Operations strategies against demand uncertainty : process flexibility and real options

Han, Zhiguang

2014

Han, Z. (2014). Operations strategies against demand uncertainty : process flexibility and real options. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/61338>

<https://doi.org/10.32657/10356/61338>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**OPERATIONS STRATEGIES AGAINST DEMAND
UNCERTAINTY: PROCESS FLEXIBILITY AND REAL
OPTIONS**

HAN ZHIGUANG

NANYANG BUSINESS SCHOOL

2014

Operations Strategies against Demand Uncertainty: Process Flexibility and Real Options

Han Zhiguang

Nanyang Business School

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

Doctor of Philosophy

2014

I would like to dedicate this thesis to my wife, without whose constant love and support throughout these years I would not have been able to accomplish this; and to my parents who always stand by me and taught me courage, patience and integrity from birth.

Acknowledgements

First of all, I would like to express my heartfelt gratitude to my supervisors Dr. Chen Shaoxiang and Dr. Geoffrey Chua for the guidance and help during my doctoral study. Without their continuous support, collaboration and encouragement, this thesis would have been a task impossible.

In the past few years, I benefited a lot from Dr. Chen and Dr. Chua, both of whom are excellent but of different styles. Dr. Chen cannot emphasize more on methodology thinking. As he told me many times, a PhD holder should be a philosopher in her/his own field and principles and methodologies come before mathematics. His innovative ideas, insights into industries help me lay the foundations of my research. Dr Chua on the other hand always encouraged me to think independently and I was given full freedom to explore the areas of interest. Also I could not articulate how many times I have discussed my research with him. His patience, mentorship as well as friendship are invaluable to me. Working with him in the past three years has been a pleasant experience. Thanks, Geoff.

I am also grateful to my defense committee members, Dr. Li Zhi-Feng and Dr Arvind Sainathan for their helpful suggestions and feedback. I would also like to thank other faculty members in OM division, Dr Viswanathan, Wang, Rohit and Chien-Ming, Fang who have given me useful advice. I also thank the two anonymous external examiners for their valuable comments and suggestions that helped improve this thesis. Special thanks go to the staff in PhD office, Bee Hua, Nisha and Karen for their assistance and help.

Finally I wish to thank the PhD students especially Liu Yan in OM division. They made my PhD life in NTU enjoyable and colorful.

Abstract

As markets are increasingly volatile today, this thesis aims to study two strategies, process flexibility and real options, that can help hedge against demand uncertainty which has posed big challenges for many manufacturing systems.

We consider the process flexibility design problem for balanced but non-identical systems. While several methods and indices have been proposed in the literature, they are mostly time-consuming or do not explicitly take into account the demand variance information. In this thesis, we analytically study the effect of demand variance on the system performance and the long chain efficiency. We then develop a variance-based Hub-and-Chain method (VHC) to generate flexibility designs for the balanced but non-identical case. Numerical tests show that VHC outperforms the long chain by 45% on average and outperforms the constraint sampling method by 25% on average. Moreover, it only utilizes 1% of the computational requirements of the constraint sampling method. We then extend VHC to the unbalanced case. We benchmark VHC against the constraint sampling method and the expansion heuristic. We find that our method produces better flexibility

designs than both existing methods in 99% of all the scenarios considered. Finally, we test the distributional robustness of VHC and show it is robust for the lognormal distribution and the gamma distribution.

In the second essay, we prove in a multistage system the existence of a sparse process flexibility structure that is able to achieve $1 - \epsilon$ benefits of full flexibility and identify the corresponding conditions which show the number of links needed to accomplish such benefits is related to the number of stages. We also prove the dual problem of our model is a 0-1 integer program. Assuming only the moment information known, we then solve the dual problem by a completely positive cross-moment model (CPCMM).

In the third essay, we consider a two-period newsvendor problem in a B2C setting. To mitigate the demand risk, the manufacturer introduces a real-option-selling scheme in period 1 despite the spot market in period 2. Customers are uncertain about their future valuations for the product and may prefer to buy the option, the right to purchase one unit product at a fixed price in period 2. By inducing customers to buy the option, the manufacturer obtains more accurate demand information for period 2. We study the joint option pricing and production planning decision faced by the manufacturer. In the first part of this essay, we introduce a fluid model where the number of customers who exercise the option in period 2 is deterministic given a specific option. We first show in the single-option and single-market-segment setting

the option-selling scheme is always optimal in comparison with the newsvendor. Moreover, surprisingly we find the optimal strike price constantly equals to the unit production cost. We then extend the fluid model to the two independent market segments setting with three subcases as follows, single option for high segment, single option for both segments, one option for each segment. The findings of interest are subcase 1 is equivalent to a single-option and single-market-segment problem along with a separate newsvendor problem and subcase 2 is a classic Kalman filtering problem while subcase 3 is similar to a mechanism design problem. The dominance of option selling over the newsvendor is again confirmed. In addition the condition of subcase 3 dominating subcase 1 can be characterized by the ratios of segment sizes and marginal profit and cost. In the second part of this essay, we concentrate on a binomial model where the number of customers who will exercise the option is binomial but can be approximated by the normal distribution. The market has correlated informed and uninformed segments with homogeneous valuation distribution. The purpose of this part is threefold, to incorporate the advance selling into the option-selling framework and show the dominance of the latter, to show the dominance of option selling over the newsvendor in the expedite setting and a counterintuitive result - having more information is not always optimal.

Contents

Contents	vi
List of Figures	ix
List of Tables	x
Nomenclature	x
1 Introduction	1
2 Hub and Chain: Variance-Based Design of Process Flexibility	9
2.1 Introduction	9
2.2 Literature Review	18
2.3 Model and Settings	22
2.4 Main Results	24
2.5 Bounds on Flexibility Design Efficiency	37
2.6 The Hub-and-Chain Method	41
2.6.1 Method 1 - No Budget Specified	44
2.6.2 Method 2 - Set Partitioning with Budget	45

2.7	Simulation and Analysis	47
2.8	Extension to Unbalanced Case	56
2.8.1	Numerical Results	59
2.9	Distributional Robustness	61
2.10	Conclusions	65
3	Multistage Process Flexibility	68
3.1	Introduction	68
3.2	The Model	69
3.3	Numerical Results	81
3.4	Worst Performance Evaluation with Unknown Distribution	84
3.4.1	Completely Positive Representation	85
3.4.2	CPPCMM for Process Flexibility	88
3.4.3	Numerical Test	91
3.4.4	Flexibility Structure Design	93
3.5	Conclusion	94
4	Production Planning Using Real Options	96
4.1	Introduction	96
4.2	Literature Review	100
4.3	Model Settings - Single Option	104
4.3.1	Single Market Model	106
4.3.1.1	Advance Selling	112
4.3.2	Single Option with Two Market Segments	113
4.3.2.1	Case 1.	117

CONTENTS

4.3.2.2	Case 2.	118
4.3.2.3	Case 3.	123
4.4	Two Options with Two Market Segments	125
4.4.1	Comparison with The Newsvendor and Single Option . . .	133
4.5	Binomial Model with Informed and Uninformed Customers	136
4.5.1	All Customers Are Informed	139
4.5.1.1	When Lost Sale Cost is Formidable	141
4.5.1.2	The Expedite Model	142
4.5.1.3	Comparing with The Newsvendor	143
4.5.2	Not All Customers Are Informed	145
4.5.2.1	When Lost Sale Cost is Formidable	148
4.5.2.2	The Expedite Model	152
4.5.2.3	Comparing Option Selling to The Newsvendor . .	153
4.5.2.4	Is Informing All of The Customers Optimal? . . .	154
4.6	Conclusions	156
5	Final Remarks and Future Research	158
	References	162

List of Figures

2.1.1 Introduction example	15
2.5.1 The three-point distribution	39
2.5.2 The normal distribution	40
2.7.1 How the pairs are grouped.	50
2.7.2 The hub-and-chain structure formed.	51
2.7.3 Structure efficiency of the three methods	53
2.8.1 A counter example	57
3.3.1 Average shortfalls of multistage systems	83
4.3.1 Manufacturer's decision process	105
4.3.2 Customer's decision process	107
4.3.3 Indifference curve	108
4.3.4 Optimal decision for single market	111
4.3.5 Indifference valuation curves of two segments	116
4.3.6 Optimal decision when only high segment buy	120
4.4.1 Optimal decision for the two-market and two-option case	132

List of Tables

2.1	Flexibility design approaches taking into account variance or not .	21
2.2	An Example of Demand Scenario	48
2.3	The results of different methods (1)	52
2.4	The Results of different methods (2)	54
2.5	The Results of the second method with 40 links	55
2.6	The results of the lognormal distributions	63
2.7	The results of the gamma distributions	64
3.1	Performances and CPU time of the SDP	93
3.2	Benchmark performances by simulation	93
4.1	Notations	106

1

Introduction

There is no business model that can prevail forever. A brief review of the history of the electronic industry or more specifically and recently the mobile phone market can demonstrate that. In the past five years Nokia's market share has kept shrinking in contrast to the fast rise of Apple and Samsung. In 2008 rarely people even an expert in the electronic industry would have forecasted that 4 years later Nokia's Symbian system would account for 1.2% of the market share in Q4 2012 with the figure of 57.1% in Q2, 2008 (Source: Gartner). Nowadays Apple and Samsung are competing fiercely for the mobile market share. But we still are not able to predict the future and what we only know is the one who can adapt quickly and swiftly to the turbulent market will win eventually.

Based on our observations on some companies and industries, we believe that fast advancing information and production technology, short lifecycle and customization together can, at least partially, characterize the current market and are pushing the players in business to innovate business models.

Firstly, the Internet is changing the landscape and environment of business. Take online sales as an example. It is gradually substituting, if not supplementing, the brick-and-mortar stores due to its huge advantages in cost, speed as well as convenience compared to traditional retailing. In the last decade, the Internet has boosted the success of companies like Dell, Amazon in US and Taobao in China and is still continually reforming the business environment. In 2012 the transactions on the websites Taobao.com and Tmall.com, both of which are subsidiaries of the e-commerce company Alibaba in China, amounted to 1.1 trillion yuan (\$170 billion), more than ebay and Amazon combined ([Anonymous, 2013](#)). More importantly, thanks to the Internet for the first time in history a manufacturer or retailer can reach, communicate with and obtain feedback from her mass customers easily and quickly.

Secondly, the lifecycles of many products, particularly like PCs, electronics, and even automobiles, are becoming shorter and shorter. For example the lifecycle of a car model on average is 10.6 years during the period 1970-1983 while it has decreased to 5.6 years during 1998-2006 ([Volpato and Stocchetti, 2008](#)). It is the fierce competition, we believe, that on one hand prompts manufacturers to introduce new models to market sooner, and on the other hand forces manufacturers to look for approaches to squeeze the production organization time before the release of new products otherwise competitors would catch up and take away market share. If it took years for Apple to possess full capacity to produce ipad1 after its launch in 2010, this would have left sufficient time for other PC manufacturers to launch their own tablets and of course the huge success of Apple would not have existed in the first place. Thus streamlined operations, like product innovation,

are crucial to a company's success in the fierce competition.

Thirdly, customers are no longer satisfied with products from mass production and instead prefer something special. The demand for customization has generated more niches than ever before and therefore drives companies to stretch product lines or expand product mixes to satisfy them. Wide product mixes and long product lines would increase the types of customers that companies can target, whereas they also mean tremendous pressure on the operations and require the manufacturing systems to be flexible enough to accommodate the miscellaneous product variants. As the number of product variants has significantly increased in the past decades, continuing to rely on inventory to guarantee service level will pile up the warehouse and consequently will consume much more working capital than before, not to mention the obsolete cost and holding cost.

