \left. + \int_{x_M(q_1)}^{\infty} f_2(q_1 + M|x_2) f_1(x_2|x_1) dx_2 \right] < 0. \end{aligned}$$

Hence, $w_1(q_1, x_1)$ is concave in q_1 .

Q.E.D.

Theorem 4.2. Given x_1 , the optimal ordering quantity q_1^* at Stage 1 satisfies

$$\begin{aligned} \int_0^{x_L(q_1^*)} [(r - s)F_2(q_1^*|x_2) - (r - c_2)] f_1(x_2|x_1) dx_2 \\ - \int_{x_M(q_1^*)}^{\infty} [(r - c_2) - (r - s)F_2(q_1^* + M|x_2)] f_1(x_2|x_1) dx_2 \\ - (c_2 - c_1) = 0. \end{aligned} \quad (4.14)$$

Proof. As $w_1(q_1, x_1)$ is concave in q_1 , the existence of q_1^* is proved by showing that Equation (4.13) can equal zero. According to Equations (4.8) and (4.11), $x_L(q_{min}) = 0$.

Hence, according to Equation (4.13),

$$\begin{aligned} \frac{\partial w_1(q_1, x_1)}{\partial q_1} \Big|_{q_1=q_{min}} &= (c_2 - c_1) + \int_{x_M(q_{min})}^{\infty} [(r - c_2) - (r - s)F_2(M|x_2)]f_1(x_2|x_1)dx_2 \\ &> 0. \end{aligned}$$

In addition, since $x_L(q_1) \leq x_M(q_1)$ and $F_2(q_1 + M|x_2) \geq F_2(q_1|x_2)$,

$$\begin{aligned} \frac{\partial w_1(q_1, x_1)}{\partial q_1} &= c_2 - c_1 - \int_0^{x_L(q_1)} [(r - s)F_2(q_1|x_2) - (r - c_2)]f_1(x_2|x_1)dx_2 \\ &\quad + \int_{x_M(q_1)}^{\infty} [(r - c_2) - (r - s)F_2(q_1 + M|x_2)]f_1(x_2|x_1)dx_2 \\ &\leq c_2 - c_1 - \int_0^{x_L(q_1)} [(r - s)F_2(q_1|x_2) - (r - c_2)]f_1(x_2|x_1)dx_2 \\ &\quad + \int_{x_L(q_1)}^{\infty} [(r - c_2) - (r - s)F_2(q_1|x_2)]f_1(x_2|x_1)dx_2. \end{aligned} \quad (4.15)$$

Note that Equation (4.15) is the first derivative of $w_1(q_1, x_1)$ when $M = 0$. If $M = 0$, the ordering problem becomes a standard newsvendor problem, where the optimal ordering quantity q_{10}^* exists. Hence, when $q_1 = q_{10}^*$, Equation (4.15) is equal to zero, implying that

$$\frac{\partial w_1(q_1, x_1)}{\partial q_1} \Big|_{q_1=q_{10}^*} \leq 0.$$

Because $w_1(q_1, x_1)$ is concave in q_1 , $\partial w_1(q_1, x_1)/\partial q_1$ decreases with q_1 .

Consequently, there exists $q_{min} < q_1^* \leq q_{10}^*$ such that $\partial w_1(q_1, x_1)/\partial q_1|_{q_1=q_1^*} = 0$. Q.E.D.

Comparing Equation (4.14) with the first-order condition (e.g. Equation (4.4)) in the standard newsvendor model, the expected overage cost c_0 , expected underage cost c_u and postponement cost c_p can be quantified by

$$c_0 = \int_0^{x_L(q_1)} [(r - s)F_2(q_1|x_2) - (r - c_2)] f_1(x_2|x_1) dx_2, \quad (4.16)$$

$$c_u = \int_{x_M(q_1)}^{\infty} [(r - c_2) - (r - s)F_2(q_1 + M|x_2)] f_1(x_2|x_1) dx_2, \quad (4.17)$$

$$c_p = (c_2 - c_1).$$

In Equations (4.16) and (4.17), an increase in q_1 increases the expected overage cost c_0 but decreases the expected underage cost c_u . Intuitively, if the retailer orders too much at Stage 1, any of the initial order cannot be cancelled based on the updated information at Stage 2, resulting in the overage cost. On the other hand, if the retailer orders too little at Stage 1, he may not order enough at Stage 2 due to the quantity constraint M , leading to the underage cost. Therefore, the optimal quantity q_1^* is determined by balancing the overage cost, underage cost and postponement cost.

Corollary 4.1. *The optimal regular ordering quantity q_1^* decreases with c_1 and M , but it increases with c_2 , s and r .*

Proof. Let $J(q_1^*, M, c_1, c_2, s, r, x_1) = \partial w_1(q_1, x_1) / \partial q_1 |_{q_1=q_1^*}$. From Equation (4.13), the following derivatives are obtained:

$$\begin{aligned} \frac{\partial J}{\partial q_1^*} &= \frac{\partial^2 w_1(q_1^*, x_1)}{\partial^2 q_1^*} < 0; \\ \frac{\partial J}{\partial c_1} &= -1 < 0; \end{aligned} \quad (4.18)$$

$$\frac{\partial J}{\partial M} = -(r - s) \int_{x_M(q_1^*)}^{\infty} f_2(q_1^* + M|x_2) f_1(x_2|x_1) dx_2 < 0;$$

$$\frac{\partial J}{\partial c_2} = 1 - \int_0^{x_L(q_1^*)} f_1(x_2|x_1) dx_2 - \int_{x_M(q_1^*)}^{\infty} f_1(x_2|x_1) dx_2 = \int_{x_L(q_1^*)}^{x_M(q_1^*)} f_1(x_2|x_1) dx_2 > 0;$$

$$\frac{\partial J}{\partial s} = \int_0^{x_L(q_1^*)} F_2(q_1^*|x_2) f_1(x_2|x_1) dx_2 + \int_{x_M(q_1^*)}^{\infty} F_2(q_1^* + M|x_2) f_1(x_2|x_1) dx_2 > 0;$$

$$\begin{aligned} \frac{\partial J}{\partial r} &= \int_0^{x_L(q_1^*)} (1 - F_2(q_1^*|x_2)) f_1(x_2|x_1) dx_2 \\ &\quad + \int_{x_M(q_1^*)}^{\infty} (1 - F_2(q_1^* + M|x_2)) f_1(x_2|x_1) dx_2 > 0. \end{aligned}$$

According to the implicit function theorem, $\frac{\partial q_1^*}{\partial x} = -\frac{\partial J}{\partial x} / \frac{\partial J}{\partial q_1^*}$. Hence, $\frac{\partial q_1^*}{\partial c_1} < 0$, $\frac{\partial q_1^*}{\partial M} < 0$,

$\frac{\partial q_1^*}{\partial c_2} > 0$, $\frac{\partial q_1^*}{\partial s} > 0$ and $\frac{\partial q_1^*}{\partial r} > 0$. This completes the proof. Q.E.D.

The results of Corollary 4.1 are reasonable. If it is cheaper to order earlier but more expensive to order later, the retailer should order more products earlier. Besides, higher revenue and salvage value imply that the retailer can make more profit for the sold products and suffer less loss for the unsold ones, and thus he would like to increase the ordering quantity. Furthermore, due to the smaller available quantity M for the emergency order, the retailer will be forced to increase the initial ordering quantity even with high demand uncertainty at Stage 1.

Theorem 4.3 *The optimal expected profit $v_1(x_1)$ increases with M . The increasing rate of $v_1(x_1)$ with M is*

$$\frac{dv_1(x_1)}{dM} = \int_{x_M(q_1^*)}^{\infty} [(r - c_2) - (r - s)F_2(q_1^* + M|x_2)]f_1(x_2|x_1)dx_2 \quad (4.19)$$

$$= \int_0^{x_L(q_1^*)} [(r - s)F_2(q_1^*|x_2) - (r - c_2)]f_1(x_2|x_1)dx_2 - (c_2 - c_1). \quad (4.20)$$

Furthermore, the increasing rate decreases as M increases. For any fixed M , the increasing rate increases with c_1 .

Proof. Note that $\partial v_1(x_1)/\partial q_1^* = 0$ from Theorem 4.2. According to Equation (4.12),

$$\begin{aligned} \frac{dv_1(x_1)}{dM} &= \frac{\partial v_1(x_1)}{\partial M} + \frac{\partial v_1(x_1)}{\partial q_1^*} \frac{\partial q_1^*}{\partial M} \\ &= \frac{\partial v_1(x_1)}{\partial M} \\ &= (w_2(q_1^*, M, x_M(q_1^*)) - w_2(q_1^*, M, x_M(q_1^*)))f_1(x_M(q_1^*)|x_1) \frac{\partial x_M(q_1^*)}{\partial M} \\ &\quad + \int_{x_M(q_1^*)}^{\infty} \frac{\partial w_2(q_1^*, M, x_1)}{\partial M} f_1(x_2|x_1)dx_2 \\ &= \int_{x_M(q_1^*)}^{\infty} [(r - c_2) - (r - s)F_2(q_1^* + M|x_2)]f_1(x_2|x_1)dx_2 \geq 0, \end{aligned}$$

where the last equality follows from Equations (4.3) and (4.5). From Equation (4.14), Equation (4.20) is obtained.

By Corollary 4.1, q_1^* decreases with M and c_1 . Furthermore, in Equations (4.8) and (4.9), $x_L(q_1^*)$ and $x_M(q_1^*)$ increase with q_1^* . Thus, $x_L(q_1^*)$ decreases with M , and $x_M(q_1^*)$ decreases with c_1 . According to Equation (4.20), the increasing rate decreases with M . According to Equation (4.19), it increases with c_1 for any fixed M . Q.E.D.

Definition 4.1 *The increasing rate of the optimal expected profit $v_1(x_1)$ with M denotes the magnitude of the impact of M on the profit. Thus, the emergency order size impact factor I_f is defined as*

$$I_f = \frac{dv_1(x_1)}{dM}.$$

Given the emergency order size impact factor I_f , the retailer can figure out whether he should focus on supply sources to ensure high availability of the emergency order in advance. Theorem 4.3 shows that the increasing rate is equal to c_u in Equation (4.17) when $q_1 = q_1^*$. Hence, I_f can be measured by the balanced expected underage cost. Furthermore, I_f becomes larger when M is smaller and c_1 is higher. Note that the relative cost $(c_2 - c_1)$ paid for the emergency order becomes smaller as c_1 increases. Hence, if the emergency order is scarce and cheap, the retailer should be sensitive to its availability.

4.2.3 Study of Two Important Demand Distributions

A special case of the martingale model of forecast evolution (MMFE) (R. G. Brown, 1959; Hausman, 1969; Heath & Jackson, 1994) is used in this two-stage setting. The studies (T. Wang, et al., 2012; Y. Wang & Tomlin, 2009) with the MMFE model show that forecast becomes more accurate as it is close to the selling season. Two types of the MMFE are considered here: the normal (ND) and lognormal (LND) distributions. For ND, the successive demand forecasts differences, i.e., $(X_2 - X_1)$ and $(X - X_2)$, are assumed to be normally distributed with parameters $(0, \sigma_1)$ and $(0, \sigma_2)$. For LND, the successive forecasts ratios, i.e., X_2/X_1 and X/X_2 , are lognormally distributed with parameters $(-0.5\sigma_1^2, \sigma_1)$ and $(-0.5\sigma_2^2,$

σ_2). Therefore, the distribution functions are as follows:

$$F_1(x_2|x_1) = \begin{cases} \Phi\left(\frac{x_2-x_1}{\sigma_1}\right) & \text{(ND)} \\ \Phi\left(\frac{\ln x_2 - \ln x_1 + 0.5\sigma_1^2}{\sigma_1}\right) & \text{(LND)} \end{cases}, \quad (4.21)$$

$$F_2(x|x_2) = \begin{cases} \Phi\left(\frac{x-x_2}{\sigma_2}\right) & \text{(ND)} \\ \Phi\left(\frac{\ln x - \ln x_2 + 0.5\sigma_2^2}{\sigma_2}\right) & \text{(LND)} \end{cases}.$$

According to Equation (4.21), under both ND and LND, $F_2^{-1}\left(\frac{r-c_2}{r-s} \middle| x_2\right)$ strictly increases with x_2 . Thus, the preceding results can be used. Substituting Equation (4.21) into Theorem 4.1 yields the following parameters related to q_2^* :

$$q_2^o(q_1, x_2) = \begin{cases} x_2 + \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\sigma_2 - q_1 & \text{(ND)} \\ x_2 \exp\left(-0.5\sigma_2^2 + \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\sigma_2\right) - q_1 & \text{(LND)} \end{cases}, \quad (4.22)$$

$$x_L(q_1) = \begin{cases} q_1 - \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\sigma_2 & \text{(ND)} \\ q_1 \exp(0.5\sigma_2^2 - \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\sigma_2) & \text{(LND)} \end{cases}, \quad (4.23)$$

$$x_M(q_1) = \begin{cases} M + q_1 - \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\sigma_2 & \text{(ND)} \\ (M + q_1) \exp(0.5\sigma_2^2 - \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\sigma_2) & \text{(LND)} \end{cases}. \quad (4.24)$$

For q_1^* , Equation (4.14) can be written as follows:

$$(c_2 - c_1) - \int_0^{x_L(q_1^*)} \left[(r-s)\Phi\left(\frac{q_1^* - x_2}{\sigma_2}\right) - (r-c_2) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2$$

$$+ \int_{x_M(q_1^*)}^{\infty} \left[(r-c_2) - (r-s)\Phi\left(\frac{q_1^* + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 = 0 \quad \text{(ND);} \quad (4.25)$$

$$(c_2 - c_1) - \int_0^{x_L(q_1^*)} \left[(r-s)\Phi\left(\frac{\ln q_1^* - \ln x_2 + 0.5\sigma_2^2}{\sigma_2}\right) - (r-c_2) \right]$$

$$\phi\left(\frac{\ln x_2 - \ln x_1 + 0.5\sigma_1^2}{\sigma_1}\right) \frac{1}{\sigma_1 x_2} dx_2$$

$$\begin{aligned}
 & + \int_{x_M(q_1^*)}^{\infty} \left[(r - c_2) - (r - s) \Phi \left(\frac{\ln(q_1^* + M) - \ln x_2 + 0.5\sigma_2^2}{\sigma_2} \right) \right] \\
 & \phi \left(\frac{\ln x_2 - \ln x_1 + 0.5\sigma_1^2}{\sigma_1} \right) \frac{1}{\sigma_1 x_2} dx_2 = 0 \quad \text{(LND)}. \quad (4.26)
 \end{aligned}$$

Based on these equations, the following theorem shows the effects of demand characteristics on the optimal ordering quantities and corresponding profits.

Theorem 4.4

1. For *ND* and *LND*,

- (a) *the optimal regular ordering quantity q_1^* and expected profit $v_1(x_1)$ increase with x_1 ; and*
- (b) *the optimal emergency ordering quantity q_2^* and expected profit $v_2(q_1, x_2)$ increase with x_2 .*

2. In particular, for *ND*,

- (a) *q_1^* is positive linear with x_1 , and the emergency order size impact factor I_f is independent of x_1 ; and*
- (b) *For any positive σ_2 , $I_f < I_f^0$, where I_f^0 denotes the limit of I_f as σ_2 approaches zero.*

Proof. The proof is provided in the Appendix.

Theorem 4.4-1 shows that the retailer orders more and expects higher profit when the demand forecast is higher as higher demand forecast indicates potentially higher demand during the selling season. Theorem 4.4-2(a) indicates that the ordering quantity q_1^* increases

linearly with the demand forecast x_1 at Stage 1, but the impact of the maximum ordering quantity I_f at Stage 2 is not affected by x_1 under ND. Hence, in response to the increased initial demand forecast, the retailer only needs to increase the initial ordering quantity correspondingly without considering the emergency ordering quantity constraint under ND. Theorem 4.4-2(b) states the impact of the availability of the emergency order becomes the strongest when the demand uncertainty is almost resolved at Stage 2 under ND. In this situation, it is an optimal policy to ensure the availability of the emergency order.

4.3 Numerical Analysis

In this section, numerical tests are conducted to support and further extend the analysis in Section 4.2. According to the preceding statements, an algorithm is proposed to calculate the optimal regular ordering quantity q_1^* . Moreover, the influence of the availability of the emergency order on the performance of the two-stage model is investigated.

Note that it is difficult to determine q_1^* by solving Equation (4.14) directly. According to Corollary 4.1, when $M = 0$, q_1^* achieves the maximum value (denoted by q_{10}^*), which is a newsvendor-type solution. Let L denote the left hand side of Equation (4.14), i.e.,

$$L = \int_0^{x_L(q_1)} [(r - s)F_2(q_1|x_2) - (r - c_2)] f_1(x_2|x_1) dx_2 - \int_{x_M(q_1)}^{\infty} [(r - c_2) - (r - s)F_2(q_1 + M|x_2)] f_1(x_2|x_1) dx_2 - (c_2 - c_1). \quad (4.27)$$

According to Equation (4.13) and Lemma 4.2, as the ordering quantity q_1 decreases from q_{10}^* , L decreases. When L becomes zero, the value of q_1^* is determined. Hence, it is easy to

determine q_1^* at any value of M by running a linear search through Equation (4.14) from q_{10}^* . The absolute value of L is the error of the following algorithm for q_1^* .

Step 1. Initialization: set up all the parameters.

Step 2. Calculate q_{10}^* according to the traditional newsvendor model with the given demand distributions.

Step 3. Set $q_1 = q_{10}^*$ and calculate L according to Equation (4.27). Set the step size $s = -1$ and indicator $I = L$.

Step 4. Set $q_1 = q_1 + s$ and update L according to Equation (4.27). If $L * I < 0$, go to step 5; otherwise, repeat step 4.

Step 5. Set $s = -0.1s$ and $I = L$. If $|s| > 0.0001$, go to step 4; otherwise, $q_1^* = q_1$.

The demand distributions follow ND or LND. Based on the presentation in Section 4.2.3, $X|x_1$ is normally distributed with parameters $(x_1, \sqrt{\sigma_1^2 + \sigma_2^2})$ for ND, and it is lognormally distributed with parameters $(\ln x_1 - 0.5(\sigma_1^2 + \sigma_2^2), \sqrt{\sigma_1^2 + \sigma_2^2})$ for LND. Hence, if $M = 0$, the newsvendor-type regular ordering quantity is

$$q_{10}^* = \begin{cases} x_1 + \sqrt{\sigma_1^2 + \sigma_2^2} \Phi^{-1}\left(\frac{r - c_1}{r - s}\right) & \text{(ND)} \\ x_1 \exp(-0.5(\sigma_1^2 + \sigma_2^2)) + \sqrt{\sigma_1^2 + \sigma_2^2} \Phi^{-1}\left(\frac{r - c_1}{r - s}\right) & \text{(LND)} \end{cases}.$$

In the base case scenario, the demand parameters are set as $x_1 = 300$, $\sigma_1 = 30$ and $\sigma_2 = 6$ under ND, and $x_1 = 100$, $\sigma_1 = 1$ and $\sigma_2 = 0.2$ under LND. Other parameters in the base case are set as follows: $r = 3$, $c_1 = 1$, $c_2 = 2$ and $s = 0.2$. Based on the preceding analysis and algorithm, numerical experiments are conducted to study the influence of the emergency order by observing q_1^* and I_f . Note that the smallest step size to search q_1^* is

10^{-3} , that the error of the algorithm is at most order of 10^{-5} , and that the average time it takes for each scenario is around 1 second. Tables 4.1–4.7 and Figures 4.2–4.15 show the changes in q_1^* and I_f with M , r , c_1 , c_2 , s , x_1 , σ_1 and σ_2 . From these numerical results, the following observations and insights are obtained:

Table 4.1 The Impact of r on q_1^* and I_f for Different M under ND and LND

ND	$r = 2.5$		$r = 3$		$r = 3.5$		$r = 4.0$		$r = 4.5$	
	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	311.97		317.31		321.37		324.61		327.31	
5	311.05	10.56%	315.81	18.48%	319.45	24.43%	322.39	29.09%	324.84	32.85%
10	310.22	8.68%	314.43	15.45%	317.69	20.68%	320.33	24.86%	322.55	28.31%
15	309.49	6.98%	313.19	12.67%	316.08	17.19%	318.45	20.88%	320.45	23.96%
20	308.84	5.50%	312.09	10.18%	314.65	14.00%	316.76	17.18%	318.55	19.89%
25	308.30	4.23%	311.14	7.99%	313.39	11.15%	315.25	13.83%	316.84	16.15%
30	307.84	3.18%	310.33	6.12%	312.30	8.66%	313.94	10.87%	315.35	12.81%
35	307.48	2.33%	309.66	4.57%	311.38	6.55%	312.82	8.31%	314.06	9.89%
40	307.19	1.66%	309.12	3.32%	310.63	4.82%	311.90	6.18%	312.99	7.41%
45	306.97	1.15%	308.70	2.34%	310.04	3.44%	311.16	4.46%	312.11	5.39%
50	306.81	0.78%	308.38	1.60%	309.59	2.38%	310.58	3.11%	311.43	3.80%
LND	$r = 2.5$		$r = 3$		$r = 3.5$		$r = 4.0$		$r = 4.5$	
M	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	88.60		105.88		121.21		135.06		147.73	
10	86.82	11.26%	102.92	19.89%	117.39	26.42%	130.60	31.52%	142.77	35.61%
20	85.29	10.05%	100.26	18.19%	113.91	24.54%	126.46	29.60%	138.12	33.72%
30	83.96	8.98%	97.90	16.62%	110.73	22.76%	122.65	27.74%	133.78	31.86%
40	82.82	8.04%	95.80	15.18%	107.86	21.07%	119.14	25.95%	129.75	30.05%
50	81.83	7.21%	93.93	13.86%	105.26	19.49%	115.93	24.24%	126.02	28.28%
60	80.97	6.47%	92.27	12.66%	102.91	18.01%	112.99	22.61%	122.57	26.58%
70	80.22	5.83%	90.80	11.56%	100.80	16.64%	110.31	21.07%	119.40	24.95%
80	79.56	5.26%	89.49	10.58%	98.89	15.37%	107.87	19.62%	116.48	23.39%
90	78.99	4.76%	88.32	9.68%	97.17	14.20%	105.65	18.27%	113.81	21.90%
100	78.48	4.32%	87.29	8.87%	95.63	13.12%	103.64	16.99%	111.36	20.50%

(1) From Tables 4.1–4.7, it is observed that the optimal regular ordering quantity q_1^* decreases as M increases, which is in accordance with Corollary 4.1. Besides, the emergency order size impact factor I_f decreases with M , which is consistent with Corollary 4.1. Hence, an increase in M brings more marginal benefit if M is smaller, otherwise it

brings less benefit. This implies that an emergency opportunity is important to the retailer even if the available emergency quantity is not large. On the other hand, if M is large enough, increasing M would bring a small profit. There exists an “effectively unlimited” M , after which I_f becomes almost zero and q_1^* barely changes. For instance, for the case of $\sigma_1 = 0.2$ in Table 4.6, when $M > 60$, I_f is zero and q_1^* remains the same. In this case, the available emergency quantity of 60 is enough for the retailer.

(2) The emergency order size impact factor I_f is larger if the price r is higher in Table 4.1 and Figures 4.2–4.3. Higher price implies that products are more profitable, and thus unsatisfied demand leads to a larger loss. In this situation, demand information accuracy becomes more important. Because more accurate demand information is obtained when the emergency order is placed, the availability of the emergency order becomes more critical to the retailer as r increases.

(3) The emergency order size impact factor I_f increases with the regular ordering cost c_1 but decreases with the emergency ordering cost c_2 in Tables 4.2–4.3 and Figures 4.3–4.7. Intuitively, if the relative cost $(c_2 - c_1)$ for the emergency order is smaller, the retailer can benefit more and thus be affected more significantly by the availability of the emergency order.

(4) The emergency order size impact factor I_f becomes smaller as the salvage value s increases in Table 4.4 and Figures 4.8–4.9. The higher the salvage value, the less loss the unsold products incur. Due to the smaller overage cost, the retailer would order more to reduce the risk of stock-out, instead of taking the more accurate demand information. This

results in a decrease in the impact of the emergency order.