With companies facing fast-changing markets, inventory, the essential tool in operations management, is hardly enough to hedge against the demand uncertainty. Managers need new solutions. At the same time, the development of manufacturing and information technologies allows companies to innovate new operations strategies, which can be categorized into two main approaches, adjusting the inner and upstream operations from the supply point of view including shortening the lead time, investing in flexible manufacturing and workforce, reserving flexible capacity, postponement, modularity etc. such that the supply can quickly respond to the rapid market changes; and deepening the integration with customers in order to obtain continuous and more accurate demand information or dynamically change the pattern of demand to fit the supply in favor of companies, including pricing, advance selling, supply chain contracts etc. This thesis

attempts to explore operations strategies in both the supply and demand aspects, which are process flexibility and real options, in order to better understand the issue of matching supply with demand - the long lasting concern in operations management.

In the first two essays of this thesis we research on process flexibility, which is defined by [Sethi and Sethi \(1990\)](#) as *relating to the set of part types that the system can produce without major setups*. Clearly process flexibility is a strategy attempting to buffer demand uncertainty from the supply aspect; and from the definition we can see that it is deeply rooted in the development of flexible manufacturing technology as any major setup or high switchover cost makes the strategy not economical. Meanwhile, the process flexibility problems usually involve two levels of decision-making - the strategic level and the operational level. [Fine and Freund \(1990\)](#) study the problem of investment in product-flexible manufacturing capacity. They formulate the problem as a two-stage stochastic program. In the first stage, the firm makes the decisions on how much flexible capacity and dedicated capacity to acquire, which are decision-makings of the strategic level. In the second stage, the firm observes the random demand and then makes the production decisions, which are decision-makings of the operational level. The problem can be solved by a backward method and the first stage decisions are eventually based on possible future states of the world which implicitly include the demand variability. However, [Fine and Freund \(1990\)](#) model how much flexibility should the firm have in two extreme cases, either dedicated capacity or fully flexible capacity. Since [Jordan and Graves \(1995\)](#) initiated the stream of research on process flexibility structure, several papers have been published in this area.

In this line of research, the central problem switches to how much *partial* flexibility should be added to the given dedicated capacity. That is, the capacity can be reconfigured to produce only part of the products when the product demand mean and variability are known and the plant capacity is fixed. In practice, this represents the case wherein the firm already has set up and operated the assembly lines for a period and intends to better hedge against the demand variability by adding flexibility rather than increasing capacity. The reason for not choosing to increase capacity may involve cost-capacity tradeoffs. Moreover, very often in the literature, the initial dedicated manufacturing systems are balanced, i.e. the product's demand mean equal to its dedicated plant capacity. Like in [Fine and Freund \(1990\)](#), the central problem studied in [Jordan and Graves \(1995\)](#) as well as this line of research is also a two-stage stochastic program. In the first stage, the firm makes the strategic decisions on the flexibility design. In the second stage, the firm makes operational decisions on the production of the flexible manufacturing system after observing the demand realization. Likewise, the strategic decisions are based on the demand variability, totally if the initial dedicated system is balanced. For example, [Jordan and Graves \(1995\)](#) propose a JG index for a product set, which can be applied to design process flexibility, is a function of the demand standard deviations.

Our focuses of the first two essays include 1) the effect of the demand variation on the performances of flexibility designs and how to use variance information to design good process flexibility structures and 2) multistage process flexibility problem and the evaluation of the performance when the demand distributions are unknown. In the first essay we use an index called *group coefficient of varia-*

tion (GCV) to design process flexibility for both of the balanced and unbalanced bipartite systems. The numerical results demonstrate that our method performs much better than the existing methods in the literature. In fact we show in our numerical study that little room is left for future improvement. In the second essay, first we show there exists a sparse structure which can achieve almost the same performance as full flexibility and that the sparsity is related to the number of stages of the supply chain. Then we show the dual of the structure evaluation problem actually is a 0-1 integer program which can be solved by a *completely positive cross moment model* (CPCMM).

In the third essay we concentrate on using real call option selling to leverage the customers' demand information. As realized by both managers and researchers that advance selling, a discount or premium offered to induce customers to buy in advance, can reduce uncertainty in the spot market, Gallego and Şahin (2010) incorporate advance selling into the framework of real options. In other words, advance selling is a special case of real options. However, the capacity in Gallego and Şahin (2010) is considered to be fixed since the background industry is airline. In comparison, the capacity in this essay is a decision variable and it considers real options as a bridge between supply and demand. The selling result of options in period 1 updates the demand information such that the newsvendor manufacturer can better plan her production or capacity. Options as hedging instruments therefore significantly reduce the risk in the supply chain. In this sense, our essay is also related to Prasad et al. (2011), who use advance selling to obtain more accurate demand information about the remaining market through the correlation coefficient between the informed and uninformed market segments.

In this essay the newsvendor problem is modelled over two periods. Period 1 is an option-selling season and then period 2 is the spot market. We first consider a fluid model with multiple market segments and derive interesting properties. In addition, we answer the question that under what conditions the multi-option strategy is optimal. Then we use a model similar to [Prasad et al. \(2011\)](#), where the demand in spot market is binomial but can be approximated by the normal distribution, to investigate how the uncertainty of exercising options affects the optimal solutions. We also show an interesting result that in some circumstances concealing the information of option selling from part of customers is better than making the whole market informed.

To our knowledge, both process flexibility and real options in a B2C setting have already been adopted in practice. Interesting examples will be given in the following chapters. The purpose of this thesis is to further extend and justify these methods and provide theoretical foundations for them rather than introduce new concepts.

The remainder of this thesis is organized as follows. In chapter 2 we first show both analytical and numerical justifications for our GCV index and then use the index to design heuristic called hub-and-chain method for balanced but nonidentical systems; at last we extend the heuristic to unbalanced systems. Numerical study is presented to demonstrate the effectiveness of our heuristic. In chapter 3 we first prove the existence of a sparse structure and apply the constraint sampling method to the multistage system. Then we investigate the property of the dual problem. At last we show the CPCMM method and the numerical study. Chapter 4 is composed of two parts. The first part studies a fluid model for the

single-market and single-option case. We then extend the model to the two market segments case in which we have three policies as follows, single option for high segment, single option for both high and low segments, and one option for each segment. We compare among the three policies. The second part of Chapter 4 studies a binomial model with informed and uninformed market segments. We first study the case in which the whole market is informed and then the case in which part of the market is informed. Final remarks and future research are presented in Chapter 5.

2

Hub and Chain: Variance-Based Design of Process Flexibility

2.1 Introduction

Manufacturing has become more and more flexible in response to fast changing customer demand. As a result of shorter lifecycles, a firm might miss the demand peak if she starts to build a new assembly line at the time that demand emerges. Thus production facilities should be engineered to quickly adapt to alternative products.

“If the last few years have taught us anything, it is that customer wants and needs can change quickly- much more quickly than we have been equipped to respond to efficiently in the past. At Michigan Assembly, we will achieve a level of flexibility we don’t have in any other plant around the world, which will allow us to meet shifting consumer pref-

erences in real time.” – Jim Tetreault, Ford vice president of North America Manufacturing (Ford, 2010).

“Lean isn’t good enough anymore. The new reality requires being both lean and flexible.” – Dave Cole, chairman emeritus for the Center for Automotive Research in Ann Arbor, Michigan (Woodyard, 2011).

Automaker Ford has stepped toward the strategy of leaner and more flexible manufacturing in the last few years. Ford retooled its assembly lines in Chicago, Wayne, Louisville to build different models on the same platform. Take the truck plant in Louisville as an example. Ford spent \$200 millions on retooling of the plant to produce the more fuel-efficient 2009 model Expedition and Navigator SUVs alongside the popular Ford F-Series Super Duty pickups (Ford, 2009). GM’s manufacturing section is also becoming more flexible after bankruptcy (Amend, 2010).

The flexible manufacturing technology, large number of product varieties and multiple production facilities together necessitate a delicate design of process flexibility network structure. Jordan and Graves (1995) introduce the impressive chaining strategy and provide several intuitive yet very useful guidelines to design a structure. It has been demonstrated that in many circumstances the chaining strategy could achieve most benefits of full flexibility (Chou et al., 2010; Hopp et al., 2004), thereby implying that investing in full flexibility may not be cost-efficient. The chaining strategy is the stepping stone to design other complex flexibility structures.

However, the success of chaining strategy depends on three conditions. First of all, as demonstrated in Chou et al. (2010) and Simchi-Levi and Wei (2012a), the longer the chain, the worse the *flexibility design efficiency* which is defined as

$$\frac{\text{performance of given flexibility design} - \text{performance of dedicated system}}{\text{performance of full flexibility} - \text{performance of dedicated system}},$$

the same concept as chain efficiency (CE) in Chou et al. (2010). The *performance* of a flexibility design is the expected output over all random instances. Hereafter we use the terms *performance* and *expected output* interchangeably. So a chain should not be too long. Second, all products in a closed chain should be, to some extent, homogeneous and so do all plants. In other words, a chain should not include a product with extremely small demand or a plant with extremely small capacity that could block the flow of idle extra demand. To illustrate, we refer to Figure 2.1.1 (a) to show a 3×3 2-chain on the top. The circles on the left represent products while the squares on the right denote plants. Each product demand follows a two-point distribution with equal probability for each point. For example, the two demand points for product 1 are 50 and 150 with demand mean 100. The link between a product and a plant means that the connected plant is capable of processing the connected product. The numbers beside the squares represent plant capacities. Our purpose is to maximize the expected output, i.e. the expected number of products, that can be produced by the system. If we divide the graph into two sub-graphs, a 2×2 2-chain with pair 1 and 2, and the third pair, not only a link is saved but also the expected output increases from 174.9 to 184.5. Obviously this is because pair 3 serves as a bottleneck. Excess demand

generated by product 2 is not able to reach idle capacity in pair 1 and capacity of pair 1 loses its risk-pooling function. This particular example counters the insight in the literature that the long chain is more desirable. As a consequence, it might be better off to divide a long chain into short chains if its chain efficiency is under expectation. Third, the flexibility design efficiency decreases in coefficient of variation of demand (Chou et al., 2010). Large CVs indicate that it would be better to use other strategies to buffer demand variation.

After Jordan and Graves (1995), several indices and methods for identifying or designing good structures have been developed (Chou et al., 2010, 2011; Iravani et al., 2005). However, to our best knowledge none of the existing indices and methods takes into account the variance of demand explicitly. Most of the indices or methods are trying to match the demand mean with capacity. This may not work under some circumstances since the process flexibility structure, a bipartite network, is stochastic. An important question is how to incorporate the variance factor into an index and then develop a method to design good structures?

In fact, demand variation plays a key role in the design of process flexibility. There are two aspects of its impact which we next discuss.