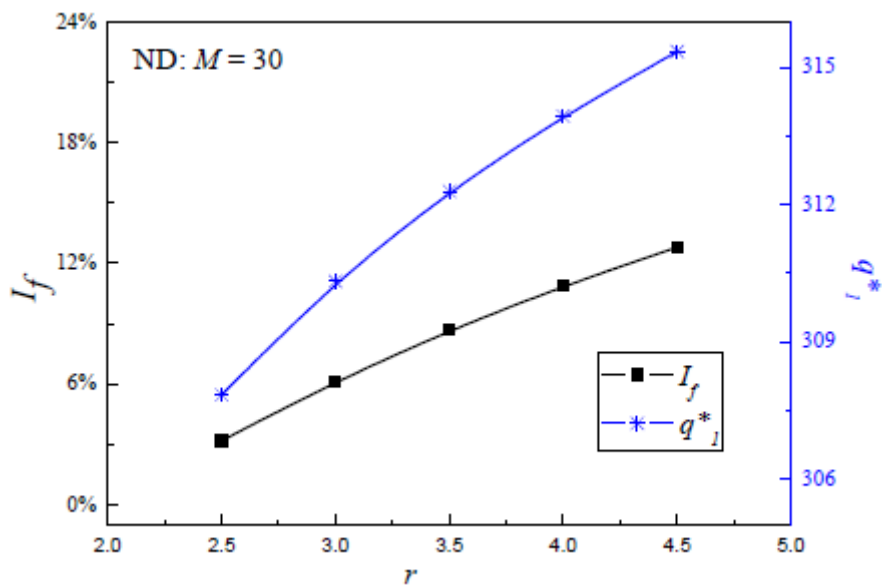


Figure 4.2 q_1^* and I_f versus r under ND

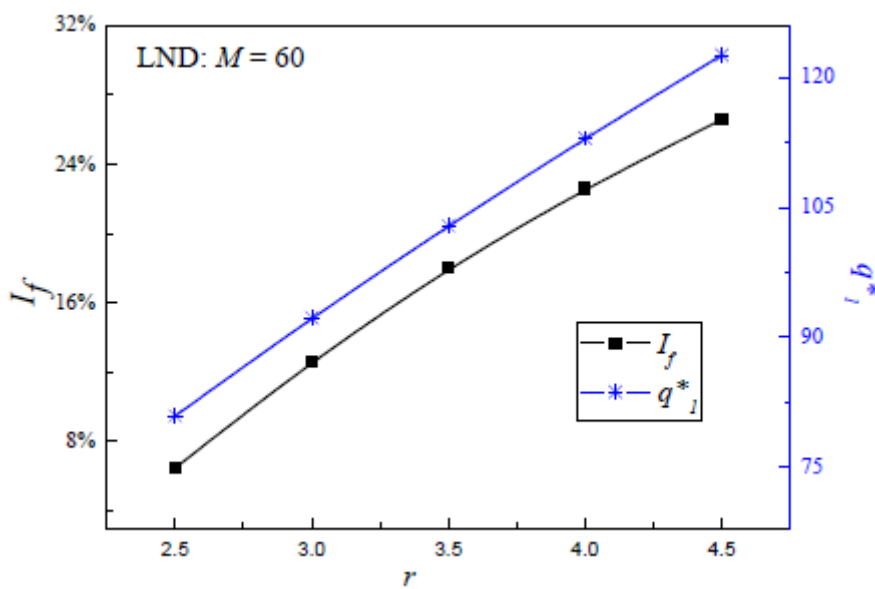


Figure 4.3 q_1^* and I_f versus r under LND

Table 4.2 The Impact of c_1 on q_1^* and I_f for Different M under ND and LND

ND	$c_1 = 0.6$		$c_1 = 0.8$		$c_1 = 1.0$		$c_1 = 1.2$		$c_1 = 1.4$	
	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	332.66		324.22		317.31		311.20		305.51	
5	331.37	7.88%	322.81	12.93%	315.81	18.48%	309.61	24.46%	303.83	30.85%
10	330.27	6.15%	321.56	10.49%	314.43	15.45%	308.11	20.96%	302.21	27.02%
15	329.36	4.67%	320.48	8.32%	313.19	12.67%	306.72	17.69%	300.67	23.36%
20	328.62	3.44%	319.56	6.44%	312.09	10.18%	305.44	14.67%	299.21	19.93%
25	328.04	2.46%	318.79	4.85%	311.14	7.99%	304.29	11.93%	297.85	16.73%
30	327.60	1.71%	318.18	3.56%	310.33	6.12%	303.26	9.51%	296.59	13.81%
35	327.28	1.15%	317.69	2.53%	309.66	4.57%	302.37	7.40%	295.45	11.18%
40	327.05	0.75%	317.33	1.75%	309.12	3.32%	301.62	5.62%	294.43	8.87%
45	326.90	0.48%	317.06	1.18%	308.70	2.34%	301.01	4.16%	293.55	6.87%
50	326.79	0.29%	316.87	0.77%	308.38	1.60%	300.51	2.99%	292.80	5.19%
LND	$c_1 = 0.6$		$c_1 = 0.8$		$c_1 = 1.0$		$c_1 = 1.2$		$c_1 = 1.4$	
M	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	176.60		133.28		105.88		86.36		71.43	
10	173.93	9.16%	130.46	14.40%	102.92	19.89%	83.27	25.61%	68.22	31.55%
20	171.50	8.52%	127.90	13.27%	100.26	18.19%	80.52	23.28%	65.38	28.52%
30	169.29	7.92%	125.61	12.21%	97.90	16.62%	78.09	21.13%	62.87	25.75%
40	167.27	7.36%	123.54	11.24%	95.80	15.18%	75.95	19.17%	60.68	23.23%
50	165.43	6.84%	121.69	10.34%	93.93	13.86%	74.06	17.39%	58.76	20.95%
60	163.76	6.35%	120.03	9.52%	92.27	12.66%	72.39	15.78%	57.08	18.90%
70	162.24	5.91%	118.54	8.77%	90.80	11.56%	70.93	14.32%	55.60	17.07%
80	160.86	5.49%	117.20	8.08%	89.49	10.58%	69.64	13.02%	54.32	15.43%
90	159.60	5.11%	115.99	7.44%	88.32	9.68%	68.50	11.85%	53.19	13.97%
100	158.46	4.75%	114.91	6.87%	87.29	8.87%	67.49	10.79%	52.19	12.67%

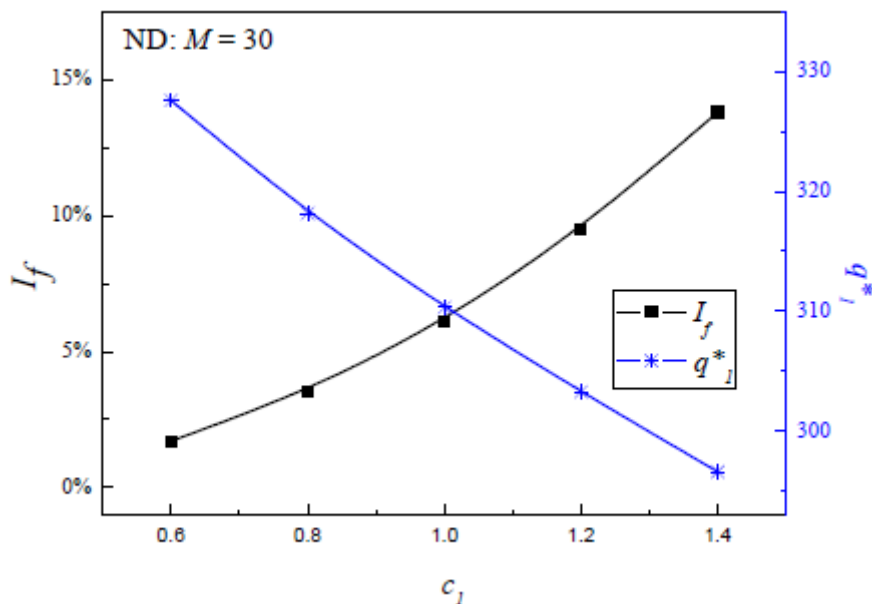


Figure 4.4 q_1^* and I_f versus c_1 under ND

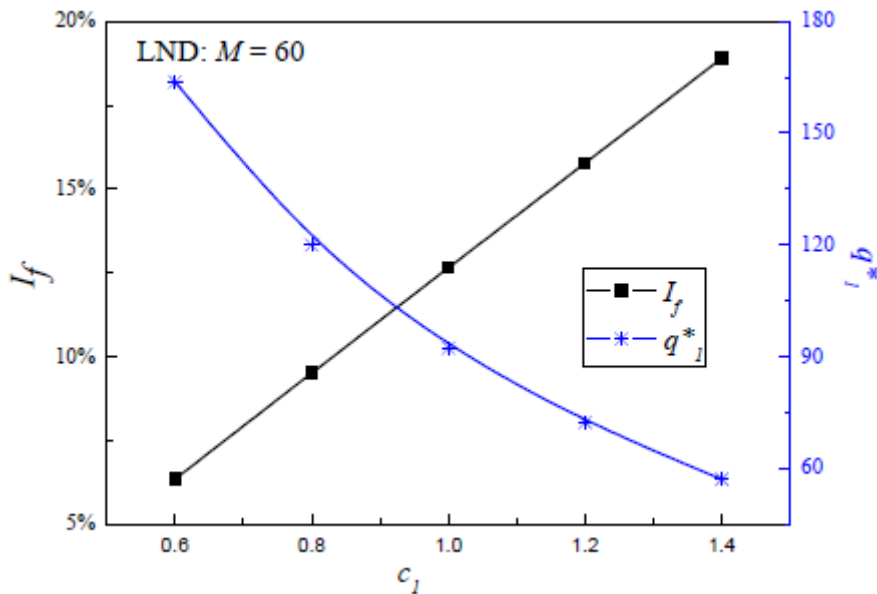


Figure 4.5 q_1^* and I_f versus c_1 under LND

Table 4.3 The Impact of c_2 on q_1^* and I_f for Different M under ND and LND

ND	$c_2 = 1.2$		$c_2 = 1.6$		$c_2 = 2.0$		$c_2 = 2.4$		$c_2 = 2.8$	
	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
M										
0	317.31		317.31		317.31		317.31		317.31	
5	314.35	40.57%	315.10	28.71%	315.81	18.48%	316.47	9.76%	317.06	2.65%
10	311.48	36.87%	313.01	25.03%	314.43	15.45%	315.72	7.82%	316.85	2.03%
15	308.70	33.23%	311.05	21.52%	313.19	12.67%	315.08	6.13%	316.68	1.52%
20	306.02	29.68%	309.24	18.21%	312.09	10.18%	314.54	4.69%	316.54	1.11%
25	303.43	26.26%	307.56	15.15%	311.14	7.99%	314.09	3.50%	316.44	0.79%
30	300.93	23.00%	306.05	12.35%	310.33	6.12%	313.74	2.55%	316.36	0.55%
35	298.54	19.91%	304.70	9.86%	309.66	4.57%	313.46	1.81%	316.30	0.37%
40	296.26	17.02%	303.52	7.68%	309.12	3.32%	313.26	1.24%	316.25	0.24%
45	294.10	14.35%	302.52	5.83%	308.70	2.34%	313.10	0.84%	316.23	0.16%
50	292.07	11.90%	301.69	4.31%	308.38	1.60%	313.00	0.55%	316.21	0.10%
LND										
M										
0	105.88		105.88		105.88		105.88		105.88	
10	100.03	42.16%	101.52	30.36%	102.92	19.89%	104.21	10.71%	105.39	2.98%
20	94.52	40.00%	97.51	28.27%	100.26	18.19%	102.76	9.62%	104.97	2.62%
30	89.36	37.84%	93.86	26.25%	97.90	16.62%	101.49	8.66%	104.61	2.32%
40	84.56	35.68%	90.54	24.33%	95.80	15.18%	100.39	7.80%	104.30	2.06%
50	80.12	33.55%	87.54	22.51%	93.93	13.86%	99.42	7.04%	104.04	1.84%
60	76.01	31.47%	84.82	20.81%	92.27	12.66%	98.58	6.36%	103.82	1.64%
70	72.25	29.45%	82.38	19.21%	90.80	11.56%	97.84	5.76%	103.62	1.47%
80	68.79	27.51%	80.18	17.73%	89.49	10.58%	97.18	5.23%	103.45	1.33%
90	65.64	25.66%	78.20	16.35%	88.32	9.68%	96.61	4.76%	103.30	1.20%
100	62.77	23.90%	76.43	15.09%	87.29	8.87%	96.10	4.34%	103.17	1.08%

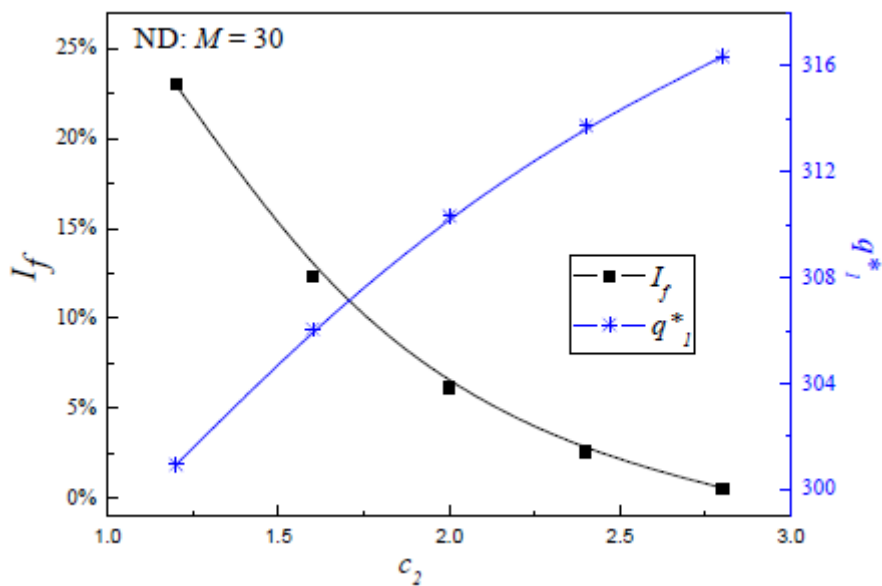


Figure 4.6 q_1^* and I_f versus c_2 under ND

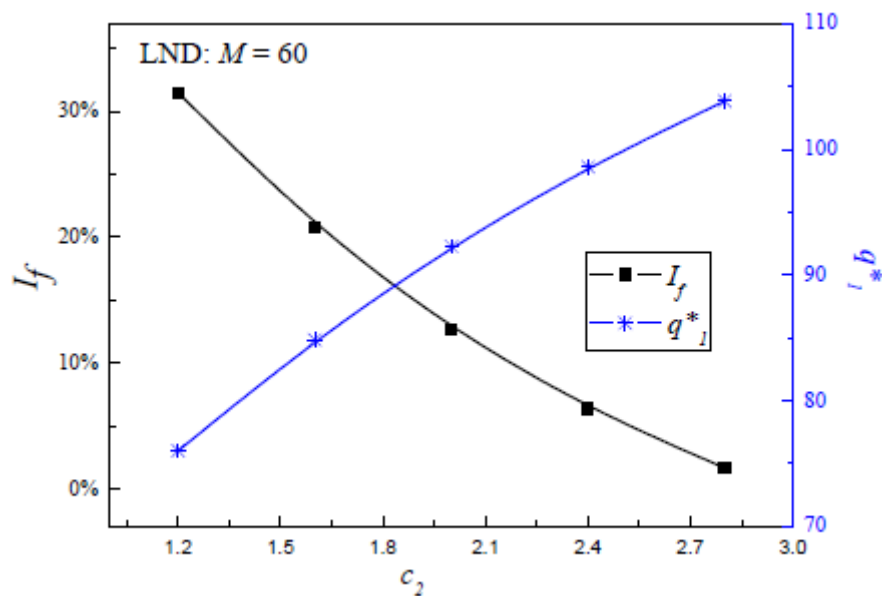


Figure 4.7 q_1^* and I_f versus c_2 under LND

Table 4.4 The Impact of s on q_1^* and I_f for Different M under ND and LND

ND	$s = 0.1$		$s = 0.2$		$s = 0.4$		$s = 0.6$		$s = 0.8$	
	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	315.14		317.31		322.53		329.60		340.85	
5	313.66	20.30%	315.81	18.48%	320.97	14.51%	327.99	10.05%	339.23	5.07%
10	312.30	17.06%	314.43	15.45%	319.56	11.96%	326.58	8.11%	337.86	3.93%
15	311.07	14.09%	313.19	12.67%	318.32	9.65%	325.36	6.39%	336.75	2.96%
20	309.96	11.40%	312.09	10.18%	317.24	7.61%	324.34	4.90%	335.86	2.15%
25	309.00	9.02%	311.14	7.99%	316.33	5.85%	323.50	3.65%	335.17	1.51%
30	308.17	6.97%	310.33	6.12%	315.57	4.38%	322.83	2.64%	334.66	1.03%
35	307.47	5.26%	309.66	4.57%	314.96	3.19%	322.32	1.85%	334.30	0.68%
40	306.91	3.86%	309.12	3.32%	314.49	2.26%	321.93	1.26%	334.04	0.43%
45	306.46	2.75%	308.70	2.34%	314.13	1.55%	321.66	0.83%	333.88	0.27%
50	306.13	1.91%	308.38	1.60%	313.87	1.03%	321.46	0.53%	333.77	0.16%
LND	$s = 0.1$		$s = 0.2$		$s = 0.4$		$s = 0.6$		$s = 0.8$	
M	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	98.48		105.88		125.97		159.45		232.01	
10	95.58	21.67%	102.92	19.89%	122.85	15.94%	156.18	11.36%	228.62	6.01%
20	93.01	19.74%	100.26	18.19%	120.03	14.71%	153.18	10.59%	225.46	5.68%
30	90.73	17.96%	97.90	16.62%	117.49	13.55%	150.44	9.86%	222.51	5.37%
40	88.71	16.34%	95.80	15.18%	115.21	12.48%	147.94	9.18%	219.77	5.07%
50	86.92	14.87%	93.93	13.86%	113.15	11.49%	145.66	8.54%	217.22	4.78%
60	85.34	13.54%	92.27	12.66%	111.31	10.57%	143.58	7.94%	214.86	4.51%
70	83.94	12.33%	90.80	11.56%	109.66	9.73%	141.70	7.38%	212.67	4.25%
80	82.70	11.25%	89.49	10.58%	108.17	8.96%	139.98	6.86%	210.64	4.01%
90	81.61	10.27%	88.32	9.68%	106.84	8.26%	138.42	6.38%	208.76	3.78%
100	80.63	9.38%	87.29	8.87%	105.64	7.62%	137.01	5.93%	207.02	3.56%

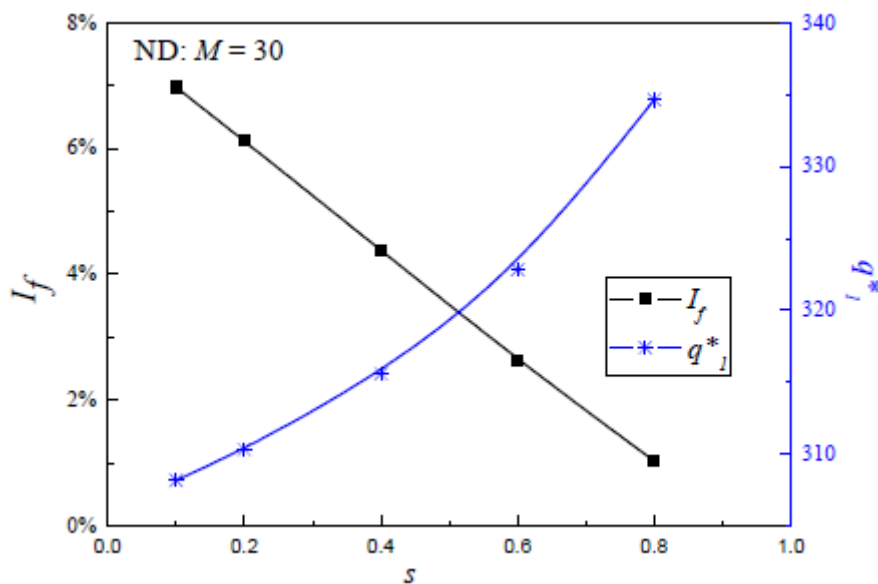


Figure 4.8 q_1^* and I_f versus s under ND

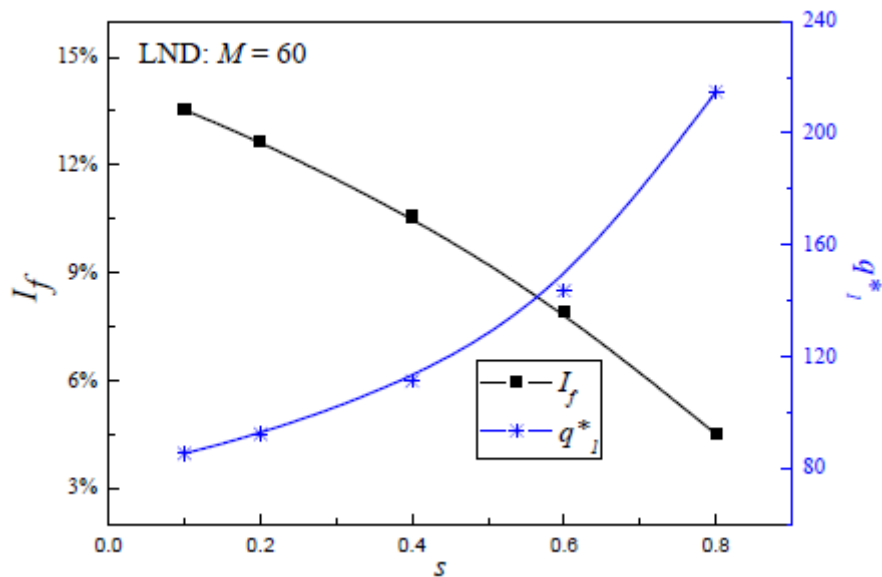


Figure 4.9 q_1^* and I_f versus s under LND

Table 4.5 The Impact of x_1 on q_1^* and I_f for Different M under ND and LND

ND	$x_1 = 180$		$x_1 = 240$		$x_1 = 300$		$x_1 = 360$		$x_1 = 420$	
	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
M										
0	197.31		257.31		317.31		377.31		437.31	
5	195.81	18.48%	255.81	18.48%	315.81	18.48%	375.81	18.48%	435.81	18.48%
10	194.43	15.45%	254.43	15.45%	314.43	15.45%	374.43	15.45%	434.43	15.45%
15	193.19	12.67%	253.19	12.67%	313.19	12.67%	373.19	12.67%	433.19	12.67%
20	192.09	10.18%	252.09	10.18%	312.09	10.18%	372.09	10.18%	432.09	10.18%
25	191.14	7.99%	251.14	7.99%	311.14	7.99%	371.14	7.99%	431.14	7.99%
30	190.33	6.12%	250.33	6.12%	310.33	6.12%	370.33	6.12%	430.33	6.12%
35	189.66	4.57%	249.66	4.57%	309.66	4.57%	369.66	4.57%	429.66	4.57%
40	189.12	3.32%	249.12	3.32%	309.12	3.32%	369.12	3.32%	429.12	3.32%
45	188.70	2.34%	248.70	2.34%	308.70	2.34%	368.70	2.34%	428.70	2.34%
50	188.38	1.60%	248.38	1.60%	308.38	1.60%	368.38	1.60%	428.38	1.60%
LND	$x_1 = 60$		$x_1 = 100$		$x_1 = 140$		$x_1 = 180$		$x_1 = 220$	
M	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	63.53		105.88		148.23		190.59		232.94	
10	60.67	18.74%	102.92	19.89%	145.22	20.40%	187.55	20.69%	229.88	20.88%
20	58.30	16.12%	100.26	18.19%	142.44	19.15%	184.69	19.70%	226.98	20.06%
30	56.36	13.86%	97.90	16.62%	139.87	17.96%	182.00	18.74%	224.21	19.26%
40	54.76	11.92%	95.80	15.18%	137.51	16.84%	179.48	17.83%	221.59	18.49%
50	53.45	10.27%	93.93	13.86%	135.34	15.78%	177.12	16.96%	219.10	17.75%
60	52.37	8.87%	92.27	12.66%	133.34	14.79%	174.91	16.12%	216.74	17.03%
70	51.48	7.69%	90.80	11.56%	131.50	13.86%	172.83	15.33%	214.50	16.35%
80	50.74	6.69%	89.49	10.58%	129.81	12.99%	170.89	14.57%	212.37	15.69%
90	50.12	5.84%	88.32	9.68%	128.26	12.17%	169.07	13.86%	210.36	15.05%
100	49.61	5.12%	87.29	8.87%	126.84	11.42%	167.37	13.18%	208.45	14.44%

(5) The optimal regular ordering quantity q_1^* increases when the price r , the emergency ordering cost c_2 and salvage value s are higher and the regular cost c_1 is lower, which is consistent with Corollary 4.1 (Tables 4.1–4.4 and Figures 4.2–4.9).

(6) Table 4.5 and Figure 4.10 show that the optimal regular ordering quantity q_1^* increases linearly but the emergency order size impact factor I_f remains the same as the initial demand forecast x_1 increases under ND, which complies with Theorem 4.4-2(a). Table 4.5 and Figure 4.11 indicate that q_1^* also increases with x_1 under LND, which is consistent with Theorem 4.4-1(a). Under LND, I_f increases with x_1 , which is different from that under ND. Because larger x_1 indicates the demand is higher during the selling season, the retailer would increase the ordering quantity. As the ordering quantity becomes larger, the restriction on the emergency ordering quantity plays a more critical role, resulting in an increase in I_f . However, under ND, the linear relationship between q_1^* and x_1 implies that the increased initial demand forecast is completely covered by the regular ordering quantity, resulting in unchangeable I_f .

(7) In Table 4.6 and Figures 4.12–4.13, the emergency order size impact factor I_f increases with the demand variability σ_1 under ND, while it increases and then decreases with σ_1 under LND. When σ_1 becomes larger, the demand forecast x_2 fluctuates more, which makes the restriction on the emergency ordering quantity tighter. As a result, the availability of the emergency order has a stronger influence as σ_1 increases. Therefore, I_f increases with σ_1 under ND, as shown in Figure 4.12. Under LND, larger σ_1 also leads to a smaller x_2 with high probability, based on the expression of $F_1(x_2|x_1)$ in Equation (4.21).

Due to the smaller x_2 , the emergency ordering quantity becomes smaller, resulting in a looser quantity restriction. Hence, the availability of the emergency order can have strong or weak influence if σ_1 is large under LND.

(8) The optimal regular ordering quantity q_1^* increases with σ_1 under ND, while it decreases under LND. As noted, if $M = 0$, all products are ordered at Stage 1. Hence, to some extent, q_{10}^* implies the total ordering quantity. Table 4.6 shows that the total ordering quantity becomes larger under ND and smaller under LND as σ_1 increases. As a result, the regular ordering quantity would increase under ND and decrease under LND with the increase in σ_1 .

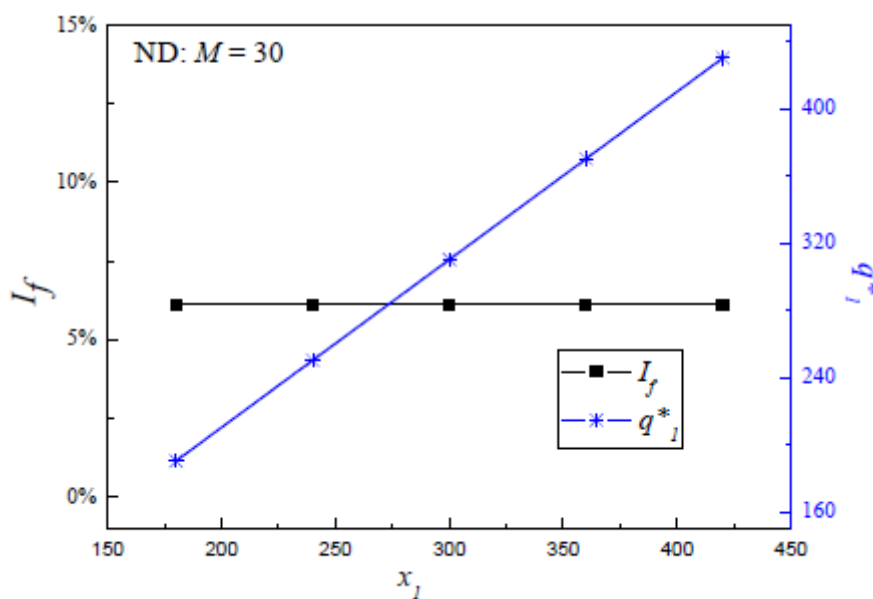


Figure 4.10 q_1^* and I_f versus x_1 under ND

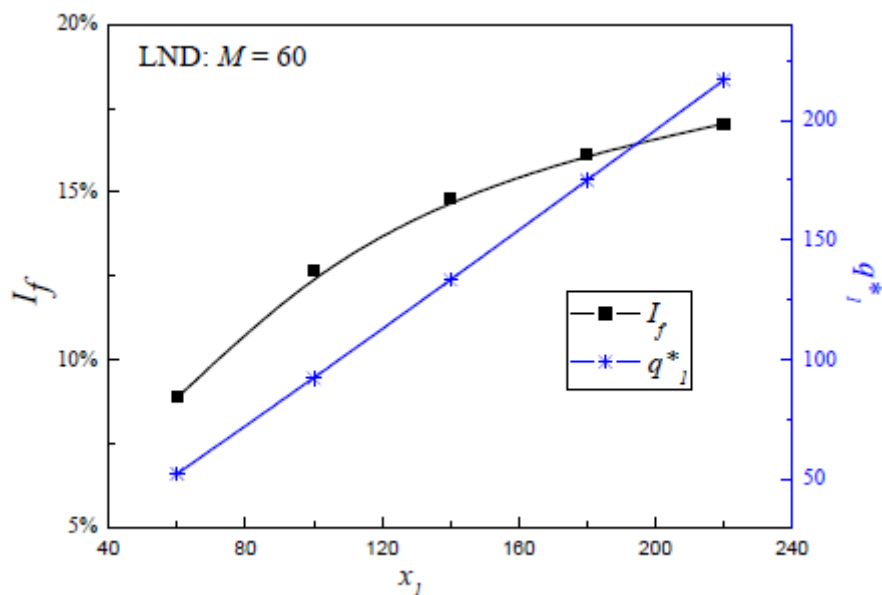


Figure 4.11 q_1^* and I_f versus x_1 under LND

Table 4.6 The Impact of σ_1 on q_1^* and I_f for Different M under ND and LND

ND	$\sigma_1 = 18$		$\sigma_1 = 24$		$\sigma_1 = 30$		$\sigma_1 = 36$		$\sigma_1 = 42$	
	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
M										
0	310.74		314.00		317.31		320.65		324.01	
5	309.44	12.87%	312.57	16.25%	315.81	18.48%	319.10	20.04%	322.42	21.19%
10	308.36	8.89%	311.31	12.74%	314.43	15.45%	317.65	17.40%	320.92	18.87%
15	307.53	5.75%	310.22	9.67%	313.19	12.67%	316.31	14.92%	319.51	16.65%
20	306.93	3.46%	309.31	7.08%	312.09	10.18%	315.09	12.63%	318.20	14.56%
25	306.53	1.92%	308.58	4.99%	311.14	7.99%	313.98	10.53%	316.98	12.60%
30	306.29	0.99%	308.02	3.36%	310.33	6.12%	312.99	8.65%	315.87	10.79%
35	306.15	0.47%	307.61	2.17%	309.66	4.57%	312.12	6.98%	314.86	9.14%
40	306.09	0.21%	307.32	1.34%	309.12	3.32%	311.37	5.54%	313.95	7.64%
45	306.05	0.08%	307.14	0.79%	308.70	2.34%	310.74	4.31%	313.14	6.31%
50	306.04	0.03%	307.02	0.45%	308.38	1.60%	310.22	3.29%	312.44	5.14%
LND	$\sigma_1 = 0.2$		$\sigma_1 = 1.0$		$\sigma_1 = 1.8$		$\sigma_1 = 2.6$		$\sigma_1 = 3.0$	
M	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f
0	112.76		105.88		54.06		14.60		5.97	
10	111.73	2.18%	102.92	19.89%	50.99	22.63%	11.99	21.05%	4.04	17.85%
20	111.24	0.89%	100.26	18.19%	48.41	20.76%	10.47	17.62%	3.37	13.76%
30	111.03	0.33%	97.90	16.62%	46.22	19.09%	9.51	15.15%	3.03	11.39%
40	110.95	0.12%	95.80	15.18%	44.37	17.59%	8.86	13.31%	2.83	9.80%
50	110.92	0.04%	93.93	13.86%	42.80	16.26%	8.40	11.90%	2.69	8.66%
60	110.91	0.01%	92.27	12.66%	41.46	15.07%	8.04	10.77%	2.59	7.79%
70	110.90	0.00%	90.80	11.56%	40.30	14.01%	7.77	9.85%	2.52	7.10%
80	110.90	0.00%	89.49	10.58%	39.30	13.07%	7.55	9.08%	2.46	6.54%
90	110.90	0.00%	88.32	9.68%	38.43	12.22%	7.37	8.43%	2.41	6.07%
100	110.90	0.00%	87.29	8.87%	37.66	11.46%	7.21	7.87%	2.37	5.67%

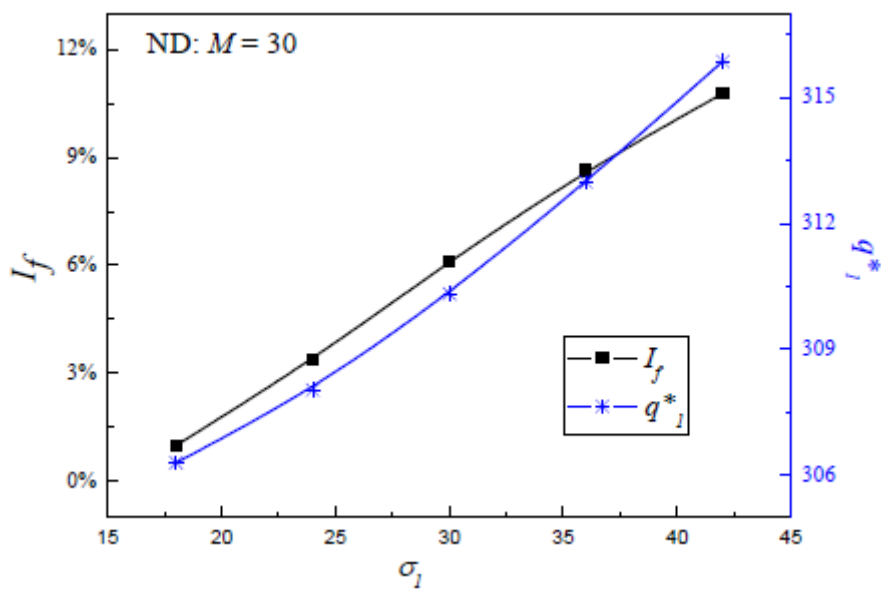


Figure 4.12 q_1^* and I_f versus σ_1 under ND

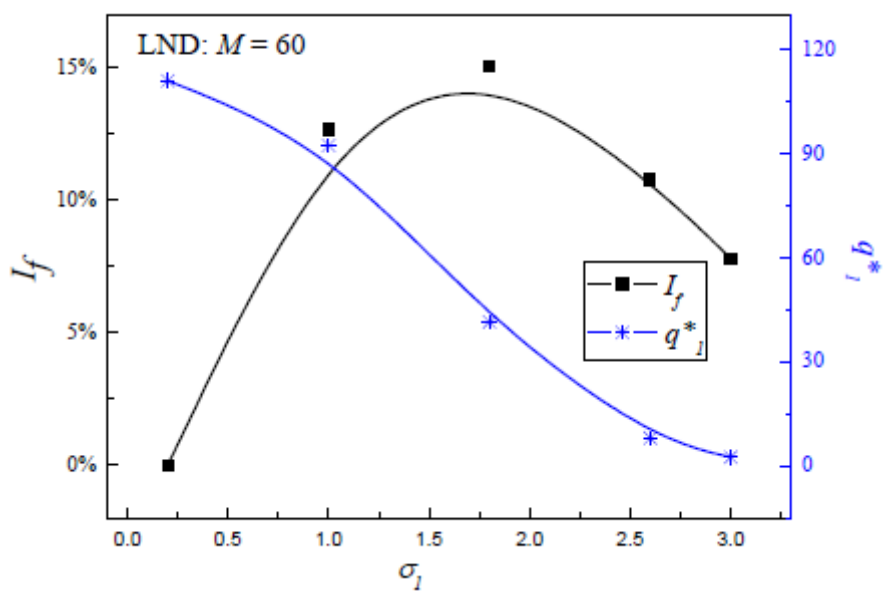


Figure 4.13 q_1^* and I_f versus σ_1 under LND

(9) The emergency order size impact factor I_f decreases with the demand variability σ_2 under both ND and LND in 4.7 and Figures 4.14–4.15. Note that σ_1 measures the

demand uncertainty that is resolved by the updated demand forecast at Stage 2, while σ_2 measures the demand uncertainty that cannot be resolved by the forecast. Hence, as the demand uncertainty σ_2 increases, the updated demand forecast at Stage 2 becomes less valuable, resulting in a decrease in I_f .

Table 4.7 The impact of σ_2 on q_1^* and I_f for different M under ND and LND

ND		$\sigma_2 = 1$		$\sigma_2 = 6$		$\sigma_2 = 16$		$\sigma_2 = 30$		$\sigma_2 = 50$	
M	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	
0	316.99		317.31		319.24		324.01		333.00		
5	315.29	23.85%	315.81	18.48%	318.13	10.04%	323.38	3.63%	332.79	0.67%	
10	313.72	20.46%	314.43	15.45%	317.16	7.93%	322.87	2.64%	332.64	0.44%	
15	312.26	17.27%	313.19	12.67%	316.33	6.12%	322.48	1.88%	332.54	0.28%	
20	310.94	14.32%	312.09	10.18%	315.64	4.61%	322.18	1.30%	332.47	0.18%	
25	309.76	11.64%	311.14	7.99%	315.08	3.38%	321.97	0.87%	332.43	0.11%	
30	308.72	9.25%	310.33	6.12%	314.64	2.41%	321.81	0.57%	332.40	0.06%	
35	307.83	7.18%	309.66	4.57%	314.31	1.67%	321.71	0.36%	332.38	0.04%	
40	307.08	5.44%	309.12	3.32%	314.07	1.12%	321.64	0.23%	332.37	0.02%	
45	306.47	4.00%	308.70	2.34%	313.89	0.74%	321.59	0.14%	332.37	0.01%	
50	305.99	2.87%	308.38	1.60%	313.77	0.47%	321.56	0.08%	332.36	0.01%	
LND		$\sigma_2 = 0.05$		$\sigma_2 = 0.2$		$\sigma_2 = 0.8$		$\sigma_2 = 1.5$		$\sigma_2 = 2.5$	
M	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	q_1^*	I_f	
0	106.76		105.88		90.92		54.62		12.23		
10	103.46	24.84%	102.92	19.89%	89.30	6.30%	54.10	1.04%	12.20	0.01%	
20	100.48	22.98%	100.26	18.19%	87.96	5.37%	53.77	0.74%	12.20	0.00%	
30	97.79	21.23%	97.90	16.62%	86.87	4.60%	53.55	0.53%	12.20	0.00%	
40	95.37	19.60%	95.80	15.18%	85.96	3.95%	53.40	0.40%	12.20	0.00%	
50	93.20	18.08%	93.93	13.86%	85.21	3.41%	53.29	0.30%	12.20	0.00%	
60	91.25	16.68%	92.27	12.66%	84.59	2.96%	53.22	0.23%	12.20	0.00%	
70	89.51	15.39%	90.80	11.56%	84.07	2.57%	53.17	0.18%	12.20	0.00%	
80	87.94	14.20%	89.49	10.58%	83.63	2.25%	53.13	0.14%	12.20	0.00%	
90	86.54	13.11%	88.32	9.68%	83.26	1.97%	53.10	0.11%	12.20	0.00%	
100	85.28	12.11%	87.29	8.87%	82.94	1.74%	53.07	0.09%	12.20	0.00%	

(10) The optimal regular ordering quantity q_1^* increases with σ_2 under ND, whereas it increases and then decreases under LND. 4.7 shows that q_{10}^* increases under ND and decreases under LND as σ_2 increases. As discussed in Part (8), q_1^* should increase with σ_2 under ND and decrease under LND. In addition, as noted in Part (9), the emergency order

becomes less important as σ_2 increases, and thus q_1^* would increase with σ_2 . Based on the combination of these two effects, q_1^* always increases under ND but may increase or decrease under LND, when σ_2 increases.

(11) The comparison of the emergency order size impact factor I_f between the two demand distributions is as follows: I_f keeps the same under ND but increases under LND as x_1 increases; I_f increases with σ_1 under ND but increases or decreases under LND; I_f decreases with σ_2 under both ND and LND. Therefore, the retailer should concentrate on the availability of the emergency order especially if (a) the initial demand forecast x_1 is high under LND, (b) the demand uncertainty σ_1 before the emergency order is high under ND, or (c) the demand uncertainty σ_2 after the emergency order is low under both ND and LND.

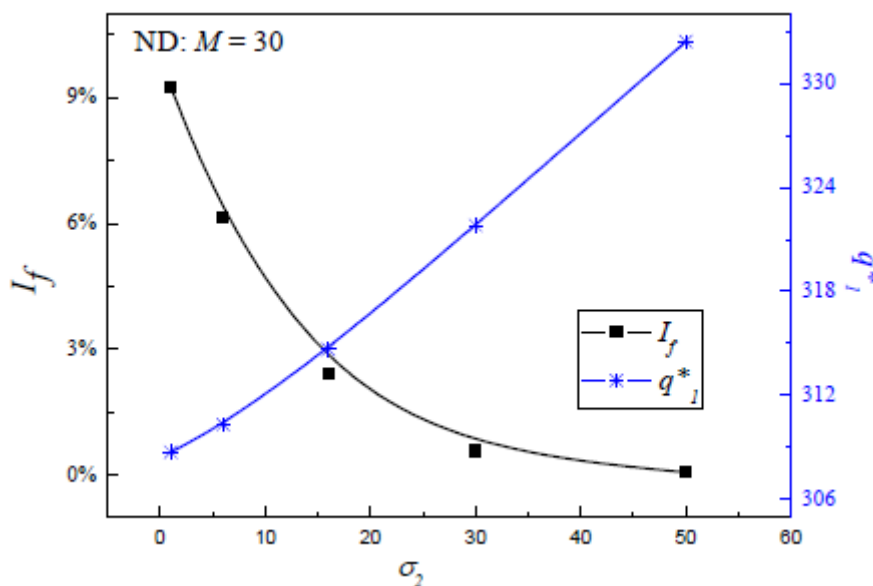


Figure 4.14 q_1^* and I_f versus σ_2 under ND

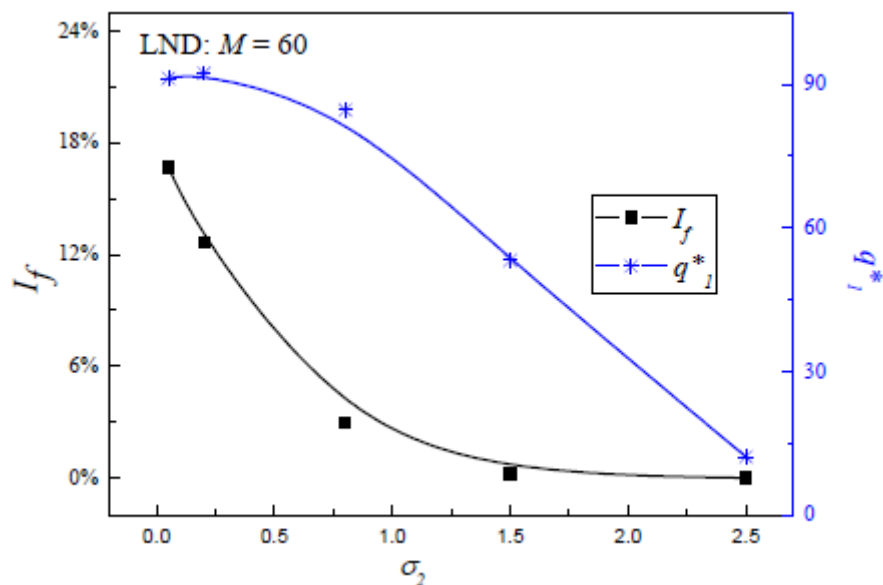


Figure 4.15 q_1^* and I_f versus σ_2 under LND

4.4 Conclusion

This study analyzes the ordering decisions when the retailer has a regular ordering opportunity and a more expensive emergency opportunity with capacity constraints and demand forecast updating. In this setting, the optimal ordering policy is obtained, and the structural properties of the optimal solutions are identified. Moreover, the changes of the regular ordering quantity and impact of the availability of the emergency order are investigated under different prices, salvage values, unit costs, maximum ordering quantity limits and demand characteristics. Based on the investigation, an algorithm is proposed to determine the optimal regular ordering quantity directly. Numerical results indicate that the availability of the emergency order is more important to the retailer when (a) the initial demand forecast is higher under LND, (b) the demand uncertainty before the emergency order is higher under ND, or (c) the demand uncertainty after the emergency order is lower

under both ND and LND.

4.5 Appendix

Proof of Theorem 4.4 1-(a). From Lemma 4.1 and Theorem 4.1,

$$\frac{dv_2(q_1, x_2)}{dx_2} = \frac{\partial v_2(q_1, x_2)}{\partial x_2} + \frac{\partial v_2(q_1, x_2)}{\partial q_2^*} \frac{\partial q_2^*}{\partial x_2} = \frac{\partial v_2(q_1, x_2)}{\partial x_2} = \frac{\partial SP(q_1 + q_2^*, x_2)}{\partial x_2}$$

Thus, for ND and LND, substituting Equations (4.7) and (4.21) into Equation (4.2) and taking

the first derivatives w.r.t. x_2 give

$$\begin{aligned} & \frac{dv_2(q_1, x_2)}{dx_2} \\ = & \begin{cases} (r-s)\Phi\left(\frac{q_1-x_2}{\sigma_2}\right) > 0, & x_2 < x_L(q_1) \\ (r-c_2) > 0, & x_L(q_1) \leq x_2 \leq x_M(q_1); \text{ and} \\ (r-s)\Phi\left(\frac{M+q_1-x_2}{\sigma_2}\right) > 0, & x_2 > x_M(q_1) \end{cases} \end{aligned} \quad (4.28)$$

$$\begin{aligned} & \frac{dv_2(q_1, x_2)}{dx_2} \\ = & \begin{cases} (r-s) \left[1 - \Phi\left(\sigma_2 - \left(\frac{\ln q_1 - \ln x_2 + 0.5\sigma_2^2}{\sigma_2}\right)\right) \right] > 0, & x_2 < x_L(q_1) \\ (r-s) \left[1 - \Phi\left(\sigma_2 - \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\right) \right] > 0, & x_L(q_1) \leq x_2 \leq x_M(q_1). \\ (r-s) \left[1 - \Phi\left(\sigma_2 - \left(\frac{\ln(M+q_1) - \ln x_2 + 0.5\sigma_2^2}{\sigma_2}\right)\right) \right] > 0, & x_2 > x_M(q_1) \end{cases} \end{aligned} \quad (4.29)$$

Equations (4.28) and (4.29) show that $v_2(q_1, x_2)$ increases with x_2 under ND and LND, respectively.

According to Equation (4.22), $q_2^o(q_1, x_2)$ increases with x_2 . Hence, in Equation (4.7), q_2^* increases with x_2 .

1-(b). In Equations (4.23) and (4.24), the expressions of $x_L(q_1)$ and $x_M(q_1)$ do not

contain x_1 . For ND (or LND), substituting Equation (4.21) and $x_2 = \bar{x} + x_1$ (or $x_2 = \bar{x}x_1$)

into Equation (4.12) and taking the first derivative w.r.t. x_1 give

$$\begin{aligned} \frac{\partial w_1(q_1, x_1)}{\partial x_1} &= \int_0^{x_L(q_1)-x_1} \frac{\partial w_2(q_1, 0, \bar{x} + x_1)}{\partial x_1} \phi\left(\frac{\bar{x}}{\sigma_1}\right) \frac{1}{\sigma_1} d\bar{x} \\ &\quad + \int_{x_L(q_1)-x_1}^{x_M(q_1)-x_1} \frac{\partial w_2(q_1, q_2^o(q_1, \bar{x} + x_1), \bar{x} + x_1)}{\partial x_1} \phi\left(\frac{\bar{x}}{\sigma_1}\right) \frac{1}{\sigma_1} d\bar{x} \\ &\quad + \int_{x_M(q_1)-x_1}^{\infty} \frac{\partial w_2(q_1, M, \bar{x} + x_1)}{\partial x_1} \phi\left(\frac{\bar{x}}{\sigma_1}\right) \frac{1}{\sigma_1} d\bar{x} > 0; \text{ or} \\ \frac{\partial w_1(q_1, x_1)}{\partial x_1} &= \int_0^{x_L(q_1)/x_1} \frac{\partial w_2(q_1, 0, \bar{x}x_1)}{\partial x_1} \phi\left(\frac{\ln\bar{x} + 0.5\sigma_1^2}{\sigma_1}\right) \frac{1}{\bar{x}\sigma_1} d\bar{x} \\ &\quad + \int_{x_L(q_1)/x_1}^{x_M(q_1)/x_1} \frac{\partial w_2(q_1, q_2^o(q_1, \bar{x}x_1), \bar{x}x_1)}{\partial x_1} \phi\left(\frac{\ln\bar{x} + 0.5\sigma_1^2}{\sigma_1}\right) \frac{1}{\bar{x}\sigma_1} d\bar{x} \\ &\quad + \int_{x_M(q_1)/x_1}^{\infty} \frac{\partial w_2(q_1, M, \bar{x}x_1)}{\partial x_1} \phi\left(\frac{\ln\bar{x} + 0.5\sigma_1^2}{\sigma_1}\right) \frac{1}{\bar{x}\sigma_1} d\bar{x} > 0. \end{aligned}$$

where the inequalities holds due to Equations (4.28) and (4.29). Hence, for both ND and

LND,

$$\frac{dv_1(x_1)}{dx_1} = \frac{\partial v_1(x_1)}{\partial x_1} + \frac{\partial v_1(x_1)}{\partial q_1^*} \frac{\partial q_1^*}{\partial x_1} = \frac{\partial v_1(x_1)}{\partial x_1} = \frac{\partial v_1(x_1)}{\partial x_1} = \frac{\partial w_1(q_1, x_1)}{\partial x_1} > 0.$$

For ND (or LND), substituting Equation (4.25) (or Equation (4.26)) $x_2 = \bar{x} + x_1$

(or $x_2 = \bar{x}x_1$) in $J(q_1^*, M, c_1, c_2, s, r, x_1)$ expression and taking the first derivative w.r.t. x_1

yield

$$\begin{aligned} \frac{\partial J}{\partial x_1} &= \int_0^{x_L(q_1^*)-x_1} (r-s) \phi\left(\frac{q_1^* - \bar{x} - x_1}{\sigma_2}\right) \phi\left(\frac{\bar{x}}{\sigma_1}\right) \frac{1}{\sigma_1\sigma_2} d\bar{x} \\ &\quad + \int_{x_M(q_1^*)-x_1}^{\infty} (r-s) \frac{1}{\sigma_2} \phi\left(\frac{q_1^* + M - \bar{x} - x_1}{\sigma_2}\right) \phi\left(\frac{\bar{x}}{\sigma_1}\right) \frac{1}{\sigma_1\sigma_2} d\bar{x} > 0; \text{ or} \end{aligned}$$

$$\begin{aligned} & \frac{\partial J}{\partial x_1} \\ &= \int_0^{x_L(q_1^*)/x_1} (r-s)\phi\left(\frac{\ln q_1^* - \ln \bar{x} - \ln x_1 + 0.5\sigma_2^2}{\sigma_2}\right)\phi\left(\frac{\ln \bar{x} + 0.5\sigma_1^2}{\sigma_1}\right)\frac{1}{\sigma_1\sigma_2 x_1 \bar{x}} d\bar{x} \\ &+ \int_{x_M(q_1^*)/x_1}^{\infty} (r-s)\phi\left(\frac{\ln(q_1^* + M) - \ln \bar{x} - \ln x_1 + 0.5\sigma_2^2}{\sigma_2}\right)\phi\left(\frac{\ln \bar{x} + 0.5\sigma_1^2}{\sigma_1}\right)\frac{1}{\sigma_1\sigma_2 x_1 \bar{x}} d\bar{x} \\ &> 0. \end{aligned}$$

According to Equation (4.18),

$$\frac{\partial q_1^*}{\partial x_1} = -\frac{\partial J}{\partial x_1} / \frac{\partial J}{\partial q_1^*} > 0.$$

Therefore, under ND and LND, $v_1(x_1)$ and q_1^* increase with x_1 .