Firstly, demand variances determine the *potential* which is defined as the difference between the performances of full flexibility and the dedicated case, of a group of product-plant pairs. A product with smaller variance contributes less excess demand and in turn results in less idle capacity of its dedicated plant. Since the key idea behind process flexibility is pooling the idle capacity and excess demand through manufacturing flexibility, it is not worthwhile to include a product-plant pair with little contribution to the pool in a flexibility network. An extreme case

is a product-plant pair associated with deterministic demand. To illustrate, in Figure 2.1.1 (c), products on the left side follow the two-point distributions with point probability $1/2$. All of the demand means as well as the plant capacities are equal to 10. As we can verify, the expected output of the scenario in Figure 2.1.1 (c) is 39.75 in comparison to 29.5 of the dedicated case. When the fifth product is also set as $(0, 20)$, the expected output of 2-chain will decrease to 37.5 in comparison to 25 of the dedicated case. Although more uncertainty is troublesome in terms of the expected output, it causes the improvement of 2-chain over the dedicated system to increase by 2.25, implying the 2-chain structure is more link or cost-efficient in the later scenario. In other words, the higher demand variation makes the option of flexibility more attractive. Meanwhile, we can see that the total expected output in Figure 2.1.1 (d) only decreases by 0.25 yet we can save one link compared to Figure 2.1.1 (c). The foregoing examples shed some light on the role that variance plays in flexibility decision-making. To raise link efficiency, a product-plant pair with relatively low demand variation are less valuable to us in design of flexibility structure therefore should not be included in the chaining system, especially when a budget of links is imposed on the problem.

Secondly, demand variation also in part affects the *flexibility design efficiency* of a bipartite network. A flexibility design with high efficiency means extra demands can drift fluently through the network and then match the idle capacity in other plants. To illustrate see Figure 2.1.1 (b). The demand also follows the two-point distributions with point probability $1/2$. For a sample path $(15, 15, 15, 5, 5, 5)$, the idle capacity on pair 6 is not able to meet the extra demand from product 1. In contrast, if the demand distributions of products 1 to 6 are $(9, 11)$

instead of (5, 15), all excess demands eventually can flow to the idle capacity on the design given the same sample path (11, 11, 11, 9, 9, 9). Other sample paths can be argued in the same fashion. As a result, any decrease in the variance of products from 1 to 5 improves the *flexibility design efficiency*.

From the previous examples and analyses, we can see firstly, to efficiently use the links, a product-plant pair with approximately deterministic demand should not be combined into a 2-chain system. And a pair with higher variation is more valuable; secondly, a long chain is not always a wise option. Once some of the pairs work as bottlenecks, we need to reconstruct the original 2-chain by dividing it into shorter sub-chains, or discarding some particular pairs. The above observations then raise two concerns. Despite the fact that a pair with higher demand variation contributes more to the risk pooling effect of a 2-chain, it also could mean a lower flexibility design efficiency and that most of the benefits of full flexibility are not achieved by the 2-chain. Meanwhile, though each of several shorter chains can achieve a higher flexibility design efficiency separately, it may not be a good structure as a whole in terms of efficiency if there are no connections among them. In this chapter we try to answer how to incorporate these observations and concerns into the flexibility design.

In this chapter, we manage to prove the insights generated by the above examples. Subsequently, we use these insights to devise a variance-based method for designing process flexibility in balanced but non-identical systems. The way our method works is to first isolate products with low demand variances into a group that will only employ a dedicated design. Then, we group the remaining products into subgroups according to proximity of demand means. The subgroup

Figure 2.1.1: Introduction example

should also not exceed a maximum amount of group demand variability, measured by a group coefficient of variation (GCV) that we introduce in this paper. Each of these subgroups will employ the long chain design. Then, a satellite product-plant pair will be selected from each subgroup. The satellite pair of the subgroup with the largest demand means will serve as the hub to which all satellite pairs of the other subgroups will be connected. As such, we refer to our method as the variance-based hub-and-chain method or VHC.

The hub-and-chain designs in this essay share some characteristics with the hub-and-spoke networks, which have been extensively studied in the airline literature. Since the advent of the US airline deregulation in 1978, the airline networks in US quickly evolved from point-to-point systems to hub-and-spoke systems which conferred great benefits to carriers and passengers. In a point-to-point system, passengers and cargo board a flight at the origin and arrive at the destination nonstop. In contrast, in a hub-and-spoke system, passengers and cargo board a flight at the origin and then transfer at the hub to another flight for their destination. It is clear that in a network with n nodes, the pure hub-and-spoke structure, i.e. assuming there is a single hub, only needs $O(n)$ routes to connect all the nodes while the point-to-point structure needs $O(n^2)$ routes. Other benefits of the hub-and-spoke structure include economy of scale, economy of traffic density, risk pooling effect at the hub and high utilization of aircraft. Thanks to the hub-and-spoke structure, carriers can significantly extend their networks to cover much more destinations than ever before (for good reviews of the hub-and-spoke systems in airlines, see [Bryan and O’Kelly \(1999\)](#); [Cook and Goodwin \(2008\)](#)). Increasingly, hub-and-spoke networks are also extended to other areas such as telecommunication ([Klincewicz, 1998](#)) and land transportation. For example, the public transport system in Singapore is constructed based on the hub-and-spoke model ([Tan and Leong, 2013](#)). Interestingly, similar concepts can also be found even in hospital networks ([Golikeri, 2011](#)).

To our best knowledge, VHC is the first method that explicitly incorporates the demand variance information into the design of process flexibility. Moreover, our numerical tests show that it is computationally efficient and produces flexibility

designs that outperform the long chain by 45% and the designs from the constraint sampling method by 25%.

We make the following contributions to the literature. First, we analytically prove the effect of changes in demand means and variances on the flexibility design performance and long chain efficiency. Second, we discover three interesting insights as follows: (1) products with low demand mean can create a bottleneck effect, (2) products with low demand variance will result in inefficient utilization of flexibility links, and (3) the long chain efficiency decreases in demand variance of any product, hence it will be a good idea to provide this product with access to more capacity. Third, we use these insights to develop an index that captures group demand variability which then forms part of our hub-and-chain method for designing flexible balanced but non-identical systems. Fourth, we also extend our method to accommodate the unbalanced case where the number of products is still equal to the number of plants, but demand mean no longer matches capacity.

The remainder of this chapter is organized as follows. We review the related literature in Section 2. In Section 3, we provide our model and the settings. We present our main analytical results in Section 4. We show the bounds on the flexibility design efficiency and introduce our index in Section 5. Then the variance-based hub-and-chain method is introduced in Section 6. Section 7 uses simulation to test the effectiveness of VHC for balanced systems. Then we extend our method to unbalanced cases in Section 8. Section 9 tests the distributional robustness of VHC for the lognormal and gamma distributions. Finally, Section 10 concludes the chapter.

2.2 Literature Review

Due to its popularity, researchers have extended the chaining strategy to various areas since [Jordan and Graves \(1995\)](#). [Graves and Tomlin \(2003\)](#) study process flexibility in supply chains. [Gurumurthi and Benjaafar \(2004\)](#) examine chaining in queuing systems and observe that chaining can achieve most benefits of full flexibility in identical systems while chaining is not superior to non-chained configuration in non-identical systems. [Wallace and Whitt \(2005\)](#) and [Iravani et al. \(2007\)](#) extend the study of the chaining strategy to cross-training in call centers. [Hopp et al. \(2004\)](#) compare the “skill chaining” cross-training strategy with the “cherry picking” strategy in the context of a serial production system. [Bish et al. \(2005\)](#) and [Muriel et al. \(2006\)](#) study the effect of flexibility on the variability of production and inventory level as well as the upstream suppliers in a make-to-order context. [Hopp et al. \(2010\)](#) study vertical flexibility, i.e. the problem of locating process flexibility or logistics flexibility along supply chains.

Other researchers concentrate on understanding the properties of chains or other sparse structures, and attempt to explain why sparse structures, especially chains, perform very well. For example, [Aksin and Karaesmen \(2007\)](#) show that the marginal return of flexibility is diminishing in the flexibility level. [Bassamboo et al. \(2010\)](#) have similar results though in the newsvendor network setting. [Chou et al. \(2010\)](#) use random walk to obtain the chain efficiency and show that the asymptotic chain efficiency decreases in the coefficient of variation, given the normal distributions. Both [Chou et al. \(2010\)](#) and [Chou et al. \(2011\)](#) prove that there exists a sparse structure that can achieve most of the benefits of full flexibility.

Simchi-Levi and Wei (2012a) prove the supermodularity of two flexible arcs in a flexibility design which is part of a long chain, on condition of balanced system and exchangeable demand. Based on this result, it is shown that the performance of a long chain can be characterized by the difference of the performances associated with two open chains. The authors also provide a matrix multiplication method to compute the performance of the long chain under identical demand with discrete and positive support. Though not in closed-form, Simchi-Levi and Wei (2012a) provide a lower bound on the long chain efficiency which depends on a variable defined as the performance of full flexibility divided by the number of products. In addition, Simchi-Levi and Wei (2012a) prove that the long chain is the optimal design in terms of performance among all the 2-flexibility designs, where each product or plant has exactly two links. Wang and Zhang (2013) obtain a lower closed-form and distribution-free bound on the asymptotic ratio of the long chain performance to that of full flexibility. The number of plants is equal to that of products in the systems studied. The random product demands follow identical and independent distributions while all plants have the same capacity. Only the information of the first and second moments are available. In the case where demand mean and plant capacity are the same, the authors present a closed-form lower bound on the asymptotic chain efficiency which only depends on the CV. Although a closed-form bound is not available when the demand mean and the plant capacity are not equal, a tractable semidefinite program to compute the long chain efficiency is provided. The authors show that the bounds, obtained based on partial demand information, are very close to those bounds which use complete distribution information. The results then are extended to k-chain.

Another line of research is developing either indices to identify good sparse structures or heuristics to design process flexibility. [Iravani et al. \(2005\)](#) introduce the mean index and eigenvalue index based on structural flexibility matrix in the setting of cross-training. [Iravani et al. \(2007\)](#) show that the cross-training structure of agents in call center can be converted to a work sharing (WS) network. They introduce the average shortest path length (APL) index to identify designs with short average waiting times. The simulation results support the authors' assertion that the smaller the APL associated with a cross-training structure, the better the performance. [Chou et al. \(2010\)](#) present an algorithm of designing process flexibility called the constraint sampling method - an idea borrowed from robust optimization. The links in the full flexibility design are viewed as the constraints in the dual of the maximum performance problem. The set of constraints is endowed with a specific probability distribution and a probability weight is assigned to each link. The algorithm is to pick links randomly according to the specific probability distribution. The approach can be time consuming because it involves simulation and sampling, which also result in highly variable outputs. [Chou et al. \(2011\)](#) and [Deng and Shen \(2013\)](#) develop computationally lightweight heuristics called expansion heuristic and enhanced chaining guideline, respectively. However, [Chou et al. \(2011\)](#) do not take into account the information on demand variance despite that the heuristic can well balance the total capacity and demand, while [Deng and Shen \(2013\)](#)'s method works well only for unbalanced systems with i.i.d. demands and same capacity. [Simchi-Levi and Wei \(2012b\)](#) propose using plant cover indices to compare the robustness, i.e. the worst-case output, among various flexibility designs. They prove that a flexibility

design with a greater plant cover index, which itself only depends on the flexibility design, is more robust than another design with a smaller index within a robust measure set, if the demand uncertainty set is symmetric. While the long chain is more robust than any other design of which each product node has two links, it is not true to generalize this result to the designs with $2n$ (n is the number of plants or plants) links in total. It should be noted that [Chou et al. \(2011\)](#) and [Simchi-Levi and Wei \(2012b\)](#) are two exceptions in the literature because both of them study process flexibility from the perspective of worst-case analysis. The two papers complement each other well in the way that the expansion heuristic works well when the deviations of the expected product demands are large while the plant cover index heuristic is effective when the demands are i.i.d. or homogenous. Besides, other indices include J-G index by [Jordan and Graves \(1995\)](#), g -measure by [Graves and Tomlin \(2003\)](#) and expansion index by [Chou et al. \(2008\)](#). In Table 2.1 we only summarize the existing works in the literature

Table 2.1: Flexibility design approaches taking into account variance or not

Variance related	Variance unrelated
Jordan and Graves (1995)	Graves and Tomlin (2003)
Chou et al. (2010)	Iravani et al. (2005)
Wang and Zhang (2013)	Iravani et al. (2007)
	Chou et al. (2008)
	Chou et al. (2011)
	Simchi-Levi and Wei (2012b)

which provide indices or heuristics for designing process flexibility. The works in the first column take into account the demand variance information, explicitly or implicitly, while those in the second column do not.

2.3 Model and Settings

Consider a manufacturing system with n products and m plants. We model this system through a bipartite graph representation $G = (A, B, \mathcal{G})$, where A denotes the set of products, B denotes the set of plants, and $\mathcal{G}$ denotes the set of edges (i, j) signifying the capability of plant j to produce product i . That is, $A = \{1, 2, \dots, m\}$, $B = \{1, 2, \dots, n\}$, and $\mathcal{G} \subseteq A \times B$. Let vector $\mathbf{D}$ denote the random demand of product i and $\mathbf{d}$ is a specific realization. Also let D_i represent the random demand of product i and d_i is its specific realization. Let c_j denote the capacity of plant j and $\mathbf{c} = [c_1, \dots, c_n]^T$. Given a demand realization $\mathbf{d}$, the maximum output can be produced by a flexibility design is as follows

$$\begin{aligned}
Z(\mathbf{d}, \mathcal{G}) = & \max \sum_{i \in A} \sum_{j \in B} x_{ij}, \\
\text{s.t. } & \sum_{j \in B} x_{ij} \leq d_i \quad \forall i \in A \\
& \sum_{i \in A} x_{ij} \leq c_j \quad \forall j \in B \\
& x_{ij} \geq 0 \quad \forall (i, j) \in \mathcal{G} \\
& x_{ij} = 0 \quad \forall (i, j) \notin \mathcal{G}
\end{aligned}$$

where x_{ij} represents the amount of i th product manufactured by j th plant.

Under random demand $\mathbf{D}$, the maximum output $Z(\mathbf{D}, \mathcal{G})$ is itself also a random variable. Hence, we define the *performance*, also referred to as expected output, of $\mathcal{G}$ to be $E[Z(\mathbf{D}, \mathcal{G})]$. For ease of exposition and whenever the context is clear, we also use the notation $[\mathcal{G}]$ to denote this quantity. In general, the prob-

lem is to generate a flexibility design $\mathcal{G}$ that maximizes $[\mathcal{G}]$ without exceeding a flexibility budget b . That is, we want to solve $\max_{\mathcal{G}: |\mathcal{G}| \leq b} [\mathcal{G}]$, where $|\mathcal{G}|$ denotes the number of edges or links of a flexibility design $\mathcal{G}$. However, this problem involves a stochastic program and is notoriously difficult and time consuming to solve. As such, most works in the literature focus on developing heuristics to solve special classes of this problem.

To classify process flexibility design problems, we introduce the following definitions. We say that the system is *balanced* if $m = n$, and the plants and products can be arranged such that $E[D_i] = c_i$ for all $i \in A$. Moreover, we call the system *identical* if the product demands are IID and the plant capacities are all equal, i.e. $D_i = D$ in distribution for all $i \in A$ and $c_j = c$ for all $j \in B$. Because of our interest in problems with heterogeneous product demand variances, we focus our study on non-identical systems. For ease of analysis and for purpose of intuition building, we first consider balanced but non-identical systems.

That said, we let n denote the system size (since $m = n$), and let $\mu_i = E[D_i] = c_i$ be the mean demand and σ_i^2 be the demand variance for product i for all $i = 1, 2, \dots, n$. We also define a few classes of flexibility designs for balanced systems of size $n \geq 2$. The dedicated design is defined as $\mathcal{D}_n = \{(i, i) | i = 1, 2, \dots, n\}$; the open chain design is defined as $\mathcal{L}_n = \mathcal{D}_n \cup \{(i, i+1) | i = 1, 2, \dots, n-1\}$; the long chain design is defined as $\mathcal{C}_n = \mathcal{L}_n \cup \{(n, 1)\}$; and the full flexibility design is defined as $\mathcal{F}_n = \{(i, j) | i, j = 1, 2, \dots, n\}$. Our goal in this chapter, as in the literature, is to analyze the performances of these flexibility designs as well as to develop new methods to improve on these designs, particularly for the balanced but non-identical case.

To this end, we further define two performance measures that will be useful in developing our new flexibility design method. We define the *potential* of a balanced system as the difference between the performances of full flexibility and the dedicated design. That is, we let the potential be $p(\mathbf{D}) = [\mathcal{F}_n] - [\mathcal{D}_n]$. Next, we define the *efficiency* of a flexibility design as the ratio of the difference between the performances of the given flexibility design and that of the dedicated design over the potential of the system. That is, we let the efficiency of $\mathcal{G}$ be $f(\mathbf{D}, \mathcal{G}) = \frac{[\mathcal{G}] - [\mathcal{D}_n]}{[\mathcal{F}_n] - [\mathcal{D}_n]}$. For example, $f(\mathbf{D}, \mathcal{C}_n)$ is equivalent to the chaining efficiency used in Chou et al. (2010).

2.4 Main Results

In this section, we examine the effect of changes in demand means and variances on the system potential and flexibility design efficiency. The results will then be used to obtain insights on how to develop methods that can generate good flexibility designs.

For analytical purpose, we consider two types of demand distributions; namely, the normal distribution and the symmetric three-point distribution. For the normal distribution, we assume that for $i = 1, 2, \dots, n$, $D_i \sim N(\mu_i, \sigma_i^2)$ truncated at 0 on the left (to avoid negative demand values) and at $2\mu_i$ on the right. Moreover σ_i/μ_i is small enough such that $Pr(D_i < 0)$ and $Pr(D_i > 2\mu_i)$ are negligible. So in our proof we use the normal distribution instead of the truncated normal distribution as an approximation. This is a reasonable assumption and is common in the literature (Chou et al., 2010; Jordan and Graves, 1995). For the symmetric

three-point distribution, we assume that for $i = 1, 2, \dots, n$,

$$D_i = \begin{cases} \mu_i - \eta_i, & \text{with probability } t_i \\ \mu_i, & \text{with probability } 1 - 2t_i \\ \mu_i + \eta_i, & \text{with probability } t_i \end{cases}$$

where $0 \leq \eta_i \leq \mu_i$ and $0 \leq t_i < \frac{1}{2}$. This is likewise a reasonable assumption because it characterizes products that have peak, shoulder, and off-peak demands. That said, higher values of η_i or t_i imply higher demand variance. For both distribution types, we further assume that the product demands are independent.

With a slight abuse of notation, we let system potential $p(\mathbf{D})$ be changed to $p(\boldsymbol{\mu}, \boldsymbol{\sigma})$ for the normal distribution and $p(\boldsymbol{\mu}, \boldsymbol{\eta}, \mathbf{t})$ for the symmetric three-point distribution. Similarly, we change the flexibility design efficiency to $f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{G})$ and $f(\boldsymbol{\mu}, \boldsymbol{\eta}, \mathbf{t}, \mathcal{G})$ whenever appropriate. We are now ready to present our results.

Theorem 2.4.1. *Consider a balanced system of size n . For the normal distribution, the system potential $p(\boldsymbol{\mu}, \boldsymbol{\sigma})$ is increasing and strictly concave in σ_k , yet independent of μ_k , $k = 1, \dots, n$. For the symmetric three-point distribution, the system potential $p(\boldsymbol{\mu}, \boldsymbol{\eta}, \mathbf{t})$ is increasing in η_k , t_k , but independent of μ_k , $k = 1, \dots, n$.*

Proof By definition of system potential, we have

$$\begin{aligned}
 p(\mathbf{D}) &= [\mathcal{F}_n] - [\mathcal{D}_n] \\
 &= E \left[\min \left(\sum_{i=1}^n D_i, \sum_{i=1}^n \mu_i \right) \right] - \sum_{i=1}^n E \left[\min \left(D_i, \mu_i \right) \right] \\
 &= \sum_{i=1}^n E \left[\left(D_i - \mu_i \right)^+ \right] - E \left[\left(\sum_{i=1}^n D_i - \sum_{i=1}^n \mu_i \right)^+ \right]
 \end{aligned}$$

For the normal distribution, it is well known that if $X \sim N(0, \sigma^2)$, then $E[X^+] = \frac{\sigma}{\sqrt{2\pi}}$. Therefore, we have the system potential $p(\boldsymbol{\mu}, \boldsymbol{\sigma}) = \frac{\sum_{i=1}^n \sigma_i}{\sqrt{2\pi}} - \frac{\sqrt{\sum_{i=1}^n \sigma_i^2}}{\sqrt{2\pi}}$. Take the derivative over σ_k and we can obtain $\frac{\partial p(\boldsymbol{\mu}, \boldsymbol{\sigma})}{\partial \sigma_k} = \frac{1}{\sqrt{2\pi}} \left(1 - \frac{\sigma_k}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \right) \geq 0$. It follows that $\frac{\partial^2 p(\boldsymbol{\mu}, \boldsymbol{\sigma})}{\partial \sigma_k^2} = -\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \left(1 - \frac{\sigma_k^2}{\sum_{i=1}^n \sigma_i^2} \right) < 0$ and $\frac{\partial p(\boldsymbol{\mu}, \boldsymbol{\sigma})}{\partial \mu_k} = 0$.

For the symmetric three-point distribution, we take $k = n$ without loss of generality. Let $[\mathcal{F}_{n-1}]$ represent the performance of the fully flexible system consisting of products and plants from 1 to $n-1$. And here we let $[\mathcal{D}_1]$ and $[\mathcal{D}_{n-1}]$ respectively represent the performances of the dedicated pair consisting of product n and plant n and the dedicated system consisting of product-plant pairs from 1 to $n-1$. When the fully flexible system consisting of products and plants from 1 to $n-1$ has excess demand, i.e. plants from 1 to $n-1$ cannot fulfill the demand of products from 1 to $n-1$, the excess demand can be satisfied or partially satisfied by the excess capacity, if any, at plant n . And the expected demand satisfied in such cases is $t_n Pr(\sum_{i=1}^{n-1} D_i > \sum_{i=1}^{n-1} \mu_i) \eta_n$. Likewise, $t_n Pr(\sum_{i=1}^{n-1} D_i < \sum_{i=1}^{n-1} \mu_i) \eta_n$ represents the expected demand satisfied when product n utilizes the excess capacity from the system consisting of products and plants from 1 to $n-1$. Therefore, $[\mathcal{F}_n] = [\mathcal{F}_{n-1}] + [\mathcal{D}_1] + t_n Pr(\sum_{i=1}^{n-1} D_i > \sum_{i=1}^{n-1} \mu_i) \eta_n + t_n Pr(\sum_{i=1}^{n-1} D_i < \sum_{i=1}^{n-1} \mu_i) \eta_n$.