2-(a). Assuming that $q_1^*(x_1^b)$ satisfies Equation (4.25) when $x_1 = x_1^b$, we will prove that the solution of Equation (4.25) is $q_1^*(x_1^a) = x_1^a - x_1^b + q_1^*(x_1^b)$ when $x_1 = x_1^a$. Substituting x_1^a and $q_1^*(x_1^a)$ into the left hand side of Equation (4.25) yields

$$\begin{aligned} & (c_2 - c_1) - \int_0^{x_L(q_1^*(x_1^a))} \left[(r-s)\Phi\left(\frac{q_1^*(x_1^a) - x_2}{\sigma_2}\right) - (r-c_2) \right] \phi\left(\frac{x_2 - x_1^a}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 \\ &+ \int_{x_M(q_1^*(x_1^a))}^{\infty} \left[(r-c_2) - (r-s)\Phi\left(\frac{q_1^*(x_1^a) + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1^a}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2. \end{aligned}$$

According to Equations (4.23) and (4.24), substituting $q_1^*(x_1^a) = x_1^a - x_1^b + q_1^*(x_1^b)$ and $x'_2 = x_2 + x_1^b - x_1^a$ in this expression gives

$$\begin{aligned} & (c_2 - c_1) - \int_0^{x_L(q_1^*(x_1^b))} \left[(r-s)\Phi\left(\frac{q_1^*(x_1^b) - x'_2}{\sigma_2}\right) - (r-c_2) \right] \phi\left(\frac{x'_2 - x_1^b}{\sigma_1}\right) \frac{1}{\sigma_1} dx'_2 \\ &+ \int_{x_M(q_1^*(x_1^b))}^{\infty} \left[(r-c_2) - (r-s)\Phi\left(\frac{q_1^*(x_1^b) + M - x'_2}{\sigma_2}\right) \right] \phi\left(\frac{x'_2 - x_1^b}{\sigma_1}\right) \frac{1}{\sigma_1} dx'_2 = 0, \end{aligned}$$

where the equality establishes because it is assumed that $q_1^*(x_1^b)$ satisfies Equation (4.25)

when $x_1 = x_1^b$. Hence, $q_1^*(x_1^a) = x_1^a - x_1^b + q_1^*(x_1^b)$ satisfies Equation (4.25) when

$x_1 = x_1^a$. This implies that q_1^* is positive linear with x_1 . In addition, note that

$$\begin{aligned} & \int_{x_M(q_1^*(x_1^a))}^{\infty} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^*(x_1^a) + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1^a}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 \\ &= \int_{x_M(q_1^*(x_1^b))}^{\infty} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^*(x_1^b) + M - x'_2}{\sigma_2}\right) \right] \phi\left(\frac{x'_2 - x_1^b}{\sigma_1}\right) \frac{1}{\sigma_1} dx'_2 \end{aligned}$$

According Equation (4.19),

$$\begin{aligned} I_f(x_1^a) &= \int_{x_M(q_1^*(x_1^a))}^{\infty} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^*(x_1^a) + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1^a}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2, \\ I_f(x_1^b) &= \int_{x_M(q_1^*(x_1^b))}^{\infty} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^*(x_1^b) + M - x'_2}{\sigma_2}\right) \right] \phi\left(\frac{x'_2 - x_1^b}{\sigma_1}\right) \frac{1}{\sigma_1} dx'_2. \end{aligned}$$

Hence, $I_f(x_1^a) = I_f(x_1^b)$, which implies that I_f does not change with x_1 .

2-(b). When σ_2 approaches zero, the optimal regular ordering quantity and emergency order size impact factor are denoted by q_1^{*0} and I_f^0 . Based on Equations (4.25) and (4.24),

Equation (4.19) under ND can be rewritten as

$$\begin{aligned} I_f^0 &= \lim_{\sigma_2 \rightarrow 0^+} \int_{x_M(q_1^{*0})}^{\infty} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^{*0} + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 \\ &= \lim_{\sigma_2 \rightarrow 0^+} \int_{M+q_1^{*0}}^{\infty} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^{*0} + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 \\ &\quad + \lim_{\sigma_2 \rightarrow 0^+} \int_{M+q_1^{*0} - \Phi^{-1}\left(\frac{r-c_2}{r-s}\right)\sigma_2}^{M+q_1^{*0}} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^{*0} + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 \\ &= \lim_{\sigma_2 \rightarrow 0^+} \int_{M+q_1^{*0}}^{\infty} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^{*0} + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 \\ &= \int_{M+q_1^{*0}}^{\infty} \lim_{\sigma_2 \rightarrow 0^+} \left[(r - c_2) - (r - s)\Phi\left(\frac{q_1^{*0} + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2. \end{aligned}$$

If $x_2 > M + q_1^{*0}$, $\lim_{\sigma_2 \rightarrow 0^+} \Phi\left(\frac{q_1^{*0} + M - x_2}{\sigma_2}\right) = 0$. Thus,

$$I_f^0 = \int_{M+q_1^{*0}}^{\infty} (r - c_2) \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2. \quad (4.30)$$

In a similar way, according to Equations (4.25) and (4.23), Equation (4.20) under ND can also be rewritten as

$$\begin{aligned} I_f^0 &= \lim_{\sigma_2 \rightarrow 0^+} \left[\int_0^{x_L(q_1^{*0})} \left[(r - s) \Phi\left(\frac{q_1^{*0} - x_2}{\sigma_2}\right) - (r - c_2) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 - (c_2 - c_1) \right] \\ &= \int_{-\infty}^{q_1^{*0}} \lim_{\sigma_2 \rightarrow 0^+} \left[(r - s) \Phi\left(\frac{q_1^{*0} - x_2}{\sigma_2}\right) - (r - c_2) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 - (c_2 - c_1). \end{aligned}$$

If $x_2 < q_1^{*0}$, $\lim_{\sigma_2 \rightarrow 0^+} \Phi\left(\frac{q_1^{*0} - x_2}{\sigma_2}\right) = 1$. Thus,

$$I_f^0 = \int_0^{q_1^{*0}} [(r - s) - (r - c_2)] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 - (c_2 - c_1). \quad (4.31)$$

If $q_1^{*0} > x_L(q_1^*)$, in Equation (4.31),

$$\begin{aligned} I_f^0 &= \int_0^{q_1^{*0}} [(r - s) - (r - c_2)] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 - (c_2 - c_1) \\ &> \int_0^{x_L(q_1^*)} \left[(r - s) \Phi\left(\frac{q_1^* - x_2}{\sigma_2}\right) - (r - c_2) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 - (c_2 - c_1) = I_f. \end{aligned}$$

Otherwise, $M + q_1^{*0} \leq x_M(q_1^*)$ since $x_M(q_1^*) = x_L(q_1^*) + M$ under ND. In this situation, in Equation (4.30),

$$\begin{aligned} I_f^0 &= \int_{M+q_1^{*0}}^{\infty} (r - c_2) \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 \\ &> \int_{x_M(q_1^*)}^{\infty} \left[(r - c_2) - (r - s) \Phi\left(\frac{q_1^* + M - x_2}{\sigma_2}\right) \right] \phi\left(\frac{x_2 - x_1}{\sigma_1}\right) \frac{1}{\sigma_1} dx_2 = I_f. \end{aligned}$$

Therefore, $I_f^0 > I_f$ whether $q_1^{*0} > x_L(q_1^*)$ or not. This completes the proofs. Q.E.D.

Chapter 5. Analysis of Reactive Production with Preprocessing Restriction in Supply Chains with Forecast Updates

This chapter considers a two-mode production system that allows the supply chain to utilize a fast reactive production with improved demand information before the selling season. However, the reactive production quantity is limited by the required preprocessing decision before the demand information is updated. The condition under which adding a reactive production is beneficial to the supply chain depends on the demand forecast updating process, besides the cost parameters. To illustrate the benefit of this system, the two-mode production system is compared with two single-mode production systems, one that is purely speculative and the other purely reactive. In addition, a pricing contract with a return policy is demonstrated to efficiently coordinate the supply chain since the contract allows the maximization and arbitrary allocation of the total supply chain profit. Contrary to the results in existing literature, this result continues to hold whether demand information is known to the manufacturer or not. Numerical examples demonstrate that the benefit of the two-mode production can be as large as 104% when the preprocessing limit exists.

5.1 Introduction

This study investigates a capacitated supply chain with two-mode production (i.e., speculative production and reactive production), where the reactive production quantity is limited by the raw preparation in advance. Before analyzing this model in detail, we would like to summarize our main findings.

First, when preprocessing limits exist, the two-mode production is not as beneficial as previously believed. Numerical examples show that if the premium paid for the delayed production is very high, the two-mode production has no value compared with the purely speculative production mode. We derive the condition under which the two-mode production is beneficial. In particular, we find that the condition depends on the demand forecast updating process, which is not the case in the model without material preprocessing (Donohue, 2000).

Second, when compared with the purely speculative production, the two-mode production can reduce the upfront investment cost and increase the available quantity when the demand information is updated. In addition, it is found that the two-mode production always outweighs the purely reactive production strategy. That is because the two-mode production enables the supply chain to enjoy the lower early production cost without decreasing the available quantity.

Third, numerical examples show that the advantage over the purely speculative production mode is large for low production delay costs (especially a low material cost), a medium sales price, high customer shortage sensitivity, high material compatibility, high resolved demand uncertainty and low unresolved demand uncertainty. In contrast, the advantage over the purely reactive production mode is large for high production delay costs (especially a high material cost), a low sales price, high customer shortage sensitivity, low material compatibility, low resolved demand uncertainty and high unresolved demand uncertainty.

Fourth, we coordinate the supply chain with the two-mode production model by using a pricing contract with a return policy in the form of (c_R, l_1, l_2, s_r) , where c_R is the reservation fee for the second order, l_1 and l_2 are the wholesale prices, and s_r is the return price. The coordination contract parameters are proved to be flexible and allow for arbitrary profit allocation between the manufacturer and buyer, leading to a possible Pareto improvement.

Fifth, the coordination contract parameters are found to remain the same whether or not the demand information is available to the manufacturer. In contrast, they change with the manufacturer's access to demand information in Donohue (2000). That's because that the reservation for the second order from the buyer reveals some demand information that complements the manufacturer's incomplete information in our model.

The remainder of this chapter is organized as follows. Section 5.2 develops the two-mode production model, characterizes related optimal decisions, and compares it with other single-mode production models. Section 5.3 analyzes decisions of the manufacturer and buyer, optimizes a pricing contract to coordinate them, and investigates the profit allocation between them. Section 5.4 summarizes the results and provides directions for future research. For the sake of clarity, all the proofs are deferred to Appendix.

5.2 Model

The sequence of events for the two-mode production model is illustrated in Figure 5.1. In Stage 1, based on the initial demand forecast, the supply chain makes speculative production

p_1^c and prepares raw material M^c for the reactive production. In Stage 2, the supply chain undertakes the high-cost reactive production $p_2^c \leq M^c$ according to the improved demand forecast. Finished products are ready before the selling season during which the demand is realized. Unused material and unsold products are salvaged at the end of the selling season.

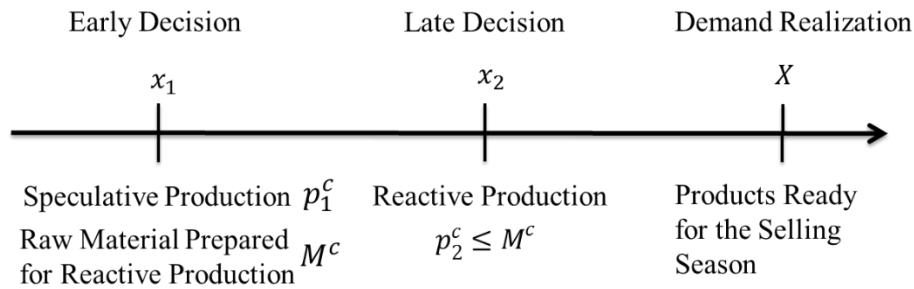


Figure 5.1 The Sequence of Events

The demand forecasts X_1 and X_2 are made in Stages 1 and 2 for the actual demand X during the selling season. Given the realized demand forecast x_1 in Stage 1, X_2 is assumed to have distribution density functions $F_1(\cdot | x_1)$ and $f_1(\cdot | x_1)$, and X has functions $F(\cdot | x_1)$ and $f(\cdot | x_1)$. Given the updated demand forecast x_2 in Stage 2, X follows distribution and density functions $F_2(\cdot | x_2)$ and $f_2(\cdot | x_2)$. The inverse distribution function $F_2^{-1}(\cdot | x_2)$ is assumed to be a differentiable and strictly increasing function with x_2 , and asymptotically go to infinity as $x_2 \rightarrow \infty$. Note that demand forecasts are assumed to be non-negative. The demand forecast updating process is identical to the one in Donohue (2000).

The parameters used in our analysis are as follows:

X : actual demand during the selling season

X_i :	demand forecast for X in Stage i
r :	unit sales price
l_i :	unit wholesale price in Stage i
s :	unit salvage value of the leftover finished product
π :	unit shortage penalty cost
c_m :	unit raw material cost
c_1 :	sum of unit early production cost and unit raw material cost
c_2 :	unit delayed production cost
$c_2 + c_m - c_1$:	unit production delay premium
s_m :	unit salvage value of the unused raw material
$c_m - s_m$:	unit material expiry cost
c_R :	unit reservation fee
A :	large quantity of available products

Additional notations are described as needed. It is assumed that $r > c_2 + c_m > c_1 > s$ and $c_1 - s > c_m > s_m$, where the second inequality states that it is more expensive to fill an order through the delayed production than the early production. The assumptions while being reasonable are designed to eliminate the trivial and irrelevant cases.

5.2.1 Centralized Supply Chain

Here, we assume that there is only one decision maker (the manufacturer) in the supply chain.

In this centralized system, the manufacturer needs to choose the early production quantity p_1^c

and amount of raw material M^c first, and the delayed production quantity $p_2^c \leq M^c$ after forecast updates. The optimal values of these decision variables can be derived through the dynamic programming. The optimal production quantity p_2^{c*} in Stage 2 is firstly determined for any given p_1^c and M^c , and then the optimal production quantity p_1^{c*} and material amount M^{c*} in Stage 1 are obtained by maximizing the total expected profit function.

Given p_1^c , M^c and the current demand forecast x_2 , the optimal expected profit function in Stage 2 is

$$v_2(p_1^c, x_2) = \max\{w_2(p_1^c, p_2^c, x_2), 0 \leq p_2^c \leq M^c\},$$

where

$$\begin{aligned} w_2(p_1^c, p_2^c, x_2) = & -c_2 p_2^c + r E_{F_2(\cdot|x_2)} [\min(p_1^c + p_2^c, X)] - \pi E_{F_2(\cdot|x_2)} [(X - p_1^c - p_2^c)^+] \\ & + s E_{F_2(\cdot|x_2)} [(p_1^c + p_2^c - X)^+] + s_m (M^c - p_2^c). \end{aligned} \quad (5.1)$$

Define $p^c = p_1^c + p_2^c$ and $\bar{w}(p^c, x_2) = -(c_2 + s_m)p^c + r E_{F_2(\cdot|x_2)} [\min(p^c, X)] - \pi E_{F_2(\cdot|x_2)} [(X - p^c)^+] + s E_{F_2(\cdot|x_2)} [(p^c - X)^+]$. Hence, (4.1) can be rewritten as follows

$$w_2(p_1^c, p_2^c, x_2) = (c_2 + s_m)p_1^c + s_m M^c + \bar{w}(p^c, x_2). \quad (5.2)$$

Because $\bar{w}(p^c, x_2)$ is a standard newsvendor form with conditional demand distribution $F_2(X|x_2)$, $w_2(p_1^c, p_2^c, x_2)$ is concave in p_2^c and the optimal production quantity p_2^{c*} is obtained by the following lemma.

Lemma 5.1. *Given p_1^c , M^c and x_2 , the optimal production quantity in Stage 2 is given by:*

$$p_2^{c*} = \begin{cases} 0 & x_2 < x_t(p_1^c) \\ \bar{p}_2^c(p_1^c) & x_t(p_1^c) \leq x_2 \leq x_t(p_1^c + M^c), \\ M^c & x_2 > x_t(p_1^c + M^c) \end{cases} \quad (5.3)$$

where $\bar{p}_2^c(p_1^c) = F_2^{-1}((r + \pi - c_2 - s_m)/(r + \pi - s)|x_2) - p_1^c$, and the threshold value $x_t(z)$ satisfies $F_2(z|x_t) = (r + \pi - c_2 - s_m)/(r + \pi - s)$.

According to Lemma 5.1, given the current demand forecast x_1 , the total expected profit function is

$$v(x_1) = \max\{w(p_1^c, M^c, x_1), p_1^c \geq 0 \text{ and } M^c \geq 0\},$$

where

$$w(p_1^c, M^c, x_1)$$

$$= -c_1 p_1^c - c_m M^c + E_{F_1(\cdot|x_1)}[v_2(p_1^c, x_2)] \quad (5.4)$$

$$= -c_1 p_1^c - c_m M^c + \int_0^{x_t(p_1^c)} w_2(p_1^c, 0, x_2) f_1(x_2|x_1) dx_2$$

$$+ \int_{x_t(p_1^c)}^{x_t(p_1^c + M^c)} w_2(p_1^c, \bar{p}_2^c, x_2) f_1(x_2|x_1) dx_2$$

$$+ \int_{x_t(p_1^c + M^c)}^{\infty} w_2(p_1^c, M^c, x_2) f_1(x_2|x_1) dx_2. \quad (5.5)$$

Hence, the optimal production quantity p_1^{c*} and amount of raw material M^{c*} in Stage 1 can be obtained in Theorem 5.1.

Theorem 5.1

- (1) Given x_1 , $w(p_1^c, M^c, x_1)$ is jointly concave in p_1^c and M^c .
- (2) The system will benefit from an additional production, i.e., $M^{c*} > 0$, if only if the following relationship for the system parameters is true:

$$c_0(p_{s1}^{c*}) > (c_2 - c_1) + c_m \quad (5.6)$$

or,

$$c_u(p_{s1}^{c*}) > c_m - s_m, \quad (5.7)$$

where

$$p_{s1}^{c*} = F^{-1}\left(\frac{r+\pi-c_1}{r+\pi-s} \mid x_1\right);$$

$$c_0(p_{s1}^{c*}) = \int_0^{x_t(p_{s1}^{c*})} [(r + \pi - s)F_2(p_{s1}^{c*} \mid x_2) - (r + \pi - c_2 - s_m)] f_1(x_2 \mid x_1) dx_2;$$

$$c_u(p_{s1}^{c*}) = \int_{x_t(p_{s1}^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*} \mid x_2)] f_1(x_2 \mid x_1) dx_2.$$

(3) If Equation (5.6) is satisfied, the optimal production quantity p_1^{c*} and amount of raw material M^{c*} satisfy:

$$\begin{aligned} \int_0^{x_t(p_1^{c*})} [(r + \pi - s)F_2(p_1^{c*} \mid x_2) - (r + \pi - c_2 - s_m)] f_1(x_2 \mid x_1) dx_2 \\ = (c_2 - c_1) + c_m, \end{aligned} \quad (5.8)$$

$$\begin{aligned} \int_{x_t(p_1^{c*} + M^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^{c*} + M^{c*} \mid x_2)] f_1(x_2 \mid x_1) dx_2 \\ = c_m - s_m. \end{aligned} \quad (5.9)$$

Otherwise, $p_1^{c*} = p_{s1}^{c*}$ and $M^{c*} = 0$.

Theorem 5.1 states that it is not always worth adding a reactive production although it enables the decision maker to access the more accurate demand information. Moreover, the condition where it is valuable to add the reactive production and the optimal production policy in Stage 1 are provided.

If Equations (5.6) and (5.7) are not satisfied, it is not profitable to add the reactive production. Note that $p_{s1}^{c*} = F^{-1}\left(\frac{r+\pi-c_1}{r+\pi-s} \middle| x_1\right)$ is the optimal production quantity in the traditional scenario where only the speculative production is allowed, i.e., $M^c = 0$. Compared with the first-order condition in the standard newsvendor model, $c_0(p_{s1}^{c*})$ in Equation (5.6) and $c_u(p_{s1}^{c*})$ in Equation (5.7) can be quantified by the expected overage cost and underage cost when $p_1^c = p_{s1}^{c*}$. If the expected overage cost is smaller than the production delay premium ($c_2 + c_m - c_1$), it is not beneficial to replace one unit of p_1^c with one unit of M^c in Stage 1. If the expected underage cost is smaller than the material expiry cost ($c_m - s_m$), it is not profitable to prepare one more unit of M^c in Stage 1. Therefore, if Equations (5.6) and (5.7) are not satisfied, no raw material is needed to prepare for the second production when $p_1^c = p_{s1}^{c*}$. In this situation, the optimal policy is to produce p_{s1}^{c*} only by using the speculative production, and the reactive production is worthless.

On the other hand, if Equation (5.6) is satisfied, then the optimal early production quantity p_1^{c*} and amount of the raw material M^{c*} can be obtained by solving Equations (5.8) and (5.9). It implies that p_1^{c*} is chosen to balance the expected overage cost and production delay premium, and that M^{c*} is chosen to balance the expected underage cost and material expiry cost.

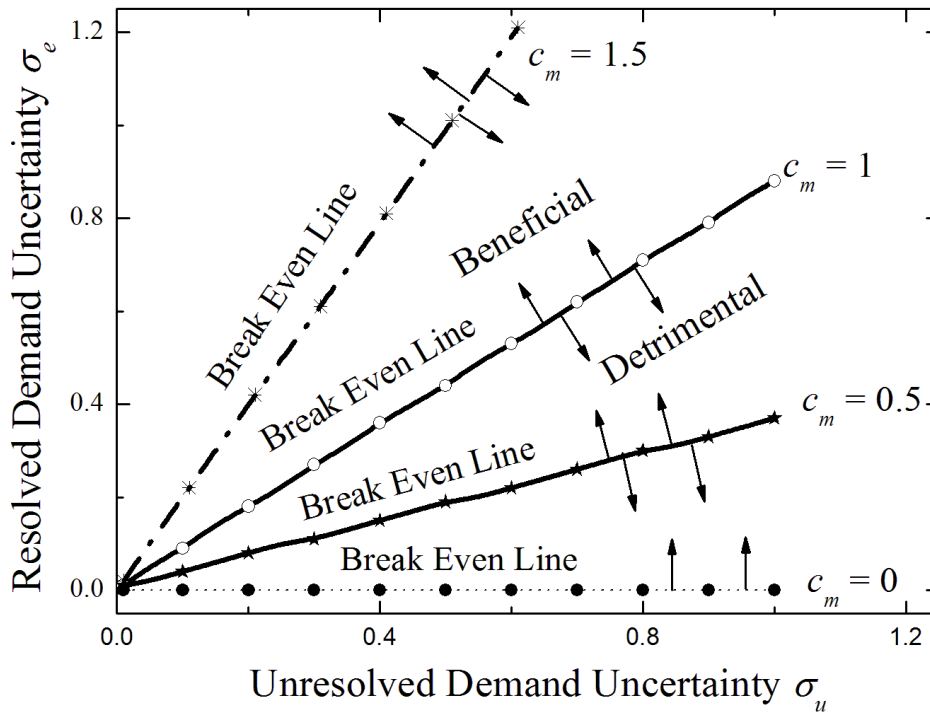


Figure 5.2 The break-Even Line in (σ_e, σ_u) Plane When $r = 10$, $c_1 = 3$, $c_2 = 3$, $c_m = 1$, $s_m = 0.05$, $s = 0.1$, $\pi = 0$, and $x_1 = 80$

In addition, an interesting finding emerged from Equations (5.6) and (5.7): the condition whether to add a reactive production is not only related to the costs, sales price and salvage values, but also related to the demand characteristics. To illustrate this, the following numerical examples are given.

We assume that the ratios of successive demand forecasts, X_2/X_1 and X/X_2 , are lognormally distributed with parameters $(-0.5\sigma_e^2, \sigma_e)$ and $(-0.5\sigma_u^2, \sigma_u)$ (a special case of martingale model of forecast evolution (R. G. Brown, 1959; Hausman, 1969; Heath & Jackson, 1994)). Hence, $X|x_1$, $X_2|x_1$ and $X|x_2$ are lognormally distributed with parameters $(\ln x_1 - 0.5(\sigma_e^2 + \sigma_u^2), \sqrt{\sigma_e^2 + \sigma_u^2})$, $(\ln x_1 - 0.5\sigma_e^2, \sigma_e)$ and $(\ln x_2 - 0.5\sigma_u^2, \sigma_u)$,

respectively. Given the lognormal distributions, the demand forecasts x_1 and x_2 correspond exactly to the expected demand observed in Stages 1 and 2, respectively. Moreover, σ_e measures the resolved demand uncertainty by the updated demand forecast in Stage 2, while σ_u measures the unresolved one. The following parameter values are used if they are not specified in numerical examples: $r = 10$, $c_1 = 3$, $c_2 = 3$, $c_m = 1$, $s_m = 0.05$, $s = 0.1$, $\pi = 0$, $x_1 = 80$, $\sigma_e = 0.9$, and $\sigma_u = 0.2$.

Figure 5.2 plots the break-even line that segments (σ_e, σ_u) plane according to whether a reactive production is beneficial or worthless when $c_m = 1.5, 1, 0.5$ and 0 . For $c_m \neq 0$ (e.g., $1.5, 1$, or 0.5), if the unresolved demand uncertainty σ_u increases, it requires a larger resolved demand uncertainty σ_e to ensure that the reactive production is beneficial. On the other hand, when $c_m = 0$, a reactive production is always beneficial to the supply chain for different σ_e and σ_u . Therefore, the condition depends on the demand forecast evolution in our model with material preprocessing, but not in the model without material preprocessing (Donohue, 2000). Moreover, a higher material cost increases the slope of the break-even line and reduces the beneficial area. It implies that the material cost strengthens the dependence of the condition on the demand forecast evolution. Hence, if material must be prepared in advance, companies should be concerned on the relative magnitude of the resolved demand uncertainty compared to the unresolved demand uncertainty when deciding whether to add a reactive production.