Meanwhile, $[\mathcal{D}_n] = [\mathcal{D}_{n-1}] + [\mathcal{D}_1]$. So we have the following result

$$\begin{aligned} [\mathcal{F}_n] - [\mathcal{D}_n] &= [\mathcal{F}_{n-1}] - [\mathcal{D}_{n-1}] + E \left[\min \left[\left(\sum_{i=1}^{n-1} D_i - \sum_{i=1}^{n-1} \mu_i \right)^+, \eta_n \right] \right] \cdot t_n \\ &\quad + E \left[\min \left[\left(\sum_{i=1}^{n-1} \mu_i - \sum_{i=1}^{n-1} D_i \right)^+, \eta_n \right] \right] \cdot t_n. \end{aligned}$$

Since $[\mathcal{F}_{n-1}] - [\mathcal{D}_{n-1}]$ is independent of η_n, t_n , and μ_n , it suffices to examine the last two expectation terms. These terms are increasing in η_n , increasing in t_n , and independent of μ_n . To prove the case of $t_k, k \neq n$, we can simply relabel the product-plant pairs to make the k th pair become the n th pair in the new index system. Then we repeat the proof for t_n . Hence, the result follows. $\square$

Theorem 2.4.2. *Consider a balanced system of size n . Given any flexibility design $\mathcal{G}$ such that the dedicated design $\mathcal{D}_n \subseteq \mathcal{G}$, $f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{G})$ is increasing in μ_k , $k = 1, \dots, n$ for the normal distribution and $f(\boldsymbol{\mu}, \boldsymbol{\eta}, \mathbf{t}, \mathcal{G})$ is increasing in μ_k , $k = 1, \dots, n$ for the symmetric three-point distribution.*

Proof For any demand instance $\mathbf{d}$, the shortfall of a system with flexibility design $\mathcal{G}$ is defined as $\sum_{i=1}^n d_i - Z(\mathbf{d}, \mathcal{G})$. Thus the expected shortfall of this system is equal to $\sum_{i=1}^n \mu_i - [\mathcal{G}]$. Let $E[S(\mathbf{D}, \mathcal{F}_n)]$, $E[S(\mathbf{D}, \mathcal{D}_n)]$ and $E[S(\mathbf{D}, \mathcal{G})]$ be the expected shortfalls of full flexibility $\mathcal{F}_n$, dedicated $\mathcal{D}_n$ and $\mathcal{G}$ designs, respectively. We first look at the normal distribution. Suppose D'_k follows $N(\mu_k + \Delta_k, \sigma_k^2)$ and let $\mathbf{D}' = [D_1, \dots, D'_k, \dots, D_n]$. Then the performance difference between $\mathcal{F}_n$ and $\mathcal{D}_n$

$$E[Z(\mathbf{D}', \mathcal{F}_n)] - E[Z(\mathbf{D}', \mathcal{D}_n)] = E[S(\mathbf{D}', \mathcal{D}_n)] - E[S(\mathbf{D}', \mathcal{F}_n)] \quad (2.1)$$

will be a constant, $\forall \Delta_k \geq 0$. This is because for the normal distribution the shortfall of either full flexibility or dedicated design will be in the form of

$$E[D - \mu]^+ = \frac{\sigma}{\sqrt{2\pi}}, \quad (2.2)$$

which only relates to σ . But for any demand sample path $(\mu_1 + x_1\sigma_1, \dots, \mu_n + x_n\sigma_n)$, the shortfall of $\mathcal{G}$ after increasing in μ_k is no more than before due to the increase of transit capacity. That is $E[S(\mathbf{D}', \mathcal{G})] \leq E[S(\mathbf{D}, \mathcal{G})]$. Thus

$$E[Z(\mathbf{D}', \mathcal{G})] - E[Z(\mathbf{D}', \mathcal{D}_n)] = E[S(\mathbf{D}', \mathcal{D}_n)] - E[S(\mathbf{D}', \mathcal{G})] \quad (2.3)$$

$$\geq E[S(\mathbf{D}, \mathcal{D}_n)] - E[S(\mathbf{D}, \mathcal{G})] \quad (2.4)$$

Consequently $f(\boldsymbol{\mu} + \boldsymbol{\Delta}, \boldsymbol{\sigma}, \mathcal{G}) \geq f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{G})$, where $\boldsymbol{\Delta} = [0, \dots, \Delta_k, \dots, 0]$. By the same token, the three-point distribution can be proved. With regard to the three-point distribution, note that keeping variance unchanged means $t\eta^2$ is a constant. $\square$

Theorem 2.4.2 can be viewed as a result of the relationship between the coefficient of variation and the flexibility design efficiency from the ‘ $\boldsymbol{\mu}$ ’ perspective. Consistent with the numerical demonstration in the literature (Chou et al., 2010), Theorem 2.4.2 indicates that reducing CV by expanding ‘ $\boldsymbol{\mu}$ ’ improves the flexibility design efficiency. In other words, a flexibility design benefits from any increase of the capacity and the expected demand, especially of the bottleneck pair, because of the risk pooling effect of the flexible capacity.

Corollary 2.4.1. *Consider a balanced system of size n . Given any flexibility*

design $\mathcal{G}$ such that the dedicated design $\mathcal{D}_n \subseteq \mathcal{G}$, for the normal distribution $f(\min_i \mu_i \cdot \mathbf{1}, \boldsymbol{\sigma}, \mathcal{G}) \leq f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{G}) \leq f(\max_i \mu_i \cdot \mathbf{1}, \boldsymbol{\sigma})$; for the symmetric three-point distribution, $f(\min_i \mu_i \cdot \mathbf{1}, \boldsymbol{\eta}, \mathbf{t}, \mathcal{G}) \leq f(\boldsymbol{\mu}, \boldsymbol{\eta}, \mathbf{t}, \mathcal{G}) \leq f(\max_i \mu_i \cdot \mathbf{1}, \boldsymbol{\eta}, \mathbf{t}, \mathcal{G})$.

Corollary 2.4.1 shows that the flexibility design efficiency is bounded below by a system with the same variance but smaller $\boldsymbol{\mu}$. If the gap between $\min_i \mu_i$ and other μ_i is large, the bound could be very loose. However, the bound will be tighter and tighter as the gap narrows.

Next we study the effect of variance on the flexibility design efficiency for balanced and identical systems. We use μ instead of $\boldsymbol{\mu}$ and so do σ , η and t for simplicity.

Proposition 2.4.1. *Consider a balanced and identical system of size n . Given any flexibility design $\mathcal{G}$ such that the dedicated design $\mathcal{D}_n \subseteq \mathcal{G}$ and holding μ unchanged, $f(\mu, \sigma, \mathcal{G})$ and $f(\mu, \eta, t, \mathcal{G})$ are decreasing in σ and η for the normal and the symmetric three-point distributions, respectively.*

Proof Since for an $n \times n$ identical system with $D_i \sim N(\mu, \sigma^2)$, we can always scale it to another system with $D_i \sim N(1, \sigma^2/\mu^2)$ and capacity $c_i = 1$. So $f(\mu, \sigma, \mathcal{G}) = f(1, \sigma/\mu, \mathcal{G})$. As a consequence, $f(\mu, \sigma + \Delta\sigma, \mathcal{G}) = f(1, (\sigma + \Delta\sigma)/\mu, \mathcal{G}) = f(\sigma/(\sigma + \Delta\sigma), \sigma/\mu, \mathcal{G})$, where $\Delta\sigma$ is nonnegative. According to Theorem 2.4.2, $f(\sigma/(\sigma + \Delta\sigma), \sigma/\mu, \mathcal{G}) \leq f(1, \sigma/\mu, \mathcal{G})$. The three-point distribution can be shown in the same fashion by fixing t $\square$

Proposition 2.4.1 investigates the relationship between demand variation and the flexibility design efficiency for balanced and identical systems. It states that an identical system with less demand variation has a higher flexibility design efficiency.

In what follows, we examine the effect of changes in demand variance on the long chain efficiency. For balanced but non-identical systems, this analysis proves to be unwieldy, especially for the case of the normal distribution. To obtain some insights as well as to isolate the effect of demand variance, we consider the symmetric three-point distribution with homogeneous support but heterogeneous tail probabilities, i.e. $\mu_i = \mu, \eta_i = \eta = \mu$, for all $i = 1, 2, \dots, n$. Without loss of generality, we can further assume $\mu = 1$. Hence, we have

$$D_i = \begin{cases} 0, & \text{with probability } t_i \\ 1, & \text{with probability } 1 - 2t_i \\ 2, & \text{with probability } t_i \end{cases}$$

As such, we can simplify the notation for the long chain efficiency into $f(\mathbf{t}, \mathcal{C}_n)$. Our interest is to see how $f(\mathbf{t}, \mathcal{C}_n)$ behaves as t_k increases or decreases. To this end, we must first characterize the performances $[\mathcal{D}_n]$, $[\mathcal{C}_n]$, and $[\mathcal{F}_n]$ of the dedicated, long chain, and full flexibility designs, respectively.

From here onwards, whenever the integer k appears in a statement, we are actually referring to some $i \in \{1, 2, \dots, n\}$ congruent to k modulo n . For instance, plant $n + 1$ refers to plant 1, while product 0 refers to product n . That said, we first present the following preliminary results that will help prove our subsequent main results.

Lemma 2.4.1. *For any balanced system of size n , we have*

$$[\mathcal{C}_n] = \sum_{i=1}^n ([\mathcal{C}_n \setminus \{(i, i + 1)\}] - [\mathcal{C}_n \setminus \{(i - 1, i), (i, i), (i, i + 1)\}]).$$

Proof This is straight-forward from Theorem 3 in Simchi-Levi and Wei (2012).

□

Lemma 2.4.2. *Consider any balanced system of size n with the symmetric three-point distribution with $\mu_i = 1, \eta_i = 1$, for all $i = 1, 2, \dots, n$. For any $i = 1, 2, \dots, n$, we have*

$$[\mathcal{C}_n \setminus \{(i, i+1)\}] - [\mathcal{C}_n \setminus \{(i-1, i), (i, i), (i, i+1)\}] = \mathbf{q}^0 \left(\prod_{k=1, k \neq i}^n \mathbf{A}_k \right) \boldsymbol{\pi}^i$$

where $\mathbf{q}^0 = [1, 0]$, $\mathbf{A}_k = \begin{bmatrix} 1-t_k & t_k \\ t_k & 1-t_k \end{bmatrix}$, and $\boldsymbol{\pi}^i = [1-t_i, 1]^T$.

Proof The proof is similar to those of Lemma 3 and Theorem 5 in Simchi-Levi and Wei (2012). □

Lemma 2.4.3. *For $j \geq i$, $\Pi_{k=i}^j \mathbf{A}_k = \begin{bmatrix} \frac{1+\Pi_{k=i}^j(1-2t_k)}{2} & \frac{1-\Pi_{k=i}^j(1-2t_k)}{2} \\ \frac{1-\Pi_{k=i}^j(1-2t_k)}{2} & \frac{1+\Pi_{k=i}^j(1-2t_k)}{2} \end{bmatrix}$.*

Proof We prove this result by induction. First, observe that the lemma is true for $j = i$ by definition of $\mathbf{A}_i$. Suppose the lemma is true for $j = j_0$. We want to show that it is also true for $j = j_0 + 1$.

$$\begin{aligned} \Pi_{k=i}^{j_0+1} \mathbf{A}_k &= \Pi_{k=i}^{j_0} \mathbf{A}_k \cdot \mathbf{A}_{j_0+1} \\ &= \begin{bmatrix} \frac{1+\Pi_{k=i}^{j_0}(1-2t_k)}{2} & \frac{1-\Pi_{k=i}^{j_0}(1-2t_k)}{2} \\ \frac{1-\Pi_{k=i}^{j_0}(1-2t_k)}{2} & \frac{1+\Pi_{k=i}^{j_0}(1-2t_k)}{2} \end{bmatrix} \cdot \begin{bmatrix} 1-t_{j_0+1} & t_{j_0+1} \\ t_{j_0+1} & 1-t_{j_0+1} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1+\Pi_{k=i}^{j_0+1}(1-2t_k)}{2} & \frac{1-\Pi_{k=i}^{j_0+1}(1-2t_k)}{2} \\ \frac{1-\Pi_{k=i}^{j_0+1}(1-2t_k)}{2} & \frac{1+\Pi_{k=i}^{j_0+1}(1-2t_k)}{2} \end{bmatrix} \end{aligned}$$

□

Lemma 2.4.4. *Consider a balanced system of size n , For the symmetric three-point distribution with $\mu_i = 1, \eta_i = 1$, for all $i = 1, 2, \dots, n$, we have*

$$[\mathcal{F}_n] = [\mathcal{F}_{n-1}] + 1 - t_n \cdot P_{n-1}(n-1).$$

where $P_k(w) = \Pr\{\sum_{i=1}^k D_i = w\}$ is the probability that the total demand for the first k products is w .