5.2.2 Comparison with Single-Mode Production

In addition to the two-mode production strategy, the manufacturer can follow two extreme strategies. One is to make all the production in Stage 1 before more demand information is revealed. This model is defined as the purely speculative production. The other extreme is to delay all the production to take advantage of more demand information in Stage 2 while raw material is prepared in Stage 1. This model is referred to as the purely reactive production. The purely reactive production should be preferred if the delay production cost is low and demand forecast accuracy is improved significantly in Stage 2. Otherwise, the purely speculative production would be preferred. Compared with these two single-mode production models, the benefit of the two-mode production could be highlighted.

Lemma 5.2.

(a) *In the two-mode production, the optimal supply chain expected profit is*

$$\begin{aligned}
 v(x_1) = & (r + \pi - s) \left[\int_0^{x_t(p_1^{c*})} \int_0^{p_1^{c*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right. \\
 & + \int_{x_t(p_1^{c*})}^{x_t(p_1^{c*} + M^{c*})} \int_0^{F_2^{-1}\left(\frac{r + \pi - c_2 - s m}{r + \pi - s} | x_2\right)} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \\
 & \left. + \int_{x_t(p_1^{c*} + M^{c*})}^{\infty} \int_0^{p_1^{c*} + M^{c*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right] - \pi E(X|x_1). \quad (5.10)
 \end{aligned}$$

(b) *In the purely speculative production, the optimal production quantity p_{s1}^{c*} is written as*

$$p_{s1}^{c*} = F^{-1} \left(\frac{r + \pi - c_1}{r + \pi - s} | x_1 \right).$$

The optimal supply chain expected profit is

$$v_s(x_1) = (r + \pi - s) \int_0^{p_{s1}^{c*}} xf(x|x_1)dx - \pi E(X|x_1).$$

(c) In the purely reactive production, the optimal amount of raw material M_r^{c*} and production quantity p_{r2}^{c*} satisfy the following equations:

$$\int_{x_t(M_r^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(M_r^{c*}|x_2)]f_1(x_2|x_1)dx_2 = c_m - s_m; \tag{5.11}$$

$$p_{r2}^{c*} = \begin{cases} \bar{p}_{r2}^c & x_2 \leq x_t(M_r^{c*}) \\ M_r^{c*} & x_2 > x_t(M_r^{c*}) \end{cases},$$

where $\bar{p}_{r2}^c = F_2^{-1}((r + \pi - c_2 - s_m)/(r + \pi - s)|x_2)$. The optimal supply chain expected profit is

$$v_r(x_1) = (r + \pi - s) \left[\int_0^{x_t(M_r^{c*})} \int_0^{F_2^{-1}\left(\frac{r+\pi-c_2-s_m}{r+\pi-s}|x_2\right)} xf_2(x|x_2)dx f_1(x_2|x_1)dx_2 + \int_{x_t(M_r^{c*})}^{\infty} \int_0^{M_r^{c*}} xf_2(x|x_2)dx f_1(x_2|x_1)dx_2 \right] - \pi E(X|x_1).$$

As expected, the optimal solutions in the purely speculative production are newsvendor-type expressions with the distribution function $F(.|x_1)$. The purely reactive production problem is similar in spirit to the two-stage problem proposed by H. Chen, Y. F. Chen, C.-H. Chiu, T.-M. Choi, and S. Sethi (2010b). They analyzed the optimization problem in terms of reactive capacity in Stage 1, and the ordering quantity and sales price in Stage 2. Comparing these two single-mode production models with the proposed two-mode production model leads to the following theorem.

Theorem 5.2. *The outcomes of the comparison among the three models are*

(a) *If Equation (5.6) is satisfied, then*

$$p_1^{c*} < p_{s1}^* < p_1^{c*} + M^{c*} = M_r^{c*};$$

$$v(x_1) > \max[v_s(x_1), v_r(x_1)].$$

(b) *If Equation (5.6) is not satisfied, then*

$$p_1^{c*} = p_{s1}^{c*} > M_r^{c*} \text{ and } M^{c*} = 0;$$

$$v(x_1) = v_s(x_1) > v_r(x_1).$$

Theorem 5.2(a) shows that the profit under the two-mode production is the largest among the three models if Equation (5.6) is satisfied. The driver of the profit advantage can be analyzed by observing the production quantity and raw material amount in Stage 1. In the two-mode production, the early production quantity p_1^{c*} is smaller than p_{s1}^{c*} in the purely speculative production, while the maximum quantity of available products $A = p_1^{c*} + M^{c*}$ is larger than $A = p_{s1}^{c*}$. Note that A is the maximum supply in Stage 2 when the demand forecast is updated. Thus, A can be used to measure the fill rate. Compared with the purely speculative production, the two-mode production can effectively reduce the upfront investment cost due to $c_m < c_1$ while at the same time ensuring a higher fill rate. Compared with the purely reactive production, the early production quantity is larger in the two-mode production (i.e., $p_1^{c*} > 0$), while A is the same (i.e., $A = p_1^{c*} + M^{c*} = M_r^{c*}$). Thus, in contrast to the purely reactive production, the two-mode production can take efficient advantage of the cheap early production without the sacrifice of the fill rate. These

observations demonstrate the value of the two-mode production.

Theorem 5.2(b) states that it may be worthless to add a reactive production. Specifically, the advantage of the reactive production may be lost due to the high production delay premium $(c_2 + c_m - c_1)$ and low value of the updated demand forecast (see discussion in Theorem 5.1). Due to the high production delay premium, both the maximum quantity and corresponding profit in the purely reactive production could be lower than those in the purely speculative production (see Figure 5.3).

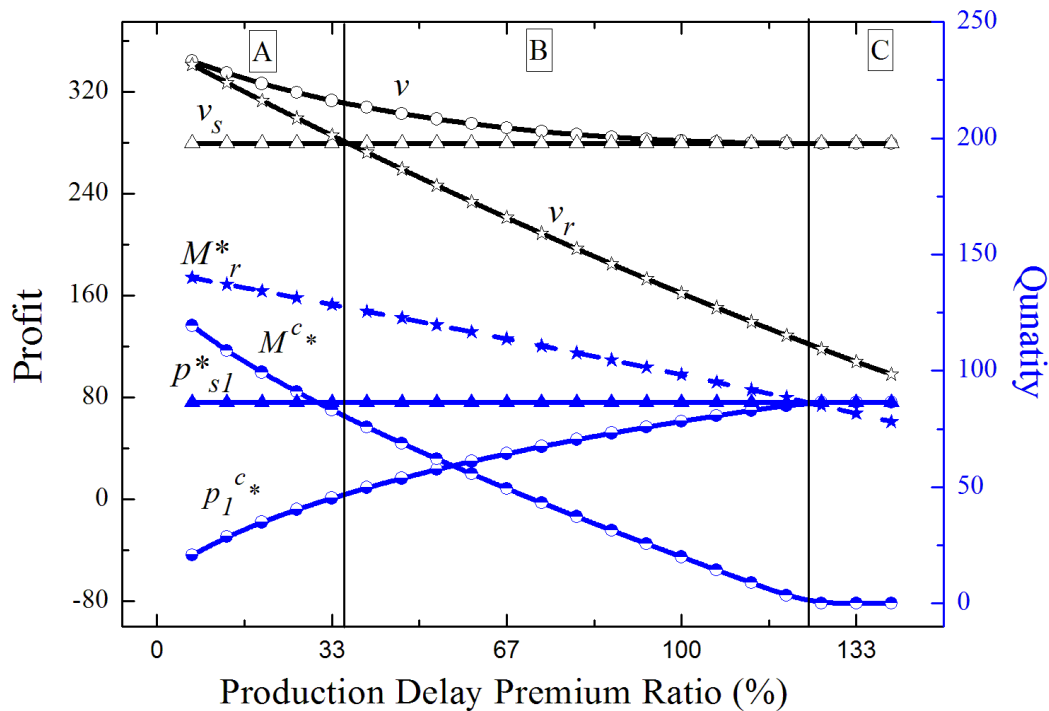


Figure 5.3 Profits and Quantities versus Production Delay Premium Ratio $(c_2 + c_m - c_1)/c_1$ for Three Models

To better illustrate Theorem 5.2, numerical examples are given. Figure 5.3 shows that the optimal production decisions and expected profits change with the production delay premium

ratio $(c_2 + c_m - c_1)/c_1$ in the three models. When the production delay premium ratio is not high (Areas A and B), the optimal expected profit in the two-mode production is the largest among the three production modes. This is because that the two-mode production allows taking advantage of both the cheap production mode and the cheap updated demand information. When the production delay premium ratio is very low (Area A), the purely reactive production is more profitable than the purely speculative production as improved demand information is acquired at low expense. When the production delay premium ratio is very high (Area C), the early production quantity p_1^{c*} in the two-mode production is equal to p_{s1}^* in the purely speculative production, and the material preparation M^{c*} becomes zero. Due to the high production delay premium, the reactive production could be harmful. This can also be seen from the lowest profit and fill rate $A = M_r^{c*} < p_{s1}^*$ in the purely reactive production. Thus, it is worthless to add a reactive production if the production delay premium ratio is very high.

5.2.3 Numerical Examples

To better understand the two-mode production, its advantages over the single-mode production are investigated in various settings. The benefit of the two-mode production over the purely speculative production (or the purely reactive production) is measured by $P_s = [(v - v_s)/v_s] \times 100\%$ (or $P_r = [(v - v_r)/v_r] \times 100\%$). The following results reveal the effects of the additional costs paid for the production delay, sales price, customer shortage sensitivity, material compatibility and demand uncertainties on the performance of the

two-mode production.

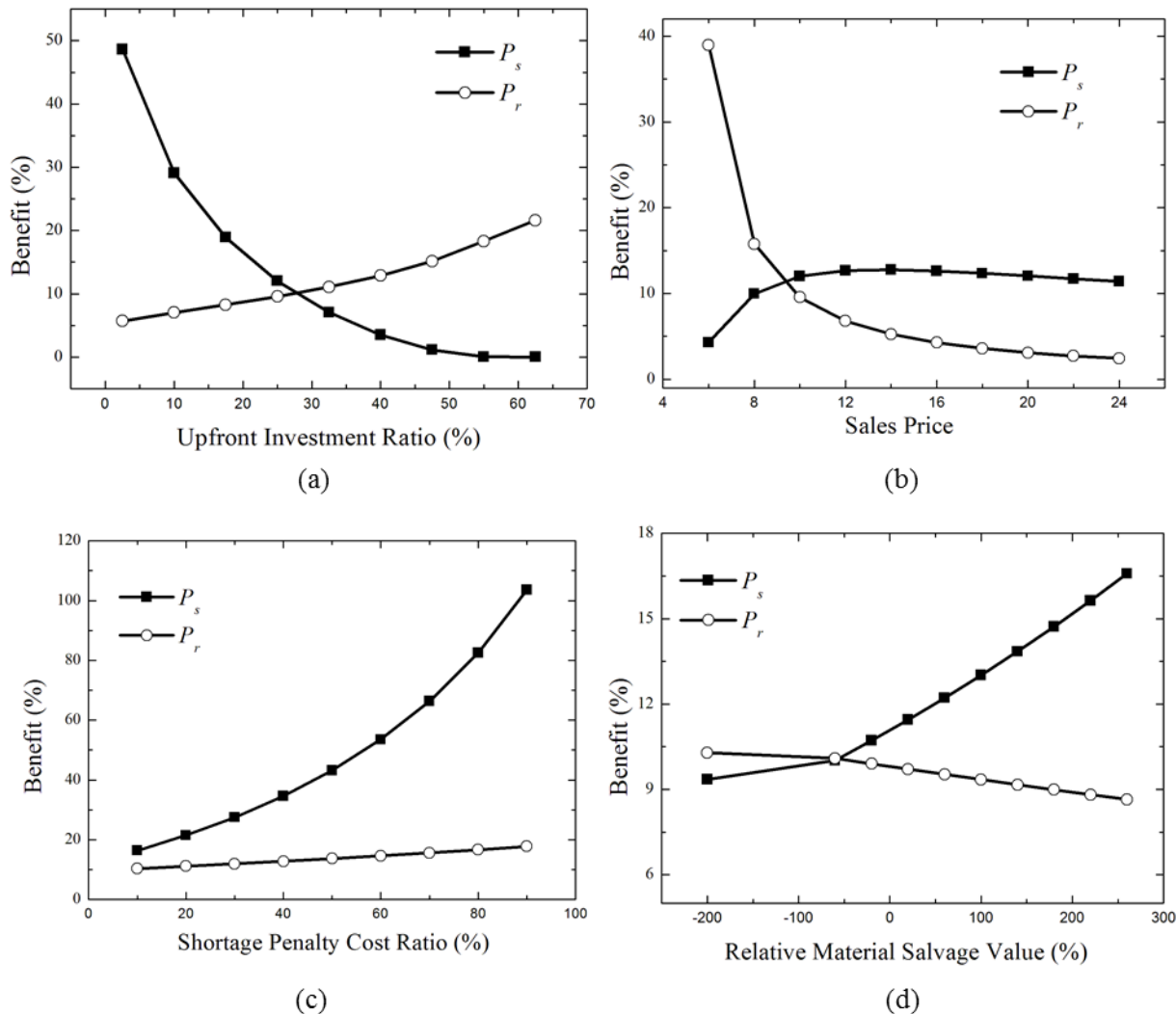


Figure 5.4 Effects of Upfront Investment Ratio c_m/c_2 , Sales Price r , Shortage Penalty Cost Ratio π/r and Relative Material Salvage Value V on Benefits P_s and P_r

Effects of Additional Costs of Production Delay: The delayed production is more expensive because of the upfront investment cost c_m in material and higher delayed production cost c_2 . Due to the increased costs of the delayed production, both of c_m and c_2

increase the early production quantity $p_1^{c^*}$, and decrease the amount of the material M^{c^*} and expected profit v in the two-mode production. As a result, the advantage of the two-mode production over the purely speculative production P_s decreases with c_m and c_2 . However, both of c_2 and c_m increase the advantage of the two-production over the purely reactive production P_r due to the higher production cost saving incurred by the early production.

To illustrate which one plays a more significant role, the effect of the upfront investment ratio c_m/c_2 when $c_m + c_2 = 4$ is shown in Figure 5.4(a). The direction of the effect of c_m/c_2 on the performance of the two-mode production is similar to c_m . It means that c_m plays a more significant role than c_2 . There are two reasons for this. First, if the material prepared for the delayed production is used, both c_m and c_2 would be paid; otherwise, only c_m would be paid. Second, the decision on the upfront investment (related to c_m) is made before the improved demand information; however, the decision on the reactive production (related to c_2) is made based on the improved demand information. Hence, a higher c_m has a more harmful impact on the two-mode production than a higher c_2 . Figure 5.4(a) also shows that the range of P_s changing with c_m/c_2 is larger than that of P_r . The benefit P_s lies on the low upfront investment cost before more accurate demand information is obtained, i.e., $c_m < c_1$, and thus it mainly depends on c_m . In contrast, the advantage P_r is caused by the cheap early production cost, i.e., $c_1 - c_m < c_2$, and thus it is affected by both c_m and c_2 . Therefore, the profit would be improved significantly from the speculative production to the two-mode production if the upfront investment cost c_m accounts for a small part or the delayed production cost c_2 accounts for a large part.

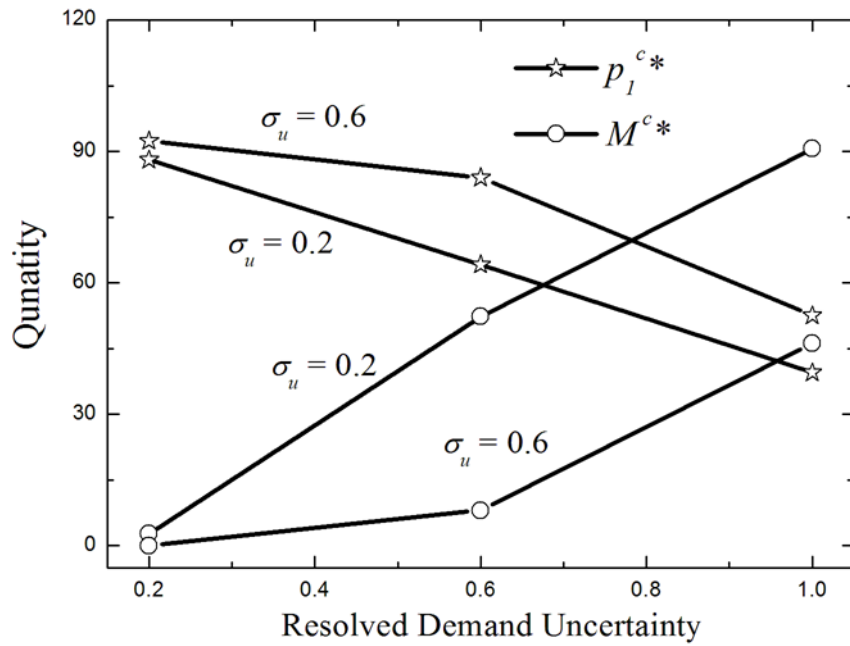
Effect of Sales Price: A higher sales price r encourages companies to make more production and investment in the material. Thus, the early production quantity and amount of the material increase with r in the two-mode production. Figure 5.4(b) shows that the advantage of the two-mode production P_s increases first and then decreases with r , compared to the purely speculative production. Note that a higher r leads to a higher underage cost. When r is not high, the value of the delayed production increases with r because the improved demand information enables companies to make the right quantity decision to reduce the higher underage cost. However, when r is very high, companies will simply produce more to hedge against the high underage cost instead of using the delayed production for the improved demand information. Hence, P_s becomes largest for products with a medium sales price. In addition, compared to the high price r , the production cost saving from the purely reactive production to the two-mode production becomes relatively small. Thus, the benefit of the two-mode production P_r is smaller for products with a higher sales price.

Effect of Customer Shortage Sensitivity: The shortage penalty cost π is used to measure the customer shortage sensitivity. Figure 5.4(c) illustrates the effect of the shortage penalty cost ratio π/r . When π/r is higher, the initial production quantity and amount of the material become larger to avoid the occurrence of shortage in the two-mode production. As shown in Figure 5.4(c), the benefits P_s and P_r of the two-mode production over the single-mode production increase with π/r . Compared to the purely speculative production, the higher fill rate A in the two-mode production (see discussion in Theorem 5.2(a)) can not

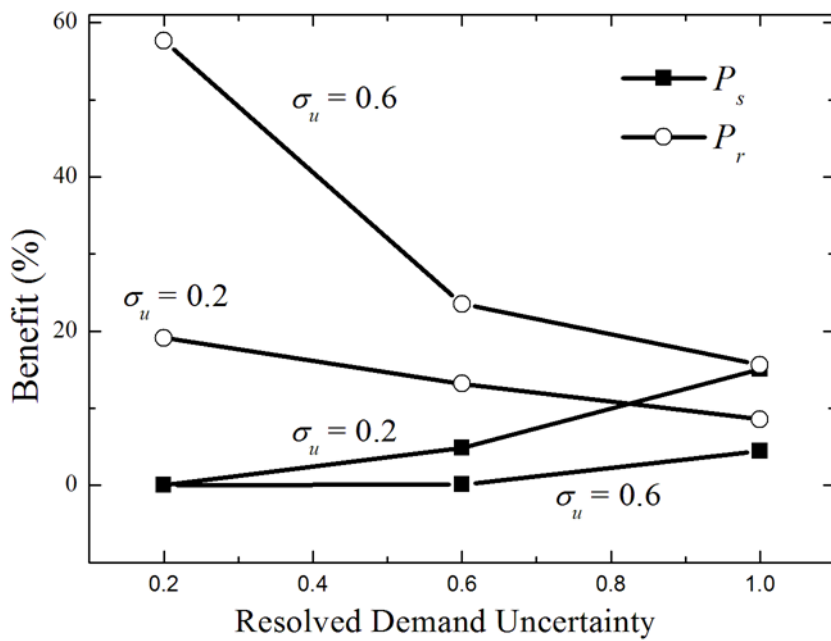
only reduce the total shortage penalty cost but also improve the incoming by selling more products. Hence, the benefit P_s increases more significantly as π/r increases. As shown in Figure 5.4(c), P_s could reach 104% when $\pi/r = 90\%$. In addition, a higher shortage penalty cost ratio increases the early production quantity, resulting in higher production cost saving from the purely reactive production to the two-mode production. Therefore, the two-mode production is more valuable than the single-mode production if stock-outs incur a higher cost.

Effect of Material Compatibility: The material compatibility could be measured by the relative material salvage value $V = (s_m - s)/|s| \times 100\%$. The salvage value could be negative, which refers to the deposable cost. Figure 5.4(d) illustrates how V impacts the two-mode production. If the material compatibility is higher, it is better to delay the conversion of the material to the finished product until demand information is updated. Thus, when V is higher, fewer products should be produced in Stage 1, and more material should be prepared for the later conversion, resulting in a decrease in p_1^{C*} and an increase in M^{C*} . The two-mode production allows the conversion delay, resulting in an increase in its benefit P_s over the purely speculative production if V increases. In the clinical trials supply chain, V could be high because the over-purchased raw material may be used for other drugs production while the overproduced trial drugs would be discarded. Hence, adding a reactive production would be beneficial to the clinical trials supply chain. In addition, due to the decreased early production quantity, the production cost saving becomes smaller from the purely reactive production to the two-mode production, leading to a decrease in the benefit

P_r if V is higher.



(a)



(b)

Figure 5.5 Effects of Resolved and Unresolved Demand Uncertainties σ_e and σ_u on

Quantities p_1^{c*} and M^{c*} and Benefits P_s and P_r

Effect of Demand Uncertainties: Figure 5.5(a) depicts that the early production quantity p_1^{c*} decreases and the amount of the material M^{c*} increases as the resolved demand uncertainty σ_e increases or the unresolved demand uncertainty σ_u decreases. If σ_e is higher or σ_u is lower, the proportion that can be resolved by the updated demand information becomes larger, implying that the value of acquiring the updated demand information is higher. Hence, a higher σ_e or lower σ_u decreases p_1^{c*} and increases M^{c*} in the two-mode production. Due to the higher value of the updated demand information, the benefit of the two-mode production becomes larger compared to the purely speculative production, if σ_e is higher or σ_u is lower. Due to the smaller early production quantity, the production cost saving from the purely reactive production to the two-mode production decreases if σ_e is higher or σ_u is lower. Therefore, a higher σ_e or lower σ_u increases the benefit P_s of the two-mode production over the purely speculative production, and decreases the benefit P_r over the purely reactive production, as shown in Figure 5.5(b).

These numerical results help to identify situations when the improvement is large if companies move from the single-mode production to the two-mode production. Table 5.1 summarizes how the parameters affect the benefit of the two-mode production. Compared to the purely speculative production, the benefit is more significant if the additional costs for production delay are low (in particular for the material cost), the sales price is medium, the customer shortage sensitivity is high, the material compatibility is high, the resolved demand uncertainty is high, and the unresolved demand uncertainty is low. Compared to the purely reactive production, the benefit is more dramatic if the additional costs for production delay

are high (in particular for the material cost), the sales price is low, the customer shortage sensitivity is high, the material compatibility is low, the resolved demand uncertainty is low, and the unresolved demand uncertainty is high.

Table 5.1 Summary-Production Strategies

		Low	Medium	High
Additional Costs for Production Delay	Material Cost c_m	S → T		R → T
	Delayed Production Cost c_2	S → T		R → T
	Upfront Investment Ratio c_m/c_2	S → T		R → T
Sales Price	Sales Price r	R → T	S → T	
Customer Shortage Sensitivity	Shortage Penalty Cost Ratio π/r			S → T
				R → T
Material Compatibility	Relative Material Salvage Value V	R → T		S → T
Demand Uncertainty	Resolved Demand Uncertainty σ_e	R → T		S → T
	Unresolved Demand Uncertainty σ_u	S → T		R → T

S → T means that the benefit is large if companies move from the purely speculative production to the two-mode production. R → T means that the benefit is large if companies move from the purely reactive production to the two-mode production.

5.3 Supply Chain Coordination

In this section, we assume that the manufacturer has a downstream buyer. In the decentralized system, the manufacturer and buyer take part in a Stackelberg Game where the manufacturer announces a contract first and the buyer makes ordering decisions based on the contract. A pricing contract with a return policy is proposed and optimized to achieve the supply chain coordination, i.e., the manufacturer's and buyer's decisions are consistent with the centralized optimal decisions. Under this contract, the allocation of the supply chain profit between the

two parties is also investigated to check whether both of the manufacturer and buyer can be better off.

5.3.1 Decentralized Decisions

A pricing contract with a return policy in the form of (c_R, l_1, l_2, s_r) , is proposed, where c_R is the reservation fee the buyer pays the manufacturer, l_1 and l_2 are the wholesale prices the manufacturer charges the buyer, and s_r is the return price the manufacturer pays the buyer for each unsold unit. Return policies are commonly used in the literature (Gerard P Cachon, 2003; H. Chen, et al., 2010b; Donohue, 2000; Pasternack, 1985). Under this contract, the decisions of the manufacture and buyer are shown in Figure 5.6. Given the contract and demand forecasts, the buyer makes ordering decisions to maximize his local expected profit. Based on the demand forecasts and buyer’s ordering quantities, the manufacturer optimizes his own production decisions. Similar to the aforementioned assumptions, to avoid the unreasonable cases, it is assumed that $r > c_R + l_2 > l_1$ and $l_1 - s_r > c_R$. It is also assumed that demand information is known to both the manufacturer and buyer.

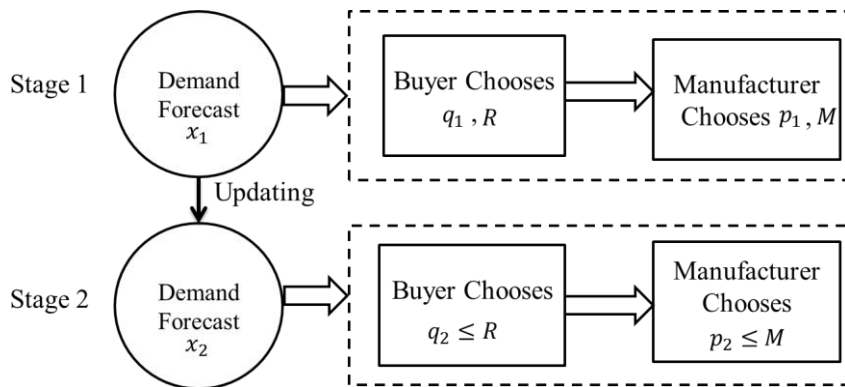


Figure 5.6 Manufacturer and Buyer’s Decisions under the Contract

5.3.2 Buyer

Under the pricing contract, the buyer should determine the first ordering quantity q_1 and reservation R in Stage 1, and the second ordering quantity $q_2 < R$ based on the improved demand information in Stage 2. The buyer's optimal total expected profit is given by

$$v^b(x_1) = \max\{w^b(q_1, R, x_1), q_1 \geq 0 \text{ and } R \geq 0\},$$

where

$$w^b(q_1, R, x_1) = -l_1 q_1 - c_R R + E_{F_1(\cdot|x_1)} [v_2^b(q_1, x_2)];$$

$$v_2^b(q_1, x_2) = \max\{w_2^b(q_1, q_2, x_2), 0 \leq q_2 \leq R\};$$

$$w_2^b(q_1, q_2, x_2) = -l_2 q_2 + r E_{F_2(\cdot|x_2)} [\min(q_1 + q_2, X)] - \pi E_{F_2(\cdot|x_2)} [(X - q_1 - q_2)^+] \\ + s_r E_{F_2(\cdot|x_2)} [(q_1 + q_2 - X)^+].$$

Compared with Equations (5.1) and (5.4), the above equations show that the buyer's problem is similar to the centralized problem with $s_m = 0$. Hence, the following lemma on the buyer's optimal ordering decisions is given.

Lemma 5.3.

(a) Given q_1, R and x_2 , the buyer's optimal ordering quantity q_2^* in Stage 2 is given by:

$$q_2^* = \begin{cases} 0 & x_2 < x_t^b(q_1) \\ \bar{q}_2(q_1) & x_t^b(q_1) \leq x_2 \leq x_t^b(q_1 + R), \\ R & x_2 > x_t^b(q_1 + R) \end{cases}$$

where $\bar{q}_2(q_1) = F_2^{-1}((r + \pi - l_2)/(r + \pi - s_r)|x_2) - q_1$, and $x_t^b(z)$ satisfies $F_2(z|x_t^b) = (r + \pi - l_2)/(r + \pi - s_r)$.

(b) Given x_1 , if

$$\int_0^{x_t^b(q_1^{n*})} [(r + \pi - s_r)F_2(q_1^{n*}|x_2) - (r + \pi - l_2)] f_1(x_2|x_1)dx_2 > (l_2 - l_1) + c_R, \quad (5.12)$$

where $q_1^{n*} = F^{-1}\left(\frac{r+\pi-l_1}{r+\pi-s_r} \middle| x_1\right)$, the buyer's optimal ordering quantity q_1^* and reservation

R^* in Stage 1 satisfy:

$$\int_0^{x_t^b(q_1^*)} [(r + \pi - s_r)F_2(q_1^*|x_2) - (r + \pi - l_2)] f_1(x_2|x_1)dx_2 = (l_2 - l_1) + c_R, \quad (5.13)$$

$$\int_{x_t^b(q_1^*+R^*)}^{\infty} [(r + \pi - l_2) - (r + \pi - s_r)F_2(q_1^* + R^*|x_2)] f_1(x_2|x_1)dx_2 = c_R. \quad (5.14)$$

Otherwise,

$$q_1^* = q_1^{n*}, \text{ and } R^* = 0. \quad (5.15)$$

5.3.3 Manufacturer

Based on the buyer's ordering decisions, the manufacturer needs to determine the optimal early production quantity p_1^* and prepare enough raw materials M^* to meet the possible maximum total orders, and produce enough products p_2^* to fill the total orders. Hence, M^* and p_2^* are set as $q_1 + R - p_1$ and $q_1 + q_2 - p_1$, and the decision concerning p_1^* becomes the key to characterizing the optimal production policy. Because the early production is cheaper, p_1^* should be not smaller than q_1 . Therefore, the manufacturer's optimal total expected profit function is

$$v^m(x_1) = \max\{w^m(p_1, x_1), q_1 \leq p_1 \leq q_1 + R\},$$

where

$$w^m(p_1, x_1) = w_1^m(p_1, x_1) + w_2^m(p_1, x_1) + w_{return}^m(x_1);$$

$$w_1^m(p_1, x_1) = c_R R + l_1 q_1 - c_1 p_1 - c_m (q_1 + R - p_1); \quad (5.16)$$

$$\begin{aligned} w_2^m(p_1, x_1) = & l_2 \left[\int_{x_t^b(q_1)}^{x_t^b(q_1+R)} \left(F_2^{-1} \left(\frac{r + \pi - l_2}{r + \pi - s_r} \middle| x_2 \right) - q_1 \right) f_1(x_2 | x_1) dx_2 \right. \\ & + \left. \int_{x_t^b(q_1+R)}^{\infty} R f_1(x_2 | x_1) dx_2 \right] \\ & + s_m \left[\int_0^{x_t^b(p_1)} (q_1 + R - p_1) f_1(x_2 | x_1) dx_2 \right. \\ & + \left. \int_{x_t^b(p_1)}^{x_t^b(q_1+R)} \left(q_1 + R - F_2^{-1} \left(\frac{r + \pi - l_2}{r + \pi - s_r} \middle| x_2 \right) \right) f_1(x_2 | x_1) dx_2 \right] \\ & + s \left[\int_0^{x_t^b(q_1)} (p_1 - q_1) f_1(x_2 | x_1) dx_2 \right. \\ & + \left. \int_{x_t^b(q_1)}^{x_t^b(p_1)} \left(p_1 - F_2^{-1} \left(\frac{r + \pi - l_2}{r + \pi - s_r} \middle| x_2 \right) \right) f_1(x_2 | x_1) dx_2 \right] \\ & - c_2 \left[\int_{x_t^b(p_1)}^{x_t^b(q_1+R)} \left(F_2^{-1} \left(\frac{r + \pi - l_2}{r + \pi - s_r} \middle| x_2 \right) - p_1 \right) f_1(x_2 | x_1) dx_2 \right. \\ & + \left. \int_{x_t^b(q_1+R)}^{\infty} (q_1 + R - p_1) f_1(x_2 | x_1) dx_2 \right]; \quad (5.17) \end{aligned}$$

$$\begin{aligned} w_{return}^m(x_1) = & -(s_r - s) \left[\int_0^{x_t^r(q_1)} \int_0^{q_1} (q_1 - x) f_2(x | x_2) dx f_1(x_2 | x_1) dx_2 \right. \\ & + \int_{x_t^r(q_1)}^{x_t^r(q_1+R)} \int_0^{F_2^{-1} \left(\frac{r + \pi - l_2}{r + \pi - s_r} \middle| x_2 \right)} \left(F_2^{-1} \left(\frac{r + \pi - l_2}{r + \pi - s_r} \middle| x_2 \right) \right. \\ & - \left. x \right) f_2(x | x_2) dx f_1(x_2 | x_1) dx_2 \\ & + \left. \int_{x_t^r(q_1+R)}^{\infty} \int_0^{q_1+R} (q_1 + R - x) f_2(x | x_2) dx f_1(x_2 | x_1) dx_2 \right]. \end{aligned}$$

$w_1^m(p_1, x_1)$ and $w_2^m(p_1, x_1)$ represent the expected profits generated in Stages 1 and 2,

respectively. In Stage 1, the manufacturer receives the revenue from the buyer's first order and reservation, and pays the costs for the early production and raw material preparation. In Stage 2, the manufacturer receives the revenue from the retailer's second order and salvage values of unused material and leftover finished products, and pays the delayed production cost. At the end of the selling season, the manufacturer needs to buy back unsold products from the buyer, resulting in the negative profit $w_{return}^m(x_1)$. Analysis on the total expected profit $w^m(p_1, x_1)$ leads to the following lemma on the optimal early production p_1^* .

Lemma 5.4. *Given q_1 , R and the policy for q_2 in Lemma 5.3(a), the manufacturer's optimal early production quantity is $p_1^* = \min(q_1 + R, \max(p_1^{n*}, q_1))$, where p_1^{n*} satisfies*

$$x_t^b(p_1^{n*}) = F_1^{-1}\left(\frac{c_2 - c_1 + c_m}{c_2 + s_m - s}\right). \quad (5.18)$$

Lemma 5.4 states that the manufacturer's early production decision not only responds directly to the buyer's first ordering decisions (q_1 , R) but also depends on the parameter values through p_1^{n*} . The left-hand side in Equation (5.18) represents the cost trade-off between the early production and delayed production, while the right-hand side $x_t^b(\cdot)$ is related to the parameters that impact on the buyer's second ordering quantity.

5.3.4 Design of Contract Prices

To coordinate the supply chain, the optimal production decision variables (p_1^* , M^* , p_2^*) in the decentralized system should be equal to (p_1^{c*} , M^{c*} , p_2^{c*}) in the centralized system. Since p_1^*

plays a key role in achieving the coordination as the values of other decision variables depend on it, the relationship between p_1^* and p_1^{c*} is firstly investigated in the following lemma.

Lemma 5.5. *The optimal early production quantity in the centralized system has the following property:*

$$p_1^{c*} > p_1^{n*}.$$

Based on Lemma 5.4 and Lemma 5.5, to achieve $p_1^* = p_1^{c*}$, the contract should be designed to incentivize the buyer to order $q_1^* = p_1^{c*}$. If $q_1^* = p_1^{c*}$, the manufacturer would produce $p_1^* = q_1^*$ and prepare $M^* = R^*$ in Stage 1, and produce $p_2^* = q_2^*$ in Stage 2. In this situation, if the contract prices are set to stimulate the buyer to reserve $R^* = M^{c*} > 0$, the manufacturer will prepare the same amount of the raw material as the optimal centralized solution, i.e., $M^* = M^{c*}$. Accordingly, an efficient contract that coordinates the supply chain is provided in the following theorem.

Theorem 5.3. *The supply chain achieves the coordination, when the contract prices (c_R, l_1, l_2, s_r) satisfy the following conditions:*

$$\begin{aligned}
 c_R &= \rho(c_m - s_m) \\
 l_1 &= (1 - \rho)(r + \pi) + \rho c_1 \\
 l_2 &= (1 - \rho)(r + \pi) + \rho(c_2 + s_m) \\
 s_r &= (1 - \rho)(r + \pi) + \rho s,
 \end{aligned}
 \tag{5.19}$$

where $0 \leq \rho \leq 1$.

The following findings can be obtained from Theorem 5.3:

1. Due to the variable ρ , the combination of the contract prices is not unique in ensuring the coordination, which is also shown in the contracts with return policies (Donohue, 2000; Pasternack, 1985). As ρ changes, the contract prices offered by the manufacturer to the buyer change, and thus the manufacturer's and buyer's profits can be flexible in the coordinated supply chain. Hence, ρ can be used by the manufacturer to control the share of the supply chain profit, and defined as the profit sharing factor.
2. The set of contract parameters is independent of demand characteristics. In addition, the coordinated conditions remain the same, even when demand information is unknown to the manufacturer (who makes his production decisions only based on the buyer's orders). This observation is different from that in Donohue (2000), where the higher predictive demand information (in Stage 2) would provide more flexibility for the manufacturer in designing the coordination contract parameters. It is because that the reservation quantity from the buyer provides additional demand information for the manufacturer to coordinate the supply chain in our model. It has been shown that sharing information about backlog or ordering policies from the downstream helps to coordinate the supply chain and thus improves the profit (Anderson, Morrice, & Lundeen, 2006; Gavirneni, et al., 1999).

3. Each wholesale (or return) price is the weighted average of the sales price, salvage value and costs that the manufacturer gains and pays directly, while the reservation fee depends only on the characteristics of the raw material. The wholesale and return prices decrease with ρ , while the reservation fee increases with ρ (due to the inequalities $0 \leq \rho \leq 1$ and $r > c_2 > c_1 > s$). Hence, the manufacturer would provide either lower wholesale prices or a lower reservation fee and a higher return price to induce the buyer to order the same as the optimal coordinated solution.

5.3.5 Profit Allocation

Besides aligning the two parties to maximize the supply chain profit, the split of the profit plays a key role in designing an efficient contract. If one party only makes a small profit under the coordinated contract, he would be unwilling to participate. Thus, it is necessary to investigate the manufacturer's and buyer's profits under the coordinated conditions, which are specified in the following theorem.

Theorem 5.4. *The manufacturer's and buyer's profits in the coordinated supply chain are*

$$v^m(x_1) = (1 - \rho)(v(x_1) + \pi E(X|x_1)),$$

$$v^b(x_1) = \rho(v(x_1) + \pi E(X|x_1)) - \pi E(X|x_1).$$

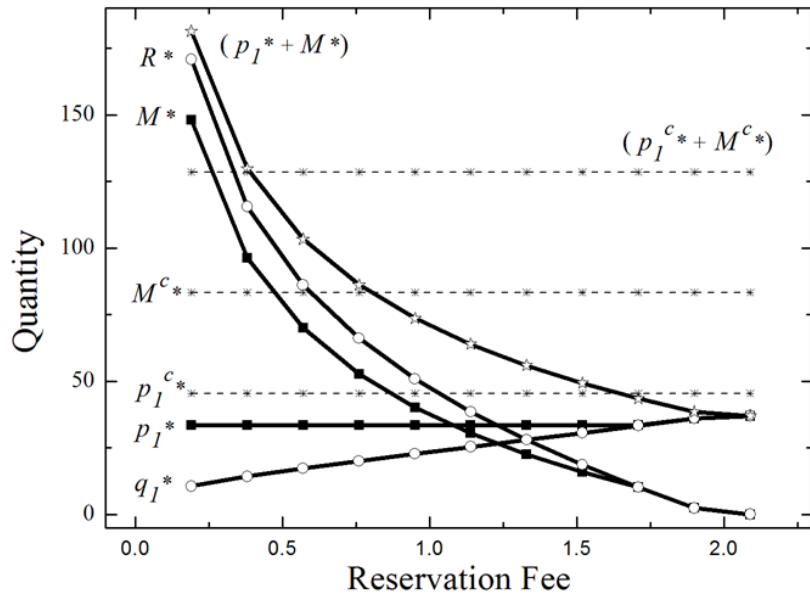
Theorem 5.4 shows that the manufacturer's profit decreases and the buyer's profit increases with the profit sharing factor ρ . Hence, the manufacturer can increase the share of the supply chain profit and exact all of the profit by decreasing ρ under the coordinated contract. In other words, the contract enables any division of the supply chain profit between

two parties, which is also found in other contracts with return policies (Gerard P Cachon, 2003; H. Chen, et al., 2010b; Donohue, 2000; Pasternack, 1985). Therefore, this contract can be used by the manufacturer to achieve Pareto improvement, which means that both the manufacturer and buyer are strictly better off.

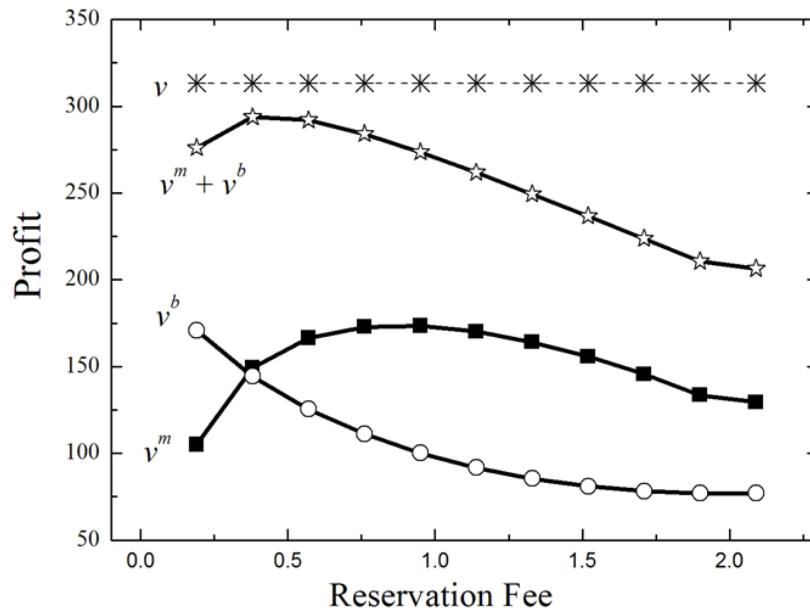
5.3.6 Numerical Illustration

To facilitate numerical study and emphasize on the effect of reservation, the following contract parameters are assumed and used in the decentralized system: $l_1 = 0.5r + 0.5c_1$, $l_2 = 0.5r + 0.5c_2$, and $s_r = s$. The reservation fee c_R could be set by the manufacturer. Without loss of generality, other parameter values are considered as stated in Section 5.2. Figure 5.7 shows how the optimal quantity decisions and expected profits of the manufacturer and buyer change with the reservation fee. The numerical results lead to the following observations about the behaviors of the buyer, manufacturer and channel.

Buyer: As the reservation fee c_R increases, the buyer's first ordering quantity q_1^* increases, and the reservation R^* and expected profit v^b decrease. When c_R is higher, the buyer would order more and reserve less (even nothing) in Stage 1. Furthermore, the decreasing rate of R^* is greater than the increasing rate of q_1^* since the cost paid for R^* is smaller than that for q_1^* (i.e., $c_R < l_1$). Hence, the maximum total ordering quantity $A = q_1^* + R^*$ decreases with c_R . It implies that a higher c_R may lead to a lower fill rate. Intuitively, the buyer's profit decreases with c_R since he needs to pay more for the reservation.



(a)



(b)

Figure 5.7 Behaviors of Buyer, Manufacturer and Channel with Reservation Fee c_R

Manufacturer: In response to the buyer's order behavior, the manufacturer's early production quantity p_1^* and material amount M^* also increases and decreases with c_R , respectively. Unlike the buyer's profit, the manufacturer's expected profit v^m first increases

and then decreases with c_R . When c_R is very low (or high), the manufacturer gains a small profit from the reservation due to the low reservation fee (or small possible total ordering quantity A from the buyer). Thus, the manufacturer would set an optimal reservation fee to maximize his profit.

Channel: Similar to the manufacturer's profit, the channel profit increases and then decreases with c_R . However, the optimal reservation fee from the perspective of the channel is different from the perspective of the manufacturer. As shown in Figure 5.7(b), the manufacturer would set a higher reservation fee due to double marginalization. In addition, Figure 5.7(a) shows that the buyer's initial ordering quantity is always smaller in the decentralized system than in the centralized system. In other words, the buyer is discouraged from ordering more in Stage 1 in the decentralized system. As a result, the manufacturer would not produce more through a cheap early production mode, leading to a lower channel profit in the decentralized system in Figure 5.7(b). Thus, the coordination conditions in Equation (5.19) stimulate the buyer to order more in Stage 1 and thus allow the manufacture to take full advantage of the cheap early production mode, in addition to avoiding double marginalization. The observation that the buyer and manufacturer produce less in the early stage in the decentralized system than in the coordinated system is similar to that in Donohue (2000). However, our numerical results also show that the buyer and manufacturer may prepare more for the later production (i.e., $M^* > M^{c^*}$ and $R^* > M^{c^*}$) if c_R is low. Due to low costs (i.e., $c_m < c_1$ and $c_R < l_1$), they may prefer over preparation for a possible high demand forecast in the later stage.

5.4 Conclusion

This study analyzes a supply chain with two-mode production, where a fast reactive production can be used to take advantage of improved demand information but the material for the production must be prepared in advance. It is demonstrated that adding this reactive production is not always beneficial to the supply chain. The condition under which the two-mode production outweighs the single-mode production is derived. The condition depends on the demand features, besides the costs, price and salvage values. Under the condition, the two-mode production has an advantage over the purely speculative production as it can not only decrease the upfront investment cost but also ensure a higher fill rate. Compared with the purely reactive production, the two-mode production is more beneficial because it allows the supply chain to enjoy the cheap early production without sacrificing the fill rate. Furthermore, a pricing contract with a return policy is proposed and optimized to coordinate the supply chain when the two-mode production is beneficial. Under this contract, the arbitrary split of the supply chain profit between the manufacturer and buyer makes Pareto improvement possible, and the manufacturer can acquire more demand information based on the buyer's reservation. These analytical results provide guidance for industrial companies. For instance, the condition Equation (5.6) is a good reference for companies to determine whether to add the reactive production. In addition, the proposed pricing contract could be an efficient contract for companies to coordinate with a downstream buyer (if it exists) as it allows maximizing the supply chain profit and allocating the profit arbitrarily.

Numerical examples reveal that if the supply chain moves from purely speculative

production, the benefit of the two-mode production is large for the low production delay costs (especially low material cost), medium sales price, high customer shortage sensitivity, high material compatibility, high resolved demand uncertainty or low unresolved demand uncertainty. If the supply chain moves from the purely reactive production, the benefit of the two-mode production is large for the high production delay costs (especially high material cost), low sales price, high customer shortage sensitivity, low material compatibility, low resolved demand uncertainty or high unresolved demand uncertainty. These numerical results enable companies to better understand when they can benefit more from the two-mode production. For example, companies who face a high shortage cost (or customer shortage sensitivity), e.g., GSK clinical trials, are encouraged to use the two-mode production if the single-mode production is used currently.

5.5 Appendix

Proof of Lemma 5.1.

$$\begin{aligned} \frac{\partial \bar{w}(p^c, x_2)}{\partial p^c} &= - \int_0^{p^c} (c_2 + s_m - s) f_2(x|x_2) dx \\ &\quad + \int_{p^c}^{\infty} (r + \pi - c_2 - s_m) f_2(x|x_2) dx; \\ \frac{\partial^2 \bar{w}(p^c, x_2)}{\partial^2 p^c} &= -(r + \pi - s) f_2(p^c|x_2) < 0. \end{aligned} \tag{5.20}$$

Note that $\bar{w}(p^c, x_2)$ is concave in p^c . According to Equation (5.2), $w_2(p_1^c, p_2^c, x_2)$ is concave in p_2^c . Hence, if there is no constraint, the optimal production quantity $p_2^{c*} = \bar{p}_2^c$, which is a newsvendor-type expression (i.e., $\bar{p}_2^c(p_1^c) = F_2^{-1}((r + \pi - c_2 - s_m)/(r + \pi - s)|x_2) - p_1^c$). If $0 \leq p_2^c \leq M^c$, p_2^{c*} is given by

$$p_2^{c*} = \begin{cases} 0 & \bar{p}_2^c < 0 \\ \bar{p}_2^c & 0 \leq \bar{p}_2^c \leq M^c \\ M^c & \bar{p}_2^c > M^c \end{cases} \quad (5.21)$$

Because $F_2^{-1}(\cdot|x_2)$ is a differentiable and strictly increasing function with x_2 , there exist threshold values of $x_t(p_1^c)$ and $x_t(p_1^c + M^c)$ satisfying $F_2(p_1^c|x_t) = (r + \pi - c_2 - s_m)/(r + \pi - s)$ and $F_2(p_1^c + M^c|x_t) = (r + \pi - c_2 - s_m)/(r + \pi - s)$ such that Equation (5.21) can be rewritten as Equation (5.3). Q.E.D.

Proof of Theorem 5.1.