Proof By conditioning on D_n , we get

$$\begin{aligned} [\mathcal{F}_n] &= [\mathcal{F}_{n-1}] + \Pr\{D_n = 0\} \cdot \Pr\left\{\sum_{i=1}^{n-1} D_i > n-1\right\} \cdot 1 + \Pr\{D_n = 1\} \cdot 1 \\ &\quad + \Pr\{D_n = 2\} \cdot \left[\Pr\left\{\sum_{i=1}^{n-1} D_i \geq n-1\right\} \cdot 1 + \Pr\left\{\sum_{i=1}^{n-1} D_i < n-1\right\} \cdot 2\right] \\ &= [\mathcal{F}_{n-1}] + t_n \cdot \frac{1 - P_{n-1}(n-1)}{2} + (1 - 2t_n) \\ &\quad + t_n \cdot \left[\frac{1 + P_{n-1}(n-1)}{2} + \frac{1 - P_{n-1}(n-1)}{2} \cdot 2\right] \\ &= [\mathcal{F}_{n-1}] + 1 - t_n \cdot P_{n-1}(n-1) \end{aligned}$$

The second equation is due to the symmetry of the distribution of $\sum_{i=1}^{n-1} D_i$. □

It is now time to present our next main result, which characterizes the performances of the dedicated, long chain, and full flexibility designs.

Theorem 2.4.3. *Consider a balanced system of size n . For the symmetric three-point distribution with $\mu_i = 1, \eta_i = 1$, for all $i = 1, 2, \dots, n$, we have*

$$(a) \ [\mathcal{D}_n] = n - \sum_{i=1}^n t_i$$

(b) $[\mathcal{C}_n] = n - \frac{1}{2} \sum_{i=1}^n t_i - \frac{1}{2} K(n) \sum_{i=1}^n \frac{t_i}{1-2t_i}$, where $K(n) = \prod_{i=1}^n (1 - 2t_i)$ is the probability that all products have demand equal to the mean.

(c) $[\mathcal{F}_n] = n - \sum_{i=1}^n t_i \cdot P_{i-1}(i-1)$.

Proof For (a), the result follows easily from the fact that $[\mathcal{D}_n] = \sum_{i=1}^n E[\min(D_i, 1)]$.

For (b), we combine Lemma 2.4.1 and Lemma 2.4.2 to obtain $[\mathcal{C}_n] = \sum_{i=1}^n \mathbf{q}^0 \Pi_{k=1, k \neq i}^n \mathbf{A}_k \boldsymbol{\pi}^k$.

By applying Lemma 2.4.3 and performing standard matrix multiplication, we get

$$[\mathcal{C}_n] = n - \frac{1}{2} \sum_{i=1}^n t_i - \frac{1}{2} \sum_{i=1}^n t_i \cdot \prod_{k=1, k \neq i}^n (1 - 2t_k).$$

By definition of $K(n)$, the results follows.

For (c), the result follows by applying Lemma 2.4.4 recursively and because $[\mathcal{F}_1] = 1$. $\square$

Corollary 2.4.2. *Consider a balanced system of size n . For the symmetric three-point distribution with $\mu_i = 1, \eta_i = 1$, for all $i = 1, 2, \dots, n$, we have the following partial derivatives and difference equations.*

- (a) $\frac{\partial}{\partial t_n} [\mathcal{D}_n] = -1$
- (b) $\frac{\partial}{\partial t_n} [\mathcal{C}_n] = -\frac{1}{2} - \frac{1}{2} K(n-1) + K(n-1) \sum_{i=1}^{n-1} \frac{t_i}{1-2t_i}$
- (c) $\frac{\partial}{\partial t_n} [\mathcal{F}_n] = -P_{n-1}(n-1)$
- (d) $[\mathcal{D}_n] - [\mathcal{D}_{n-1}] = 1 + t_n \cdot \frac{\partial}{\partial t_n} [\mathcal{D}_n] = 1 - t_n$
- (e) $[\mathcal{C}_n] - [\mathcal{C}_{n-1}] = 1 + t_n \cdot \frac{\partial}{\partial t_n} [\mathcal{C}_n] = 1 + t_n \left(-\frac{1}{2} - \frac{1}{2} K(n-1) + K(n-1) \sum_{i=1}^{n-1} \frac{t_i}{1-2t_i} \right)$
- (f) $[\mathcal{F}_n] - [\mathcal{F}_{n-1}] = 1 + t_n \cdot \frac{\partial}{\partial t_n} [\mathcal{F}_n] = 1 - t_n \cdot P_{n-1}(n-1),$

Although in Corollary 2.4.2 we only provide the results for index n , we can actually relabel the product-plant pairs to make the pair of interest be the n th pair.

For example, we can relabel pairs $1, \dots, n$ in a long chain to pairs $n, 1, \dots, n-1$. Then we have the same results for the new n th pair which is pair 1 in the initial index system. Of course, the meaning of $[\mathcal{D}_{n-1}]$ in a index system differs from the one in another index system and so do $[\mathcal{C}_{n-1}]$ and $[\mathcal{F}_{n-1}]$. Next, we present an interesting connection between the sensitivity of the long chain efficiency to demand variance (represented by tail probability) and the sensitivity of the long chain efficiency to system size.

Theorem 2.4.4. *Consider a balanced system of size n . For the symmetric three-point distribution with $\mu_i = 1, \eta_i = 1$, for all $i = 1, 2, \dots, n$, we have $f(\mathbf{t}, \mathcal{C}_n)$ is decreasing in t_i if and only if $f(\mathbf{t}, \mathcal{C}_n)$ is decreasing in n .*

Proof Without loss of generality, we can take $i = n$. By definition of $f(\mathbf{t}, \mathcal{C}_n)$ and using Corollary 2.4.2 (d)-(f), we have

$$\begin{aligned}
 & \frac{\partial}{\partial t_n} f(\mathbf{t}, \mathcal{C}_n) \leq 0 \\
 \Leftrightarrow & \frac{\partial}{\partial t_n} \frac{[\mathcal{C}_n] - [\mathcal{D}_n]}{[\mathcal{F}_n] - [\mathcal{D}_n]} \leq 0 \\
 \Leftrightarrow & \frac{([\mathcal{F}_n] - [\mathcal{D}_n]) \cdot \frac{\partial}{\partial t_n} ([\mathcal{C}_n] - [\mathcal{D}_n]) - ([\mathcal{C}_n] - [\mathcal{D}_n]) \cdot \frac{\partial}{\partial t_n} ([\mathcal{F}_n] - [\mathcal{D}_n])}{([\mathcal{F}_n] - [\mathcal{D}_n])^2} \leq 0 \\
 \Leftrightarrow & ([\mathcal{F}_n] - [\mathcal{D}_n]) \cdot \frac{([\mathcal{C}_n] - [\mathcal{D}_n]) - ([\mathcal{C}_{n-1}] - [\mathcal{D}_{n-1}])}{t_n} \\
 & - ([\mathcal{C}_n] - [\mathcal{D}_n]) \cdot \frac{([\mathcal{F}_n] - [\mathcal{D}_n]) - ([\mathcal{F}_{n-1}] - [\mathcal{D}_{n-1}])}{t_n} \leq 0 \\
 \Leftrightarrow & - ([\mathcal{F}_n] - [\mathcal{D}_n])([\mathcal{C}_{n-1}] - [\mathcal{D}_{n-1}]) + ([\mathcal{C}_n] - [\mathcal{D}_n])([\mathcal{F}_{n-1}] - [\mathcal{D}_{n-1}]) \leq 0 \\
 \Leftrightarrow & \frac{[\mathcal{C}_n] - [\mathcal{D}_n]}{[\mathcal{F}_n] - [\mathcal{D}_n]} \leq \frac{[\mathcal{C}_{n-1}] - [\mathcal{D}_{n-1}]}{[\mathcal{F}_{n-1}] - [\mathcal{D}_{n-1}]}
 \end{aligned}$$

Since the meanings of $[\mathcal{D}_{n-1}]$, $[\mathcal{C}_{n-1}]$ and $[\mathcal{F}_{n-1}]$ depend on the index system,

for $i \neq n$ we can again relabel the product-plant pairs to make the initial i th pair be the n th pair. Then we repeat the above proof and the result follows. $\square$

The analysis involved to show that the long chain efficiency decreases in system size has proven unwieldy. In fact, previous works (Chou et al., 2010; Simchi-Levi and Wei, 2012) are unable to prove this relationship even for the simplest case of balanced and identical systems. Nonetheless, several numerical studies have supported the belief that this relationship holds. To conduct similar numerical studies for our balanced but non-identical case, we first note the following recursive equations for $P_k(w)$. For $k = 1, \dots, n$,

$$\begin{aligned} P_k(1) &= (1 - 2t_k)P_{k-1}(0) + t_k P_{k-1}(1) \\ P_k(w) &= (1 - 2t_k)P_{k-1}(w - 1) + t_k P_{k-1}(w) + t_k P_{k-1}(w - 2), \quad \forall w = 2, \dots, k - 1 \\ P_k(k) &= (1 - 2t_k)P_{k-1}(k - 1) + 2t_k P_{k-1}(k - 2) \end{aligned}$$

with boundary conditions $P_0(0) = 1$, and $P_k(0) = \prod_{i=1}^k t_i$. For completeness, we can obtain the rest of the probabilities with $P_k(w) = P_k(2k - w), \forall w = k + 1, k + 2, \dots, 2k$.

The above recursive equations allow us to calculate $P_k(k)$ for all $k = 1, 2, \dots, n$, which in turn allows us to calculate $[\mathcal{F}_n]$ according to Theorem 2.4.3. From the same theorem, the values of $[\mathcal{D}_n]$ and $[\mathcal{C}_n]$ are more straight-forward to calculate. For different system sizes $n = 2, 3, \dots, 20$, we randomly generated 10,000 sets of tail probabilities $t_i \in (0, 0.5), \forall i = 1, 2, \dots, n$ and calculated the long chain efficiency $f(\mathbf{t}, \mathcal{C}_n)$. For all instances, we find that $f(\mathbf{t}, \mathcal{C}_n)$ is, indeed, decreasing in n . Hence, we also have good reason to believe that $f(\mathbf{t}, \mathcal{C}_n)$ is decreasing in t_n

(equivalently, demand variance).

The next result follows easily from Theorem 2.4.4, and also provides upper and lower bounds on the long chain efficiency.

Corollary 2.4.3. *Consider a balanced system of size n . For the symmetric three-point distribution with $\mu_i = 1, \eta_i = 1$ for all $i = 1, 2, \dots, n$, suppose $f(\mathbf{t}, \mathcal{C}_n)$ is decreasing in n . Then, $f(\max_i t_i \cdot \mathbf{1}, \mathcal{C}_n) \leq f(\mathbf{t}, \mathcal{C}_n) \leq f(\min_i t_i \cdot \mathbf{1}, \mathcal{C}_n)$.*

For our next result, we refer to the edges in $\mathcal{G}$ as flexibility links. We then examine the efficiency of link usage for products with low demand variance. Interestingly, we find that link efficiency decreases and goes to zero as variance decreases and goes to zero. This implies that when a product has very low demand variance, including it in the long chain does not cause any harm to the system performance, but the benefit is minimal, resulting in a poorly utilized flexibility link.