Part (1). According to Equation (5.5) and Lemma 5.1, using Leibniz's rule, the first derivative of $w(p_1^c, M^c, x_1)$ in relation to p_1^c is as follows:

$$\begin{aligned} \frac{\partial w(p_1^c, M^c, x_1)}{\partial p_1^c} &= -c_1 + (w_2(p_1^c, 0, x_2) - w_2(p_1^c, M^c, x_2))f_1(x_t(p_1^c)|x_1)x'_t(p_1^c) \\ &\quad + (w_2(p_1^c, M^c, x_2) - w_2(p_1^c, M^c, x_2))f_1(x_t(p_1^c + M^c)|x_1)x'_t(p_1^c + M^c) \\ &\quad + \int_0^{x_t(p_1^c)} \frac{\partial w_2(p_1^c, 0, x_2)}{\partial p_1^c} f_1(x_2|x_1)dx_2 \\ &\quad + \int_{x_t(p_1^c)}^{x_t(p_1^c + M^c)} \frac{\partial w_2(p_1^c, \bar{p}_2^c(p_1^c), x_2)}{\partial p_1^c} f_1(x_2|x_1)dx_2 \\ &\quad + \int_{x_t(p_1^c + M^c)}^{\infty} \frac{\partial w_2(p_1^c, M^c, x_2)}{\partial p_1^c} f_1(x_2|x_1)dx_2. \end{aligned} \quad (5.22)$$

From Equation (5.2) and Lemma 5.1, it is seen that

$$\begin{aligned} \frac{\partial w_2(p_1^c, 0, x_2)}{\partial p_1^c} &= (c_2 + s_m) + \frac{\partial \bar{w}(p_1^c, x_2)}{\partial p_1^c} \\ &= (c_2 + s_m) + (r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^c|x_2); \end{aligned} \quad (5.23)$$

$$\frac{\partial EP_2(p_1^c, M^c, x_2)}{\partial p_1^c} = (c_2 + s_m) + \frac{\partial \bar{w}(p_1^c + M^c, x_2)}{\partial p_1^c}$$

$$= (c_2 + s_m) + (r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^c + M^c | x_2); \quad (5.24)$$

$$\begin{aligned} \frac{\partial EP_2(p_1^c, \bar{p}_2^c(p_1^c), x_2)}{\partial p_1^c} &= (c_2 + s_m) + \frac{\partial \bar{w}(p_1^c + \bar{p}_2^c(p_1^c), x_2)}{\partial p_1^c} \\ &= (c_2 + s_m); \end{aligned} \quad (5.25)$$

Substituting Equations (5.23), (5.24) and (5.25) into Equation (5.22) yields

$$\begin{aligned} \frac{\partial w(p_1^c, M^c, x_1)}{\partial p_1^c} &= (c_2 + s_m) - c_1 + \int_0^{x_t(p_1^c)} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^c | x_2)] f_1(x_2 | x_1) dx_2 \\ &+ \int_{x_t(p_1^c + M^c)}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^c + M^c | x_2)] f_1(x_2 | x_1) dx_2; \end{aligned} \quad (5.26)$$

In a similar way, the first derivative of $w(p_1^c, M^c, x_1)$ in relation to M^c is as follows:

$$\begin{aligned} \frac{\partial w(p_1^c, M^c, x_1)}{\partial M^c} &= -c_m + s_m + (w_2(p_1^c, M^c, x_2) - w_2(p_1^c, M^c, x_2)) f_1(x_t(p_1^c + M^c) | x_1) x'_t(p_1^c + M^c) \\ &+ \int_{x_t(p_1^c + M^c)}^{\infty} [(r + \pi - c_2) - (r + \pi - s)F_2(p_1^c + M^c | x_2)] f_1(x_2 | x_1) dx_2 \\ &= -c_m + s_m \\ &+ \int_{x_t(p_1^c + M^c)}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^c + M^c | x_2)] f_1(x_2 | x_1) dx_2. \end{aligned} \quad (5.27)$$

Note that $F_2(z | x_t(z)) = (r + \pi - c_2 - s_m) / (r + \pi - s)$. The second derivatives are as

follows:

$$\frac{\partial^2 w(p_1^c, M^c, x_1)}{\partial M^c \partial p_1^c} = - \int_{x_t(p_1^c + M^c)}^{\infty} (r + \pi - s) f_2(p_1^c + M^c | x_2) f_1(x_2 | x_1) dx_2 < 0;$$

$$\frac{\partial^2 w(p_1^c, M^c, x_1)}{\partial p_1^c \partial M^c} = - \int_{x_t(p_1^c + M^c)}^{\infty} (r + \pi - s) f_2(p_1^c + M^c | x_2) f_1(x_2 | x_1) dx_2 < 0;$$

$$\begin{aligned} & \frac{\partial^2 w_1(p_1^c, M^c, x_1)}{\partial^2 p_1^c} \\ &= - \int_0^{x_t(p_1^c)} (r + \pi - s) f_2(p_1^c | x_2) f_1(x_2 | x_1) dx_2 \\ & \quad - \int_{x_t(p_1^c + M^c)}^{\infty} (r + \pi - s) f_2(p_1^c + M^c | x_2) f_1(x_2 | x_1) dx_2 \\ & < 0; \end{aligned}$$

$$\frac{\partial^2 w(p_1^c, M^c, x_1)}{\partial^2 M^c} = - \int_{x_t(p_1^c + M^c)}^{\infty} (r + \pi - s) f_2(p_1^c + M^c | x_2) f_1(x_2 | x_1) dx_2 < 0.$$

Then,

$$\begin{aligned} & \frac{\partial^2 w(p_1^c, M^c, x_1)}{\partial^2 p_1^c} \frac{\partial^2 w(p_1^c, M^c, x_1)}{\partial^2 M^c} - \frac{\partial^2 w(p_1^c, M^c, x_1)}{\partial M^c \partial p_1^c} \frac{\partial^2 w(p_1^c, M^c, x_1)}{\partial p_1^c \partial M^c} \\ &= (r + \pi - s)^2 \int_0^{x_t(p_1^c)} f_2(p_1^c | x_2) f_1(x_2 | x_1) dx_2 \int_{x_t(p_1^c + M^c)}^{\infty} f_2(p_1^c + M^c | x_2) f_1(x_2 | x_1) dx_2 > 0. \end{aligned}$$

Therefore, the Hessian matrix is negative definite, and $w(p_1^c, M^c, x_1)$ is jointly concave in p_1^c and M^c .

Parts (2). Assuming that M^c is unconstrained in the two-mode production (i.e., M^c could be negative and non-negative). Since $w(p_1^c, M^c, x_1)$ is jointly concave in p_1^c and M^c , the optimal solutions p_1^{c*} and M^{c*} are obtained by making Equations (5.26) and (5.27) equal to zero. Namely, p_1^{c*} and M^{c*} satisfy

$$\int_0^{x_t(p_1'^{c*})} [(r + \pi - s)F_2(p_1'^{c*}|x_2) - (r + \pi - c_2 - s_m)] f_1(x_2|x_1) dx_2 \quad (5.28)$$

$$= (c_2 - c_1) + c_m,$$

$$\int_{x_t(p_1'^{c*} + M'^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*} + M'^{c*}|x_2)] f_1(x_2|x_1) dx_2 = c_m - s_m. \quad (5.29)$$

Note that the threshold value $x_t(\cdot)$ and demand distribution $F_2(\cdot|x_2)$ are increasing functions, and demand forecasts are non-negative. (The optimal solutions of $p_1'^{c*}$ and $(p_1'^{c*} + M'^{c*})$ by solving (5.28) and (5.29) are non-negative. Therefore, $M'^{c*} < 0$ is discussed in this study, without considering $p_1'^{c*} < 0$).

Firstly, we prove the sufficiency of the condition Equation (5.6). Assume that Equation (5.6) is true, i.e.,

$$\int_0^{x_t(p_{s1}^{c*})} [(r + \pi - s)F_2(p_{s1}^{c*}|x_2) - (r + \pi - c_2 - s_m)] f_1(x_2|x_1) dx_2 \quad (5.30)$$

$$> (c_2 - c_1) + c_m.$$

Then, according to Equation (5.28),

$$p_1'^{c*} < p_{s1}^{c*} = F^{-1}\left(\frac{r + \pi - c_1}{r + \pi - s} \middle| x_1\right).$$

Hence,

$$(r + \pi - c_1) + (c_m - s_m) - (r + \pi - s)F(p_1'^{c*}|x_1) > c_m - s_m. \quad (5.31)$$

Note that

$$(r + \pi - c_1) + (c_m - s_m) - (r + \pi - s)F(p_1'^{c*}|x_1)$$

$$= (r + \pi - c_2 - s_m) + [(c_2 - c_1) + c_m] - \int_0^{\infty} (r + \pi - s)F_2(p_1'^{c*}|x_2) f_1(x_2|x_1) dx_2$$

$$\begin{aligned}
 &= \int_0^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*}|x_2)]f_1(x_2|x_1)dx_2 \\
 &\quad - \int_0^{x_t(p_1'^{c*})} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*}|x_2)]f_1(x_2|x_1)dx_2 \\
 &= \int_{x_t(p_1'^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*}|x_2)]f_1(x_2|x_1)dx_2,
 \end{aligned}$$

where the last second equality is established by Equation (5.28). Thus, Equation (5.31) could be rewritten as

$$\int_{x_t(p_1'^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*}|x_2)]f_1(x_2|x_1)dx_2 > c_m - s_m.$$

Comparing the above equation with Equation (5.29), we can derive that

$$p_1'^{c*} < p_1'^{c*} + M^{c*}.$$

Therefore, if Equation (5.6) is true, $M^{c*} > 0$.

Then, we turn our attention to the necessity of the condition Equation (5.6). $M^{c*} > 0$ implies that $p_1'^{c*} + M^{c*} > p_1'^{c*}$. Hence, according to Equation (5.29)

$$\begin{aligned}
 c_m - s_m &< \int_{x_t(p_1'^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*}|x_2)]f_1(x_2|x_1)dx_2 \\
 &= \int_0^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*}|x_2)]f_1(x_2|x_1)dx_2 \\
 &\quad - \int_0^{x_t(p_1'^{c*})} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1'^{c*}|x_2)]f_1(x_2|x_1)dx_2 \\
 &= (r + \pi - c_2 - s_m) + (c_2 - c_1) + c_m - \int_0^{\infty} (r + \pi - s)F_2(p_1'^{c*}|x_2)f_1(x_2|x_1)dx_2 \\
 &= (r + \pi - c_1) + (c_m - s_m) - (r + \pi - s)F(p_1'^{c*}|x_1), \tag{5.32}
 \end{aligned}$$

where and the second equality is obtained according to Equation (5.28). The inequality (5.32) can be rewritten as

$$p_1^{c*} < F^{-1} \left(\frac{r + \pi - c_1}{r + \pi - s} \middle| x_1 \right) = p_{s1}^{c*}.$$

Based on Equation (5.28), this condition can be rewritten as Equation (5.30), i.e., Equation (5.6). Therefore, if $M^{c*} > 0$, then Equation (5.6) is true.

Note that

$$\begin{aligned} & \int_0^{x_t(p_{s1}^{c*})} [(r + \pi - s)F_2(p_{s1}^{c*}|x_2) - (r + \pi - c_2 - s_m)] f_1(x_2|x_1) dx_2 \\ &= (r + \pi - c_1) - (r + \pi - c_2 - s_m) \\ &+ \int_{x_t(p_{s1}^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*}|x_2)] f_1(x_2|x_1) dx_2. \end{aligned}$$

Hence, the following inequality can be derived from the condition Equation (5.30):

$$\int_{x_t(p_{s1}^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*}|x_2)] f_1(x_2|x_1) dx_2 > c_m - s_m. \quad (5.33)$$

From Equation (5.33), Equation (5.7) is obtained. Therefore, Equation (5.6) can be written as Equation (5.7).

Parts (3). If Equation (5.6) is satisfied, then $M^{c*} > 0$. Thus, the optimal production quantity $p_1^{c*} = p_1^{c*}$ and amount of raw material $M^{c*} = M^{c*}$. Thus, Equations (5.8) and (5.9) can be obtained based on Equations (5.28) and (5.29).

If Equation (5.6) is not satisfied, $M^{c*} < 0$. Because $w(p_1^c, M^c, x_1)$ is jointly concave in p_1^c and M^c , $\sup_{p_1^c} w(p_1^c, M^c, x_1)$ is concave in M^c . Therefore, if M^c is constrained (i.e., $M^c \geq 0$), $M^{c*} = 0$. If $M^{c*} = 0$, the optimal production quantity p_1^{c*} is the optimal solution of the traditional Newsvendor scenario where only the speculative production is allowed, i.e., $p_1^{c*} = p_{s1}^{c*}$. This completes the proofs. Q.E.D.

Proof of Lemma 5.2.

Part (a). The standard newsvendor form $\bar{w}(p^c, x_2)$ in Equation (5.2) can be rewritten as

$$\begin{aligned} \bar{w}(p^c, x_2) &= [r + \pi - (c_2 + s_m)]p^c - \pi \int_0^\infty x f_2(x|x_2) dx - (r + \pi \\ &\quad - s) \int_0^{p^c} (p^c - x) f_2(x|x_2) dx \\ &= p^c [(r + \pi - c_2 - s_m) - (r + \pi - s) F_2(p^c|x_2)] \\ &\quad - \pi \int_0^\infty x f_2(x|x_2) dx - (r + \pi - s) \int_0^{p^c} x f_2(x|x_2) dx. \end{aligned}$$

Hence, according to Equations (5.2) and (5.3) and Lemma 5.1,

$$\begin{aligned} w(p_1^c, M^c, x_1) &= (c_2 + s_m - c_1) p_1^c - (c_m - s_m) M^c \\ &\quad + E_{F_1(\cdot|x_1)} [\bar{w}(p_1^c + p_2^{c*}, x_2)]. \end{aligned} \tag{5.34}$$

Using Lemma 5.1,

If $x_2 < x_t(p_1^c)$,

$$\begin{aligned} \bar{w}(p_1^c + 0, x_2) &= p_1^c [(r + \pi - c_2 - s_m) - (r + \pi - s) F_2(p_1^c|x_2)] \\ &\quad + (r + \pi - s) \int_0^{p_1^c} x f_2(x|x_2) dx - \pi \int_0^\infty x f_2(x|x_2) dx; \end{aligned} \tag{5.35}$$

If $x_t(p_1^c) \leq x_2 \leq x_t(p_1^c + M^c)$,

$$\begin{aligned} \bar{w}\left(p_1^c + F_2^{-1}\left(\frac{r + \pi - c_2 - s_m}{r + \pi - s} \middle| x_2\right), x_2\right) &= F_2^{-1}\left(\frac{r + \pi - c_2 - s_m}{r + \pi - s} \middle| x_2\right) \\ &\quad \left[(r + \pi - c_2 - s_m) - (r + \pi - s) F_2\left(F_2^{-1}\left(\frac{r + \pi - c_2 - s_m}{r + \pi - s} \middle| x_2\right) \middle| x_2\right) \right] \\ &\quad + (r + \pi - s) \int_0^{F_2^{-1}\left(\frac{r + \pi - c_2 - s_m}{r + \pi - s} \middle| x_2\right)} x f_2(x|x_2) dx - \pi \int_0^\infty x f_2(x|x_2) dx \end{aligned}$$

$$= (r + \pi - s) \int_0^{F_2^{-1}\left(\frac{r+\pi-c_2-s_m}{r+\pi-s}|x_2\right)} x f_2(x|x_2) dx - \pi \int_0^\infty x f_2(x|x_2) dx; \quad (5.36)$$

If $x_2 > x_t(p_1^c + M^c)$,

$$\begin{aligned} \bar{w}(p_1^c + M^c, x_2) &= (p_1^c + M^c)[(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^c + M^c|x_2)] \\ &\quad + (r + \pi - s) \int_0^{p_1^c + M^c} x f_2(x|x_2) dx - \pi \int_0^\infty x f_2(x|x_2) dx. \end{aligned} \quad (5.37)$$

Hence, using Theorem 5.1-(3)

$$\begin{aligned} v_s(x_1) &= w(p_1^{c^*}, M^{c^*}, x_1) \\ &= (c_2 + s_m - c_1)p_1^{c^*} - (c_m - s_m)M^{c^*} - p_1^{c^*}(c_2 - c_1 + c_m) + (p_1^{c^*} + M^{c^*})(c_m - s_m) \\ &\quad + (r + \pi - s) \left[\int_0^{x_t(p_1^{c^*})} \int_0^{p_1^{c^*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right. \\ &\quad + \int_{x_t(p_1^{c^*})}^{x_t(p_1^{c^*} + M^{c^*})} \int_0^{F_2^{-1}\left(\frac{r+\pi-c_2-s_m}{r+\pi-s}|x_2\right)} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \\ &\quad + \left. \int_{x_t(p_1^{c^*} + M^{c^*})}^\infty \int_0^{p_1^{c^*} + M^{c^*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right] \\ &\quad - \pi \int_0^\infty \int_0^\infty x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \\ &= (r + \pi - s) \left[\int_0^{x_t(p_1^{c^*})} \int_0^{p_1^{c^*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right. \\ &\quad + \int_{x_t(p_1^{c^*})}^{x_t(p_1^{c^*} + M^{c^*})} \int_0^{F_2^{-1}\left(\frac{r+\pi-c_2-s_m}{r+\pi-s}|x_2\right)} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \\ &\quad + \left. \int_{x_t(p_1^{c^*} + M^{c^*})}^\infty \int_0^{p_1^{c^*} + M^{c^*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right] - \pi E(X|x_1). \end{aligned}$$

Part (b). Because the single speculative production mode is used, $M^c = 0$. Substituting

$M^c = 0$ and $p_1^{c^*} = p_{s1}^{c^*}$ into Equation (5.26) yields

$$(c_2 + s_m) - c_1$$

$$+ \int_0^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*}|x_2)] f_1(x_2|x_1) dx_2 = 0. \quad (5.38)$$

Rearranging terms leads to

$$\int_0^{\infty} F_2(p_{s1}^{c*}|x_2) f_1(x_2|x_1) dx_2 = F(p_{s1}^{c*}|x_1) = \frac{r + \pi - c_1}{r + \pi - s}.$$

Hence, by substituting $M^c = 0$ and $p_1^{c*} = p_{s1}^{c*}$ to Equations (5.35), (5.37) and (5.34),

$v_s(x_1)$ can be written as

$$\begin{aligned} v_s(x_1) &= (c_2 + s_m - c_1)p_{s1}^{c*} \\ &+ p_{s1}^{c*} \int_0^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*}|x_2)] f_1(x_2|x_1) dx_2 \\ &+ (r + \pi - s) \int_0^{\infty} \int_0^{p_{s1}^{c*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 - \pi E(X|x_1). \end{aligned}$$

According to Equation (5.38),

$$\begin{aligned} v_s(x_1) &= (r + \pi - s) \int_0^{\infty} \int_0^{p_{s1}^{c*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 - \pi E(X|x_1) \\ &= (r + \pi - s) \int_0^{p_{s1}^{c*}} x f(x|x_1) dx - \pi E(X|x_1). \end{aligned}$$

Part (c). Because the single reactive production mode is used, $p_1^c = 0$. Substituting $p_1^c = 0$ and $M^c = M_r^{c*}$ into Equation (5.27) and Lemma 5.1 yields

$$\int_{x_t(M_r^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(M_r^{c*}|x_2)] f_1(x_2|x_1) dx_2 = c_m - s_m. \quad (5.39)$$

$$p_{r2}^{c*} = \begin{cases} \bar{p}_{r2}^c & x_2 \leq x_t(M_r^{c*}) \\ M_r^{c*} & x_2 > x_t(M_r^{c*}) \end{cases}.$$

where $\bar{p}_{r2}^c = F_2^{-1}((r + \pi - c_2 - s_m)/(r + \pi - s)|x_2)$. By substituting $p_1^c = 0$ and $M^c = M_r^{c*}$ into Equations (5.36), (5.37) and (5.34), $v_r(x_1)$ can be written as

$$v_r(x_1) = -(c_m - s_m)M_r^{c*}$$

$$\begin{aligned}
 & + M_r^{c*} \int_{x_t(M_r^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(M_r^{c*}|x_2)]f_1(x_2|x_1)dx_2 \\
 & + (r + \pi - s) \left[\int_0^{x_t(M_r^{c*})} \int_0^{F_2^{-1}\left(\frac{r+\pi-c_2-s_m}{r+\pi-s}|x_2\right)} x f_2(x|x_2)dx f_1(x_2|x_1)dx_2 \right. \\
 & \left. + \int_{x_t(M_r^{c*})}^{\infty} \int_0^{M_r^{c*}} x f_2(x|x_2)dx f_1(x_2|x_1)dx_2 \right] - \pi E(X|x_1).
 \end{aligned}$$

According to Equation (5.39)

$$\begin{aligned}
 v_r(x_1) & = (r + \pi - s) \left[\int_0^{x_t(M_r^{c*})} \int_0^{F_2^{-1}\left(\frac{r+\pi-c_2-s_m}{r+\pi-s}|x_2\right)} x f_2(x|x_2)dx f_1(x_2|x_1)dx_2 \right. \\
 & \left. + \int_{x_t(M_r^{c*})}^{\infty} \int_0^{M_r^{c*}} x f_2(x|x_2)dx f_1(x_2|x_1)dx_2 \right] - \pi E(X|x_1).
 \end{aligned}$$

This completes the proofs.

Q.E.D.

Proof of Theorem 5.2.

Part (a). Based on Theorem 5.1-(2), if Equation (5.6) is satisfied, $M^{c*} > 0$. Substituting

$p_1^{c*} = p_{s1}^{c*}$ into Equation (5.26) yields

$$\begin{aligned}
 & \frac{\partial w(p_{s1}^{c*}, M^{c*}, x_1)}{\partial p_1^c} \\
 & = (c_2 + s_m) - c_1 + \int_0^{x_t(p_{s1}^{c*})} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*}|x_2)] f_1(x_2|x_1)dx_2 \\
 & + \int_{x_t(p_{s1}^{c*} + M^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*} + M^{c*}|x_2)] f_1(x_2|x_1)dx_2.
 \end{aligned}$$

According to Equation (5.38),

$$\begin{aligned}
 & \frac{\partial w(p_{s1}^{c*}, M^{c*}, x_1)}{\partial p_1^c} \\
 & = - \int_{x_t(p_{s1}^{c*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c*}|x_2)] f_1(x_2|x_1)dx_2
 \end{aligned}$$

$$\begin{aligned}
 & + \int_{x_t(p_{s1}^{c^*} + M^{c^*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c^*} + M^{c^*} | x_2)] f_1(x_2 | x_1) dx_2 \\
 & < 0,
 \end{aligned} \tag{5.40}$$

where the inequality is established since $p_{s1}^{c^*} + M^{c^*} > p_{s1}^{c^*}$ when $M^{c^*} > 0$. From Equation (5.26), it is noted that

$$\begin{aligned}
 & \frac{\partial w(p_1^{c^*}, M^{c^*}, x_1)}{\partial p_1^c} \\
 & = (c_2 + s_m) - c_1 + \int_0^{x_t(p_1^{c^*})} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^{c^*} | x_2)] f_1(x_2 | x_1) dx_2 \\
 & \quad + \int_{x_t(p_1^{c^*} + M^{c^*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_1^{c^*} + M^{c^*} | x_2)] f_1(x_2 | x_1) dx_2 \\
 & = 0.
 \end{aligned} \tag{5.41}$$

Therefore, $\frac{\partial w(p_{s1}^{c^*}, M^{c^*}, x_1)}{\partial p_1^c} < 0$, $\frac{\partial w(p_1^{c^*}, M^{c^*}, x_1)}{\partial p_1^c} = 0$. Because $w(p_1^c, M^c, x_1)$ is concave in p_1^c , $p_1^{c^*} < p_{s1}^{c^*}$. It implies that

$$\begin{aligned}
 & - \int_0^{x_t(p_{s1}^{c^*})} [(r + \pi - s)F_2(p_{s1}^{c^*} | x_2) - (r + \pi - c_2 - s_m)] f_1(x_2 | x_1) dx_2 \\
 & < - \int_0^{x_t(p_1^{c^*})} [(r + \pi - s)F_2(p_1^{c^*} | x_2) - (r + \pi - c_2 - s_m)] f_1(x_2 | x_1) dx_2.
 \end{aligned} \tag{5.42}$$

Hence, according to Equations (5.38) and (5.41),

$$\begin{aligned}
 & - \int_{x_t(p_{s1}^{c^*})}^{\infty} [(r + \pi - c_2 - s_m) - (r + \pi - s)F_2(p_{s1}^{c^*} | x_2)] f_1(x_2 | x_1) dx_2 \\
 & < - \int_{x_t(p_1^{c^*} + M^{c^*})}^{\infty} [(r + \pi - c_2 - s_m) \\
 & \quad - (r + \pi - s)F_2(p_1^{c^*} + M^{c^*} | x_2)] f_1(x_2 | x_1) dx_2.
 \end{aligned} \tag{5.43}$$

Hence, $p_{s1}^{c^*} < p_1^{c^*} + M^{c^*}$.

It is easy to obtain that $p_1^{c*} + M^{c*} = M_r^{c*}$ by comparing Equations (5.9) and (5.11).

Therefore,

$$p_1^{c*} < p_{s1}^* < p_1^{c*} + M^{c*} = M_r^{c*};$$

According to Theorem 5.1-(1), $w(p_1^c, M^c, x_1)$ is jointly concave in p_1^c and M^c . Thus, (p_1^{c*}, M^{c*}) is the global optimal point. It implies that

$$w(p_1^{c*}, M^{c*}, x_1) > w(0, M_r^{c*}, x_1);$$

$$w(p_1^{c*}, M^{c*}, x_1) > w(p_{s1}^{c*}, 0, x_1);$$

Hence,

$$v(x_1) > \max[v_r(x_1), v_s(x_1)].$$

Part (b). Based on Theorem 5.1-(2), if Equation (5.6) is not satisfied, $M'^{c*} \leq 0$. Hence, the direction of the inequality of (5.40) is reversed. Similar to the proof of Part (a), it is obtained that $p_1'^{c*} \geq p_{s1}^{c*}$. Due to $p_1'^{c*} \geq p_{s1}^{c*}$, the directions of the inequalities of (5.42) and (5.43) are reversed, resulting in $p_{s1}^{c*} \geq p_1'^{c*} + M'^{c*} = M_r^{c*}$. Hence, $p_{s1}^{c*} > M_r^{c*}$. According to the proof of Theorem 5.1-(3), if Equation (5.6) is not satisfied, and M^c is constrained (i.e., $M^c \geq 0$), $M^{c*} = 0$ and $p_1^{c*} = p_{s1}^{c*}$, where $w(p_1^c, M^c, x_1)$ is maximized. Consequently, if $M^c \geq 0$,

$$M_r^{c*} < p_{s1}^{c*} = p_1^{c*} = p_1^{c*} + M^{c*}$$

$$v(x_1) = v_s(x_1) > v_r(x_1).$$

This completes the proofs.

Q.E.D.

Proof of Lemma 5.3. The results in (a) and (b) are easily obtained by substituting $c_1 = l_1$,

$c_2 = l_2$, $s = s_r$, $s_m = 0$, $c_m = c_R$ into Lemma 5.1 and Theorem 5.1. Q.E.D.

Proof of Lemma 5.4. Note that $w_{return}^m(x_1)$ is independent of p_1 . According to Leibniz's rule and Lemma 5.3(a), from Equations (5.16) and (5.17), the first and second derivatives of $w^m(p_1, x_1)$ in relation to p_1 are as follows:

$$\begin{aligned} \frac{\partial w^m(p_1, x_1)}{\partial p_1} &= \frac{\partial w_1^m(p_1, x_1)}{\partial p_1} + \frac{\partial w_2^m(p_1, x_1)}{\partial p_1} \\ &= c_m - c_1 - s_m \int_0^{x_t^b(p_1)} f_1(x_2|x_1) dx_2 + s \int_0^{x_t^b(p_1)} f_1(x_2|x_1) dx_2 \\ &\quad + c_2 \int_{x_t^b(p_1)}^{\infty} f_1(x_2|x_1) dx_2; \\ \frac{\partial^2 w^m(p_1, x_1)}{\partial^2 p_1} &= -(c_2 + s_m - s)f_1(x_t^b(p_1)|x_1) < 0. \end{aligned}$$

Hence, $w^m(p_1, x_1)$ is concave in p_1 during $[q_1, q_1 + R]$. Under no constraint, the optimal production quantity p_1^{n*} is obtained by making $\frac{\partial w^m(p_1, x_1)}{\partial p_1}$ equal to zero. During $[q_1, q_1 + R]$, the optimal production quantity $p_1^* = \min(q_1 + R, \max(p_1^{n*}, q_1))$. Q.E.D.