Theorem 2.4.5. *Consider a balanced system of size n with two flexibility designs, namely, $\mathcal{C}_n$ and $\mathcal{C}_{n-1} \cup \{(n, n)\}$. For the symmetric three-point distribution with $\mu_i = 1, \eta_i = 1$ for all $i = 1, 2, \dots, n$. The difference in performance $[\mathcal{C}_n] - [\mathcal{C}_{n-1} \cup \{(n, n)\}]$ is increasing in t_n and equals 0 when $t_n = 0$.*

Proof It is easy to see that $[\mathcal{C}_{n-1} \cup \{(n, n)\}] = [\mathcal{C}_{n-1}] + (1 - t_n)$. From Corollary 2.4.2(e), we have

$$[\mathcal{C}_n] - [\mathcal{C}_{n-1}] - (1 - t_n) = t_n \left(\frac{1}{2} - \frac{1}{2} K(n-1) + K(n-1) \sum_{i=1}^{n-1} \frac{t_i}{1 - 2t_i} \right).$$

Except in the trivial case when $t_i = 0$ for all $i = 1, 2, \dots, n-1$, the coefficient of t_n above is strictly positive and independent of t_n . Hence, the result follows. $\square$

2.5 Bounds on Flexibility Design Efficiency

The above results have supported the insights that we gain from Figure 2.1.1. We have shown in Theorem 2.4.1 that a system with larger demand variation has more potential. It implies that we should pay more attention to pairs with larger demand variance. Meanwhile, Theorem 2.4.5 indicates that combining pairs with small demand variance into flexibility design brings us little value for the specific symmetric three-point distribution. These findings suggest isolating the pairs with relatively low variance and assigning them dedicated design, though we are not able to analytically prove Theorem 2.4.5 for other distributions. So we can preserve links to install flexibility in other places. Second, Theorem 2.4.2 tells us that a product with low demand mean can create a bottleneck if included in a long chain design. Therefore, we will likely be better off with having a few smaller chains where each chain will have relatively similar demand means. However, by doing this we might lose output due to a reduction in risk pooling. In fact, there exists no risk pooling at all between products from different small chains. To avoid the reduction of risk pooling, we use the saving of links from the pairs with low demand variance to bridge those shorter chains.

Then the question arises is how to divide a long chain into shorter chains. To address this, we again turn to the bounds on the flexibility design efficiency like in Corollary 2.4.1. We attempt to combine Corollary 2.4.1 and Corollary 2.4.3. We do so numerically and find that for the normal distribution and symmetric

2.5 Bounds on Flexibility Design Efficiency

three-point distribution, our numerical examples consistently show that

$$f(\min_i \mu_i \cdot \mathbf{1}, \max_i \sigma_i \cdot \mathbf{1}, \mathcal{C}_n) \leq f(\min_i \mu_i \cdot \mathbf{1}, \boldsymbol{\sigma}, \mathcal{C}_n) \leq f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{C}_n)$$

$$f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{C}_n) \leq f(\max_i \mu_i \cdot \mathbf{1}, \boldsymbol{\sigma}, \mathcal{C}_n) \leq f(\max_i \mu_i \cdot \mathbf{1}, \min_i \sigma_i \cdot \mathbf{1}, \mathcal{C}_n).$$

For the symmetric three-point distribution we first randomly sample demand mean μ_i in $[50, 100]$ and then sample η_i in $[0, 50]$ and t_i from $[0, 1/2]$. By doing this we avoid the negative demand. For the normal distribution we also sample demand mean μ_i in $[50, 100]$ and then sample σ_i from interval $[0, 25]$. For each distribution we sample 100 scenarios for 20×20 system.

Figure 2.5.1 and 2.5.2 respectively show the results of the three-point and the normal distribution and they are consistent with our conjecture. In Figure 2.5.2 the lowest, lower, original, upper, highest bounds correspond to $f(\min_i \mu_i \cdot \mathbf{1}, \max_i \sigma_i \cdot \mathbf{1}, \mathcal{C}_n)$, $f(\min_i \mu_i \cdot \mathbf{1}, \boldsymbol{\sigma}, \mathcal{C}_n)$, $f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{C}_n)$, $f(\max_i \mu_i \cdot \mathbf{1}, \boldsymbol{\sigma}, \mathcal{C}_n)$, $f(\max_i \mu_i \cdot \mathbf{1}, \min_i \sigma_i \cdot \mathbf{1}, \mathcal{C}_n)$, respectively. We can see that the lower bound $f(\min_i \mu_i \cdot \mathbf{1}, \boldsymbol{\sigma}, \mathcal{C}_n)$ is already very close to the line for $f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{C}_n)$. And all bounds fluctuate in the same directions as the original one. Interestingly, for an identical system, the four bounds will finally converge to $f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{C}_n)$. We have similar observations for the three-point distribution.

Hence $f(\min_i \mu_i \cdot \mathbf{1}, \max_i \sigma_i \cdot \mathbf{1}, \mathcal{C}_n)$, to some extent, signals the magnitude of the long chain efficiency for the normal distribution. High $f(\min_i \mu_i \cdot \mathbf{1}, \max_i \sigma_i \cdot \mathbf{1}, \mathcal{C}_n)$ guarantees high $f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathcal{C}_n)$, though not vice versa.

We have shown that the long chain efficiency of a balanced but non-identical

2.5 Bounds on Flexibility Design Efficiency

Figure 2.5.1: The three-point distribution

system is bounded below by that of an identical system. The demand mean of the identical system is equal to $\min_i \mu_i$ of the non-identical system and the standard deviation is equal to $\max_i \sigma_i$ of the non-identical system. According to Theorem 2.4.1, the bound increases in $\max_i \sigma_i / \min_i \mu_i$. This gives us the insight that we can use an index analogous to coefficient of variation to check whether a long chain has already achieved most of the benefits of full flexibility. More importantly, we can use this index to guide us to divide the long chain into several short chains with high efficiency. We introduce the definition of group coefficient of variation.

Definition 2.5.1. *For a balanced but nonidentical system or subsystem G , we*

2.5 Bounds on Flexibility Design Efficiency

Figure 2.5.2: The normal distribution

define its **group coefficient of variation** as follows.

$$GCV(G) = \frac{\max_{i \in G} \sigma_i}{\min_{i \in G} \mu_i}$$

There are two reasons to use $\max_{i \in G} \sigma_i / \min_{i \in G} \mu_i$ rather than $\sum_{i \in G} \sigma_i / \sum_{i \in G} \mu_i$ or some other indices. First, when GCV is small, it clearly indicates that the bound $f(\min_i \mu_i \times \mathbf{1}, \max_i \sigma_i \times \mathbf{1}, \mathcal{C}_n)$ is already good enough and in turn signals good long chain efficiency. Second, when GCV is high, we can use the denominator of GCV to identify which pair is a bottleneck for the long chain, like in Figure 2.1.1(a). So we can exclude this pair from the long chain and increase the efficiency of the long chain. Meanwhile, the numerator shows the pair with the

largest σ_i which should be dealt with separately.

2.6 The Hub-and-Chain Method

For a balanced but non-identical system, i.e. $m = n$, we seek a flexibility design $\mathcal{G} \subseteq A \times B$ that maximizes $[\mathcal{G}]$ such that $|\mathcal{G}| \leq b = 2n$. As mentioned earlier, this is a difficult stochastic mixed integer linear program, which leads us to focus on heuristics to find good enough flexibility designs within reasonable computational time. Clearly, choosing $\mathcal{G} = \mathcal{C}_n$ involves negligible computational time and is a feasible design. However, as explained in the previous section, a product with low demand mean might cause a bottleneck and substantially reduce the effectiveness of this long chain design. As such, we might be better served using smaller chains with a provision to link together these small chains. And therein lies the inspiration for our proposed method.

In what follows, we present our variance-based hub-and-chain method, or VHC for short. The main idea is to partition the product set A into $s + 1$ subsets $G_0, G_1, \dots, G_s$ such that the subsystem corresponding to G_0 will employ a dedicated design, while the subsystems corresponding to $G_1, \dots, G_s$ will each employ a long chain design. Then, we choose a product-plant pair in each subsystem $G_1, \dots, G_s$ and call it the *satellite pair* for that subsystem. Finally, we use the links freed up by G_0 to connect the satellite pair of G_1 to those of $G_2, \dots, G_s$. We denote the resulting flexibility design as $\mathcal{H}(G_0, G_1, \dots, G_s)$.

Our problem becomes to maximize $[\mathcal{H}(G_0, G_1, \dots, G_s)]$ such that $\bigcup_{i=0}^s G_i = A$ and $G_i \cap G_j = \emptyset, \forall i, j = 0, 1, \dots, s$. This is still a difficult problem to solve for

at least two reasons. First, the objective function is the expected optimal value of another optimization problem. As a remedy, we can consider a substitute objective function which is to minimize the maximum of the GCVs of the subsets $G_1, \dots, G_s$. That is, we want to partition such that the worst group variability of all subsystems is minimized.

$$\begin{aligned}
& \min \quad z \\
& \text{s.t.} \quad \sum_{j=0}^s y_{ij} = 1, \quad \forall i = 1, \dots, n \\
& \quad \quad z \geq \text{GCV}(G_j) = \frac{\max_{i=1}^n \sigma_i y_{ij}}{\min_{i=1}^n \{\mu_i y_{ij} | y_{ij} = 1\}} \quad \forall j = 1, \dots, s \\
& \quad \quad z \geq \frac{\sigma_i}{\mu_i} \cdot y_{i0} \quad \forall i = 1, \dots, n \\
& \quad \quad s = 1 + \frac{1}{2} \sum_{i=1}^n y_{i0} \\
& \quad \quad y_{ij} \in \{0, 1\}, \quad \forall i = 1, \dots, n, \forall j = 0, 1, \dots, s \\
& \quad \quad z \geq 0; s \geq 1 \text{ and integer}
\end{aligned}$$

where y_{ij} decides whether product i is included in subset G_j or not. In addition, the third and fourth constraints ensure that we do not include too many products into G_0 . The third constraint prefers that products with low demand variability be isolated, while the fourth constraint ensures that the total number of flexibility links used stays at $2n$.

The second reason why our problem is difficult is that the integer s is also a decision variable, which depends on the number of products selected into the

2.6 The Hub-and-Chain Method

set G_0 . In fact, even if we fix s , we are still left with a difficult binary nonlinear program. To avoid the long computational time associated with the above optimization problem, we propose the VHC method detailed in Algorithm 1. Recall that A is the product set, $GCV(G_j)$ is according to Definition 2.5.1, and let $G(i)$ be the i th smallest element in set G .