Proof of Lemma 5.5. Substituting p_1^{n*} into the left hand side of Equation (5.8) yields

$$\begin{aligned} &\int_0^{x_t(p_1^{n*})} [(r + \pi - s)F_2(p_1^{n*}|x_2) - (r + \pi - c_2 - s_m)] f_1(x_2|x_1) dx_2 \\ &< \int_0^{x_t(p_1^{n*})} [(r + \pi - s) - (r + \pi - c_2 - s_m)] f_1(x_2|x_1) dx_2 \\ &= (c_2 + s_m - s)F_1(x_t^b(p_1^{n*})|x_1) \\ &= c_2 - c_1 + c_m \end{aligned}$$

$$= \int_0^{x_t(p_1^{c*})} [(r + \pi - s)F_2(p_1^{c*}|x_2) - (r + \pi - c_2 - s_m)] f_1(x_2|x_1) dx_2,$$

where the last two equalities are established based on Equations (5.18) and (5.8), respectively.

Due to the increasing functions of $x_t(\cdot)$ and $F_2(\cdot|x_2)$, $p_1^{c*} > p_1^{n*}$. Q.E.D.

Proof of Theorem 5.3. Substituting $q_1^* = p_1^{c*}$ and $R^* = M^{c*} > 0$ into Equations (5.13)

and (5.14) yields

$$\begin{aligned} & \int_0^{x_t^b(p_1^{c*})} [(r + \pi - s_r)F_2(p_1^{c*}|x_2) - (r + \pi - l_2)] f_1(x_2|x_1) dx_2 \\ & = (l_2 - l_1) + c_R, \end{aligned} \quad (5.44)$$

$$\int_{x_t^b(p_1^{c*} + M^{c*})}^{\infty} [(r + \pi - l_2) - (r + \pi - s_r)F_2(p_1^{c*} + M^{c*}|x_2)] f_1(x_2|x_1) dx_2 = c_R. \quad (5.45)$$

According to Equations (5.8), (5.9), (5.44) and (5.45), the relationships of the contract prices are obtained as follows:

$$(r + \pi - s_r) = \rho(r + \pi - s), \quad (5.46)$$

$$(r + \pi - l_2) = \rho(r + \pi - c_2 - s_m), \quad (5.47)$$

$$c_R = \rho(c_m - s_m), \quad (5.48)$$

$$(l_2 - l_1) + c_R = \rho[(c_2 - c_1) + c_m]. \quad (5.49)$$

Equations (5.46) and (5.47) show that $(r + \pi - c_2 - s_m)/(r + \pi - s) = (r + \pi - l_2)/(r + \pi - s_r)$. Comparing Lemma 5.1 and Lemma 5.3(a), $q_2^* = p_2^{c*}$ since $q_1^* = p_1^{c*}$ and $R^* = M^{c*}$. As previously noted, $p_1^* = q_1^*$, $M^* = R^*$ and $p_2^* = q_2^*$ when $q_1^* = p_1^{c*}$.

Consequently, $p_1^* = p_1^{c*} = q_1^*$, $M^* = M^{c*} = R^*$ and $p_2^* = p_2^{c*} = q_2^*$. The coordination

conditions are derived from Equations (5.46), (5.47), (5.48) and (5.49).

Q.E.D.

Proof of Theorem 5.4.

As noted previously, the buyer's problem resembles the centralized problem with $s_m = 0$.

Hence, similar to (5.10), the optimal buyer's profit is

$$\begin{aligned}
 v^b(x_1) = & (r + \pi - s_r) \left[\int_0^{x_t^b(q_1^*)} \int_0^{q_1^*} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right. \\
 & + \int_{x_t^b(q_1^*)}^{x_t^b(q_1^*+R^*)} \int_0^{F_2^{-1}\left(\frac{r+\pi-l_2}{r+\pi-s_r}|x_2\right)} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \\
 & + \left. \int_{x_t^b(q_1^*+R^*)}^{\infty} \int_0^{q_1^*+R^*} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right] - \pi E(X|x_1) \\
 & + \int_{x_t^b(q_1^*+R^*)}^{\infty} \int_0^{q_1^*+R^*} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 - \pi E(X|x_1). \tag{5.50}
 \end{aligned}$$

Substituting $q_1^* = p_1^{c*}$, $R^* = M^{c*}$ and Equation (5.19) into Equation (5.50) yields

$$\begin{aligned}
 v^b(x_1) = & \rho(r + \pi - s) \left[\int_0^{x_t(p_1^{c*})} \int_0^{p_1^{c*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right. \\
 & + \int_{x_t(p_1^{c*})}^{x_t(p_1^{c*}+M^{c*})} \int_0^{F_2^{-1}\left(\frac{r+\pi-c_2-s_m}{r+\pi-s}|x_2\right)} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \\
 & + \left. \int_{x_t(p_1^{c*}+M^{c*})}^{\infty} \int_0^{p_1^{c*}+M^{c*}} x f_2(x|x_2) dx f_1(x_2|x_1) dx_2 \right] - \pi E(X|x_1). \tag{5.51}
 \end{aligned}$$

Based on Equations (5.10) and (5.51), $v^b(x_1) = \rho(v(x_1) + \pi E(X|x_1)) - \pi E(X|x_1)$. Hence,

the manufacturer's profit is $v^m(x_1) = v(x_1) - v^b(x_1) = (1 - \rho)(v(x_1) + \pi E(X|x_1))$ in

the coordinated supply chain.

Q.E.D.

Chapter 6. Conclusions and Future Research

In the chapter, we summarize the main results and contributions of this thesis. In addition, we discuss limitations of this research and potential research opportunities for future work.

6.1 Summary of the Research

Scenario 1 investigates an extension of the conventional newsvendor model with multiple ordering opportunities under demand forecast updating and supply restrictions. The retailer prefers to postpone orders to improve demand accuracy. However, the manufacturer would charge the retailer a higher cost and limit the ordering time and quantity for a later order. These supply features have negative effects on the value of demand forecast updating. Therefore, a balance between demand forecast accuracy and supply constraints is investigated from the perspective of the retailer in supply modes A and B.

In supply mode A, the optimal ordering quantity depends on the current demand forecast; however, the current demand forecast does not affect the optimal ordering time, which is one of the two endpoints of the feasible timespan under justifiable assumptions. Hence, demand forecast updating is either valuable or valueless in supply mode A. In supply mode B, due to the ordering quantity restrictions, the optimal ordering time may depend on the demand forecast evolution and be complicated to derive. To approximate it, an algorithm is proposed by relaxing the ordering quantity restrictions appropriately.

Numerical examples show the effects of product and demand attributes on the value of demand forecast updating in supply modes A and B. Demand forecast updating is found to be

more valuable to the retailer in the two supply modes, particularly for (a) the more expensive original purchasing cost, (b) the smaller value overstock products, (c) the lower additional cost paid for lead time reduction and (d) the higher demand forecast efficiency. In addition, as the price decreases, the value of demand forecast updating keeps the same in supply mode A, and increases in most cases in supply mode B.

Different supply scenarios are compared through numerical examples. When demand uncertainty is low, the benefit of demand forecast updating decreases largely due to the increased purchasing cost. In a market with highly uncertain demand, the ordering time and quantity restrictions should be relaxed to take advantage of the demand forecast updating. In other words, a proactive retailer should find multiple suppliers and choose the shorter supply lead time when necessary. In addition, it is observed that supply mode B is more valuable to retailers than supply mode A if they face a risk-averse manufacturer, especially in a market with high demand uncertainty, and vice versa.

Scenario 2 analyzes the ordering decisions for a retailer who has a regular ordering opportunity and a more expensive emergency ordering opportunity. The emergency ordering quantity is limited due to the shorter supply lead time. Using dynamic programming, we derive the optimal ordering policy and identify the structural properties of the optimal solutions. Furthermore, we investigate how the regular ordering quantity and impact of the availability of the emergency order change with the price, salvage value, unit cost, maximum ordering quantity limit and demand characteristic. Numerical examples reveal that the availability of the emergency order is more critical to the retailer if (a) the initial demand

forecast is high under LND, (b) the demand uncertainty before the emergency order is high under ND, or (c) the demand uncertainty after the emergency order is low under both ND and LND.

Scenario 3 investigates a supply chain with two-mode production. A fast reactive production can be used to take advantage of improved demand information, in addition to the traditional speculative production. However, the material for the reactive production must be prepared in advance. We find that the supply chain is not always benefited by adding this reactive production. The condition under which the two-mode production outweighs the single-mode production depends on the demand features, besides the costs, price and salvage values. Because the two-mode production can not only decrease the upfront investment cost but also ensure a higher fill rate, it has an advantage over the purely speculative production. Compared with the purely reactive production, the two-mode production is more valuable as it allows enjoying the cheap early production without sacrificing the fill rate. Furthermore, a pricing contract with a return policy is proposed and optimized to coordinate the supply chain when the two-mode production is beneficial. This contract allows the arbitrary split of the supply chain profit between the manufacturer and buyer, which makes Pareto improvement possible.

Numerical examples show that compared to the purely speculative production, the benefit of the two-mode production is larger for the lower production delay costs (especially low material cost), medium sales price, higher customer shortage sensitivity, higher material compatibility, higher resolved demand uncertainty or lower unresolved demand uncertainty.

In contrast to the purely reactive production, the benefit of the two-mode production is larger for the higher production delay costs (especially higher material cost), lower sales price, higher customer shortage sensitivity, lower material compatibility, lower resolved demand uncertainty or higher unresolved demand uncertainty.

6.2 Limitations and Future Work

6.2.1 Manufacturer

The models about manufacture systems are simple in this thesis. In Scenario 1, manufacture systems have been modeled as a series of M/M/1 queues to analyze the relationships among the ordering cost, time and quantity. Scenario 2 and 3 assume that manufacturers charge a higher cost and impose a quantity restriction on later orders without considering how manufactures set these parameters.

In practice, manufacture systems are more complicated. It would be more practical to consider more general systems. Two potential directions can be considered as future research: (1) decomposition methods based on the stochastic independence assumption (Reiser & Kobayashi, 1974; Whitt, 1983), and (2) the concept of intrinsic ratio (K. Wu & McGinnis, 2013), which can be used to capture the dependence among stations.

In addition, to make quick response, manufactures can adopt different strategies, e.g., redesigning production processes, resorting to responsive suppliers, or procuring products directly from sport markets. Thus, it would be more complete to compare all the possible strategies manufacturers may use and provide the conditions under which one dominates the

others.

6.2.2 Two Retailers and Types of Products

We focus on only one retailer's decision and one type of products in this thesis. In practice, manufacturers may sell different types of products to different retailers and place different priorities on retailers' orders. It is interesting to investigate how manufacturers allocate the limited reactive capacity between two retailers or two types of products. It is also worthy considering the interaction among manufacturers and retailers if retailers know the limitation on the later ordering quantity.

6.2.3 Supply Uncertainty

This research only focuses on demand uncertainty. In practice, supply uncertainty exists and plays a critical role in making a strategy. For example, supply yield in agriculture is volatile due to unpredicted weather condition. The impact of supply yield uncertainty has been investigated in the literature (refer to Yano and Lee (1995)). Hence, one of the future research directions is to study the case where the maximum quantity at the second instant is uncertain. In some markets (e.g., spot markets), the supply cost varies with time (Gurnani & Tang, 1999). Another research direction is to extend the models to incorporate an uncertain second procurement cost.

References

- Adan, I., Resing, J. (2002). Queueing theory. *Department of Mathematics and Computing Science, Eindhoven University of Technology*.
- Allon, G., Van Mieghem, J. A. (2010). Global dual sourcing: Tailored base-surge allocation to near-and offshore production. *Management science*, 56, 110-124.
- Anderson, E. G., Morrice, D. J., Lundeen, G. (2006). Stochastic optimal control for staffing and backlog policies in a two-stage customized service supply chain. *Production and Operations Management*, 15, 262-278.
- Aviv, Y. (2001). The effect of collaborative forecasting on supply chain performance. *Management science*, 47, 1326-1343.
- Aviv, Y. (2002). Gaining benefits from joint forecasting and replenishment processes: The case of auto-correlated demand. *Manufacturing & Service Operations Management*, 4, 55-74.
- Aviv, Y. (2003). A time-series framework for supply-chain inventory management. *Operations research*, 51, 210-227.
- Aviv, Y. (2007). On the benefits of collaborative forecasting partnerships between retailers and manufacturers. *Management science*, 53, 777-794.
- Barnes-Schuster, D., Bassok, Y., Anupindi, R. (2002). Coordination and flexibility in supply contracts with options. *Manufacturing & Service Operations Management*, 4, 171-207.

- Bonser, J. S., Wu, S. D. (2001). Procurement planning to maintain both short-term adaptiveness and long-term perspective. *Management Science*, 47, 769-786.
- Boute, R. N., Van Mieghem, J. A. (2011). Global dual sourcing and order smoothing. In *2011 MSOM Annual Conference, Ann Arbor, MI, June* (pp. 26-28): Citeseer.
- Boyacı, T., Özer, Ö. (2010). Information acquisition for capacity planning via pricing and advance selling: When to stop and act? *Operations Research*, 58, 1328-1349.
- Brown, A., Lee, H. L. (1997). Optimal pay-to-delay capacity reservation with application to the semiconductor industry. *Department of Industrial Engineering and Engineering Management Working Paper, Stanford University, Stanford, CA*.
- Brown, R. G. (1959). Statistical Forecasting for Inventory Control. *McGraw-Hill: New York*.
- Burke, P. J. (1956). The output of a queuing system. *Operations Research*, 4, 699-704.
- Cachon, G. P. (2003). Supply chain coordination with contracts. *Handbooks in Operations Research and Management Science*, 11, 227-339.
- Cachon, G. P., Lariviere, M. A. (2005). Supply chain coordination with revenue-sharing contracts: strengths and limitations. *Management Science*, 51, 30-44.
- Chan, F. T., Chan, H. K. (2006). A simulation study with quantity flexibility in a supply chain subjected to uncertainties. *International Journal of Computer Integrated Manufacturing*, 19, 148-160.
- Chen, H.-W., Gupta, D., Gurnani, H. (2013). Fast-ship commitment contracts in retail supply chains. *IIE Transactions*, 45, 811-825.
- Chen, H., Chen, Y., Chiu, C.-H., Choi, T.-M., Sethi, S. (2010a). Coordination Mechanism for

- the Supply Chain with Leadtime Consideration and Price-dependent Demand. *European Journal of Operational Research*, 203, 70-80.
- Chen, H., Chen, Y. F., Chiu, C.-H., Choi, T.-M., Sethi, S. (2010b). Coordination mechanism for the supply chain with leadtime consideration and price-dependent demand. *European Journal of Operational Research*, 203, 70-80.
- Choi, T.-M., Li, D., Yan, H. (2004). Optimal Single Ordering Policy with Multiple Delivery Modes and Bayesian Information Updates. *Computers & Operations Research*, 31, 1965-1984.
- Choi, T.-M. J., Li, D., Yan, H. (2006). Quick Response Policy with Bayesian Information Updates. *European Journal of Operational Research*, 170, 788-808.
- Cotherix Inc. (2006). Form 10-K, filed March 14, 2006, for the Fiscal Year Ended December 31, 2005. *Lexis-Nexis Academic database (retrieved 8.11.07)*.
- Das, S. K., Abdel-Malek, L. (2003). Modeling the flexibility of order quantities and lead-times in supply chains. *International journal of production Economics*, 85, 171-181.
- Donohue, K. L. (2000). Efficient supply contracts for fashion goods with forecast updating and two production modes. *Management Science*, 46, 1397-1411.
- Dvoretzky, A., Kiefer, J., Wolfowitz, J. (1952). The inventory problem: I. Case of known distributions of demand. *Econometrica: Journal of the Econometric Society*, 187-222.
- Emmons, H., Gilbert, S. M. (1998). Note. The Role of Returns Policies in Pricing and Inventory Decisions for Catalogue Goods. *Management Science*, 44, 276-283.

- Erkoc, M., Wu, S. D. (2005). Managing High - Tech Capacity Expansion via Reservation Contracts. *Production and Operations Management*, 14, 232-251.
- Fisher, M., Raman, A. (1996). Reducing the Cost of Demand Uncertainty through Accurate Response to Early Sales. *Operations Research*, 44, 87-99.
- Fisher, M. L. (1997). What Is the Right Supply Chain for Your Product? *Harvard Business Review*, 75, 105-117.
- Fisher, M. L., Hammond, J. H., Obermeyer, W. R., Raman, A. (1994). *Making Supply Meet Demand in an Uncertain World*: Graduate School of Business Administration, Harvard University.
- Fleischhacker, A. J., Zhao, Y. (2013). Balancing Learning and Economies of Scale for Adaptive Clinical Trials. *Operations Research for Health Care*, 2, 42-51.
- Gavirneni, S., Kapuscinski, R., Tayur, S. (1999). Value of information in capacitated supply chains. *Management Science*, 45, 16-24.
- Ghemawat, P., Nueno, J. L., Dailey, M. (2003). *Zara: Fast Fashion*: Harvard Business School.
- Gilbert, K. (2005). An ARIMA supply chain model. *Management science*, 51, 305-310.
- Glock, C. H. (2012). Lead Time Reduction Strategies in a Single-vendor–single-buyer Integrated Inventory Model with Lot Size-dependent Lead Times and Stochastic Demand. *International journal of production Economics*, 136, 37-44.
- Graves, S. C. (1999). A single-item inventory model for a nonstationary demand process. *Manufacturing & Service Operations Management*, 1, 50-61.

- Graves, S. C., Kletter, D. B., Hetzel, W. B. (1998). A Dynamic Model for Requirements Planning with Application to Supply Chain Optimization. *Operations Research*, 46, S35-S49.
- Gurnani, H., Tang, C. S. (1999). Note: Optimal Ordering Decisions with Uncertain Cost and Demand Forecast Updating. *Management Science*, 45, 1456-1462.
- Hausman, W. H. (1969). Sequential decision problems: a model to exploit existing forecasters. *Management Science*, 16, B-93-B-111.
- Heath, D. C., Jackson, P. L. (1994). Modeling the Evolution of Demand Forecasts with Application to Safety Stock Analysis in Production/Distribution Systems. *IIE Transactions*, 26, 17-30.
- Huang, H., Sethi, S. P., Yan, H. (2005). Purchase contract management with demand forecast updates. *IIE Transactions*, 37, 775-785.
- Iida, T., Zipkin, P. H. (2006). Approximate solutions of a dynamic forecast-inventory model. *Manufacturing & Service Operations Management*, 8, 407-425.
- Iyer, A. V., Bergen, M. E. (1997). Quick Response in Manufacturer-retailer Channels. *Management science*, 43, 559-570.
- Jin, M., Wu, S. D. (2001). Modeling capacity reservation in high-tech manufacturing. Working paper, Lehigh University, Pennsylvania.
- Kaminsky, P., Swaminathan, J. M. (2001). Utilizing Forecast Band Refinement for Capacitated Production Planning. *Manufacturing & Service Operations Management*, 3, 68-81.

- Kanda, A., Deshmukh, S. (2008). Supply chain coordination: perspectives, empirical studies and research directions. *International journal of production Economics*, 115, 316-335.
- Karmarkar, U. S. (1987). Lot sizes, lead times and in-process inventories. *Management Science*, 33, 409-418.
- Kim, J., Benton, W. (1995). Lot size dependent lead times in a Q, R inventory system. *The Internatioal Journal of Production Research*, 33, 41-58.
- Kleinrock, L. (1965). A conservation law for a wide class of queueing disciplines. *Naval Research Logistics Quarterly*, 12, 181-192.
- Lee, H., Whang, S. (2002). The Impact of the Secondary Market on the Supply Chain. *Management science*, 48, 719-731.
- Lee, H. L., Padmanabhan, V., Whang, S. (1997). Information distortion in a supply chain: the bullwhip effect. *Management Science*, 43, 546-558.
- Lee, H. L., So, K. C., Tang, C. S. (2000). The value of information sharing in a two-level supply chain. *Management science*, 46, 626-643.
- Lee, H. L., Tang, C. S. (1997). Modelling the Costs and Benefits of Delayed Product Differentiation. *Management Science*, 43, 40-53.
- Li, J., Chand, S., Dada, M., Mehta, S. (2008). Managing inventory over a short Season: Models with two procurement opportunities. *Manufacturing & Service Operations Management*, 11, 174-184.
- Li, J., Liu, L. (2008). Supply Chain Coordination with Manufacturer's Limited Reserve Capacity: An Extended Newsboy Problem. *International Journal of Production*

- Economics*, 112, 860-868.
- Li, X., Wang, Q. (2007). Coordination mechanisms of supply chain systems. *European Journal of Operational Research*, 179, 1-16.
- Milner, J. M., Kouvelis, P. (2005). Order quantity and timing flexibility in supply chains: the role of demand characteristics. *Management Science*, 51, 970-985.
- Miyaoka, J., Hausman, W. (2004). How a base stock policy using "stale" forecasts provides supply chain benefits. *Manufacturing & Service Operations Management*, 6, 149-162.
- Oh, S., Özer, Ö. (2013). Mechanism design for capacity planning under dynamic evolutions of asymmetric demand forecasts. *Management science*, 59, 987-1007.
- Özer, Ö., Uncu, O., Wei, W. (2007). Selling to the "Newsvendor" with a forecast update: Analysis of a dual purchase contract. *European Journal of Operational Research*, 182, 1150-1176.
- Özer, Ö., Wei, W. (2004). Inventory control with limited capacity and advance demand information. *Operations Research*, 52, 988-1000.
- Pasternack, B. A. (1985). OPTIMAL PRICING AND RETURN POLICIES FOR PERISHABLE COMMODITIES. *Marketing Science*, 4, 166-161.
- Raz, G., Druehl, C. T., Blass, V. (2013). Design for the environment: Life - cycle approach using a newsvendor Model. *Production and Operations Management*.
- Reiser, M., Kobayashi, H. (1974). Accuracy of the diffusion approximation for some queuing systems. *IBM Journal of Research and Development*, 18, 110-124.
- Serel, D. A. (2009). Optimal Ordering and Pricing in a Quick Response System. *International*

- Journal of Production Economics*, 121, 700-714.
- Serel, D. A., Dada, M., Moskowitz, H. (2001). Sourcing decisions with capacity reservation contracts. *European Journal of Operational Research*, 131, 635-648.
- Sethi, S. P., Yan, H., Zhang, H.-q. (2001). Peeling layers of an onion: Inventory model with multiple delivery modes and forecast updates. *Journal of Optimization Theory and Applications*, 108, 253-281.
- Sethi, S. P., Yan, H., Zhang, H. (2006). *Inventory and supply chain management with forecast updates* (Vol. 81): Springer Science & Business Media.
- Simchi-Levi, D., Simchi-Levi, E., Kaminsky, P. (1999). *Designing and managing the supply chain: Concepts, strategies, and cases*: McGraw-Hill Companies.
- Su, J. C., Chang, Y.-L., Ferguson, M. (2005). Evaluation of postponement structures to accommodate mass customization. *Journal of Operations Management*, 23, 305-318.
- Tang, C. S., Rajaram, K., Alptekinoglu, A., Ou, J. (2004). The benefits of advance booking discount programs: Model and analysis. *Management science*, 50, 465-478.
- Thomopoulos, N. T. (2015). *Demand Forecasting for Inventory Control*: Springer.
- Treville, S., Schürhoff, N., Trigeorgis, L., Avanzi, B. (2014). Optimal Sourcing and Lead - Time Reduction under Evolutionary Demand Risk. *Production and Operations Management*, 23, 2103-2117.
- Tsay, A. A. (1999). The quantity flexibility contract and supplier-customer incentives. *Management Science*, 45, 1339-1358.
- Vlachos, D., Tagaras, G. (2001). An Inventory System With Two Supply Modes and

- Capacity Constraints. *International journal of production Economics*, 72, 41-58.
- Wang, T., Atasu, A., Kurtulus, M. (2012). A Multiordering Newsvendor Model with Dynamic Forecast Evolution. *Manufacturing & Service Operations Management*, 14, 472-484.
- Wang, Y., Tomlin, B. (2009). To Wait or Not to Wait: Optimal Ordering under Lead Time Uncertainty and Forecast Updating. *Naval Research Logistics*, 56, 766-779.
- Webb A. (2006). Ripeness sticker takes the guesswork out of picking and eating. *The Albuquerque Journal*, May 25.
- Weng, Z. K. (2004). Coordinating order quantities between the manufacturer and the buyer: A generalized newsvendor model. *European Journal of Operational Research*, 156, 148-161.
- Whitt, W. (1983). The queueing network analyzer. *Bell System Technical Journal*, 62, 2779-2815.
- Wu, J. (2005). Quantity flexibility contracts under Bayesian updating. *Computers & Operations Research*, 32, 1267-1288.
- Wu, K., McGinnis, L. (2013). Interpolation approximations for queues in series. *IIE Transactions*, 45, 273-290.
- Yan, H., Liu, K., Hsu, A. (2003). Optimal Ordering in a Dual-supplier System with Demand Forecasts Updates. *Production and Operations Management*, 12, 30-45.
- Yano, C. A., Lee, H. L. (1995). Lot sizing with random yields: A review. *Operations research*, 43, 311-334.

Yao, Z., Leung, S. C., Lai, K. K. (2008). Manufacturer's revenue-sharing contract and retail competition. *European Journal of Operational Research*, 186, 637-651.

Zhao, Z., Ball, M., Kotake, M. (2005). Optimization-Based Available-To-Promise with Multi-Stage Resource Availability. *Annals of Operations Research*, 135, 65-85.

Zhou, Y.-W., Wang, S.-D. (2009). Manufacturer-buyer coordination for newsvendor-type-products with two ordering opportunities and partial backorders. *European Journal of Operational Research*, 198, 958-974.