The VHC method involves three main steps. In Step 1, products with very low demand variance will be isolated into the subset G_0 , in which only a dedicated flexibility design will be employed. The decision on how many such products will be chosen depends on two pre-specified parameters $\theta_1, \theta_2 \in (0, 1)$. θ_1 refers to the maximum standard deviation for a single product relative to total standard deviation for all products in order for that single product to be relegated to G_0 . On the other hand, θ_2 refers to the maximum proportion of total standard deviation to be allowed for a group of products to be relegated to G_0 . Clearly, higher values of θ_1 and θ_2 will result in more products being included in subset G_0 . All remaining products will be placed in subset G_1 . In Step 2, the subset G_1 will be further divided into more subsets $G_1, G_2, \dots, G_s$ such that the GCV of each subset does not exceed a pre-specified threshold value $\theta_3 \in (0, 1)$. Our choice of θ_3 will obviously influence the number s of subsets as well as the number of elements in each subset. Finally, in Step 3, we form the hub-and-chain design. We first let the subsystem G_0 employ a dedicated design and each of the subsystems $G_1, \dots, G_s$ employ a long chain design. For each non-dedicated subsystem, we let the product-plant pair with the largest demand variance be the satellite pair for that subsystem. Then, we label the subsystem G_1 as the hub whose satellite pair will be connected (via two links) to the satellite pair of every other subsystem.

tem $G_2, \dots, G_s$. Note that the reason for choosing pairs with maximum demand variance as satellite pairs is because they are the most likely sources of excess demand and idle capacity. Furthermore, this is in line with the discussion at the end of Section 2.5 that high-variance products should be provided more access to additional capacity.

Having said that, we now demonstrate how to implement the VHC method and benchmark against two existing methods; namely, the long chain design and the constraint sampling method developed in Chou et al. (2010). Before we proceed, we observe that the number of flexibility links utilized by the hub-and-chain design is not always equal to $2n$, but largely depends on our choice of values for the parameters θ_1, θ_2 , and θ_3 . However, in most instances, the number of links deviates from $2n$ by only a small number. On one hand, a firm with no hard budget on flexibility links may want to just run the VHC method using some pre-tested parameters values. On the other hand, a firm with a strict budget of $2n$ links may still use the VHC method as a subroutine in a broader iterative method. We discuss both cases in the next subsections.

2.6.1 Method 1 - No Budget Specified

Here we present the heuristic to construct flexibility design for a balanced $n \times n$ system without a specified budget. The main steps are presented in Algorithm 1.

Let G_j denote an index vector of a group of pairs. n_j is the number of its elements. We initialize $G_1 = (1, \dots, n)$, $n_1 = n$ and $G_j = \emptyset$, $n_j = 0$, $\forall j > 1$. We use another index vector $G_0 = \emptyset$ to accept pairs associated with unacceptably low

variability. $\mathcal{G} = \mathbf{0}$ is a $n \times n$ matrix to record the sparse structure. Assignment of a value of 1 to $\mathcal{G}(i, j)$ means a link from product i to plant j has been established.

We rely on three thresholds θ_1 , θ_2 , θ_3 to assort the pairs. These threshold values were selected after a few rounds of calibration. θ_1 is the threshold of standard deviation in order to decide which pair should be expelled out of main structure and put into *dedicated* group while θ_2 is the maximum proportion of standard deviation that allowed to be separated. Of course θ_1 is not compulsory. We can solely use θ_2 to control the proportion of total standard deviation to be split and each time expel the pair associated with minimum σ . Let x represent the portion of standard deviation has already been picked out. θ_3 is the threshold for $\frac{\max_{i \in G_j} \sigma_i}{\min_{i \in G_j} \mu_i}$, the GCV of G_j . Once GCV is still greater than θ_3 , the pair associated with minimum μ of G_j is taken out and put into G_{j+1} .

2.6.2 Method 2 - Set Partitioning with Budget

In this subsection, we have a hard budget of links B , e.g. only 40 links for a 20×20 system, for $n \times n$ system. To obtain the optimal hub-and-chain design, we implement the following iterative method which uses Steps 2 and 3 of Algorithm 1 as a subroutine. That is, we no longer need the parameters θ_1 and θ_2 to determine the number $|G_0|$ of dedicated product-plant pairs. Instead, we initialize $|G_0| = 2$, and later incrementally add 2 more products to G_0 , i.e. $|G_0| \leftarrow |G_0| + 2$. For each value of $|G_0|$, we iteratively run Steps 2 and 3 of Algorithm 1, starting with $\theta_3 = \max_{i=1}^n \frac{\sigma_i}{\mu_i}$. If the hub-and-chain structure has $s = \lfloor (B + |G_0| - 2n)/2 + 1 \rfloor$ groups, stop. Otherwise, increase θ_3 by a step size of $\Delta\theta_3$ and repeat the previous

Algorithm 1 Balanced Case

Step 1: Pick products with unqualified variance.

- 1: Initialize $x \leftarrow 0$; $n_1 \leftarrow n$;
- 2: **while** $\min_{i \in G_1} \sigma_i / \sum_{i=1}^n \sigma_i < \theta_1$ and $x + \min_{i \in G_1} \sigma_i / \sum_{i=1}^n \sigma_i < \theta_2$ **do**
- 3: $x \leftarrow x + \min_{i \in G_1} \sigma_i / \sum_{i=1}^n \sigma_i$; $G_1 \leftarrow G_1 \setminus \{\arg \min_{i \in G_1} \sigma_i\}$ and $G_0 \leftarrow G_0 \cup \{\arg \min_{i \in G_1} \sigma_i\}$; $n_1 \leftarrow n_1 - 1$;
- 4: **end while**

Step 2: Assort the remaining products.

- 1: Initialize $j \leftarrow 1$; $n_{j+1} \leftarrow 0$;
- 2: **while** $G_j \neq \emptyset$ **do**
- 3: **while** $\frac{\max_{i \in G_j} \sigma_i}{\min_{i \in G_j} \mu_i} > \theta_3$ **do**
- 4: $l^* \leftarrow \arg \min_{i \in G_j} \sigma_i$; $G_j \leftarrow G_j \setminus \{l^*\}$; $G_{j+1} \leftarrow G_{j+1} \cup \{l^*\}$; $n_{j+1} \leftarrow n_{j+1} + 1$;
- 5: **end while**
- 6: $j \leftarrow j + 1$; initialize $n_{j+1} \leftarrow 0$;
- 7: **end while**
- 8: $J \leftarrow j - 1$.

Step 3: Forge the hub-and-chain structure.

- 1: **for** $j = 1$ to J **do**
 - 2: **for** $i = 1$ to $n_j - 1$ **do**
 - 3: $\mathcal{G}(G_j(i), G_j(i)) \leftarrow 1$; $\mathcal{G}(G_j(i), G_j(i+1)) \leftarrow 1$;
 - 4: **end for**
 - 5: $\mathcal{G}(G_j(n_j), G_j(1)) \leftarrow 1$; $l_j^* \leftarrow \arg \max_{i \in G_j} \sigma_i$; $\mathcal{G}(G_1(l_1^*), G_j(l_j^*)) \leftarrow 1$;
 - 6: $\mathcal{G}(G_j(l_j^*), G_1(l_1^*)) \leftarrow 1$;
 - 7: **end for**
-

process. The reason of starting from $\max_{i=1}^n CV_i$ is to make sure at least two sub-groups can be obtained. s is derived from the rationale as follows. s groups need $2(s-1)$ links to form a hub-and-chain structure and another $2(n - |G_0|)$ is necessary to form the long chain structure in each sub-group, supposing at least 2 pairs in each subgroup. Thus $2(n - |G_0|) + 2(s-1) \leq B - I$ which is equivalent to $s \leq (B + |G_0| - 2n)/2 + 1$. While a binary search might be faster, this line search for θ_3 does not depend on the system size n , hence is already

computationally very efficient. Finally, we choose the best among the hub-and-chain designs generated by different values of $|G_0|$ based on their performances according to 10,000 randomly selected demand realizations.

2.7 Simulation and Analysis

Next we simulate a 20×20 system. The simulator generates randomly a mean for each product within the range of $[2, 500]$ and based on the mean then sample a standard deviation in the range $[0, \mu/2]$ such that coefficient of variation at most is 0.5. For the sake of simplicity the means and standard deviations are supposed to be integral. So we round $\mu/2$ to the nearest integer towards zero if necessary. The plant capacity dedicated to a product is set as the mean of its corresponding product. Only 30 product-capacity scenarios are generated because we have to compare our VHC method with the constraint sampling method which is pretty time-consuming.

For each product-capacity scenario, generate 10000 demand realizations. These will be used to evaluate the hub-and-chain structure forged by our method and the structures sampled by constraint sampling. Furthermore, the performances of the dedicated, long chain as well as full flexibility structure are also computed using these demand realizations.

The constraint sampling algorithm applied here is similar to the one in [Chou et al. \(2010\)](#) which considers a transshipment case. It has been modified to accommodate our example. As a result, we let

$$\tilde{E}(x_{ij}^*) = \sum_{i=1}^{10000} \left[\frac{D_i c_j}{\max\{\sum_i D_i, \sum_j c_j\}} \right] / 10000,$$

$$\tilde{p}_{ij} = \frac{\tilde{E}(x_{ij}^*)}{\sum_{i=1}^{20} \sum_{j=1}^{20} \tilde{E}(x_{ij}^*)}.$$

In addition to that, the constraint sampling process starts from a given dedicated structure so as to guarantee the performance of a sampled structure is at least as good as the dedicated system.

In the first method, we set $\theta_1 = 0.01$, $\theta_2 = 0.1$, $\theta_3 = 0.6$. In the second method, we set the step size as 0.01 which is small enough to get the number of groups that we want. Our budget B is 40 links. At most we can exclude 12 links to meet our link budget.

The long chains in the subgroups are formed following the natural sequence of the products, i.e. product 1 connected to plant 2 and so on. Though the sequence of a long chain surely impacts its performance, numerical tests demonstrate that the performance will not vary substantially.

Table 2.2: An Example of Demand Scenario

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
μ	115	33	384	336	358	322	210	196	409	159	408	395	427	253	318	476	222	31	434	316
σ	21	16	44	110	109	63	15	3	86	15	149	73	180	93	91	43	107	4	201	36

Before we proceed to the numerical results, let us first take an example and illustrate how the heuristic is executed. Table 2.2 shows a demand scenario sampled consisting of 20 products. Demand of each product follows a specific normal distribution with mean μ and standard deviation σ . Figure 2.7.1 (a) shows product-plant pairs with pre-established dedicated links and Figure 2.7.1 (b) shows how

these product-plant pairs are grouped. Let a circle represent a product and Let a square represent a plant. We can see from Figure 2.7.1 (b) that a dedicated group including pairs 8 and 18 has been selected out of the twenty pairs. Table 2.2 shows the standard deviations of product 8 and 18 are both negligible. Thus it is not worthwhile to invest any flexibility for such products. One observation of interest is products in group 1 have larger demand mean than those in group 2, 3 and 4. This is because the heuristic always excludes the product with minimum demand mean first.

Figure 2.7.2 demonstrates the realization of step 3 of the heuristic and what the hub-and-chain structure looks like. There is a long chain in each of the subgroups except the dedicated one. Pair 19 has been elected as the hub of the whole structure to bridge all the other subgroups such that extra demand and idle capacity could be matched among the groups.

The hub-and-chain structure significantly facilitates the exchange of excess demand and idle capacity. Our numerical results for Method 1 are summarized in Table 2.3, where columns 2 to 6 contain the respective performances of the dedicated, long chain, full flexibility, constraint sampling method, and hub-and-chain designs. Columns 7, 8 and 9 present the number of non-dedicated groups, the number of products in the dedicated group, and the number of flexibility links used in the hub-and-chain design, respectively. Finally, the last three columns show the efficiency of the hub-and-chain design, as well as the improvement of this design over the long chain and the best design from the constraint sampling method. Here, we define improvement of the hub-and-chain design over a flexibil-

Figure 2.7.1: How the pairs are grouped.

ity design $\mathcal{G}$ as

$$\frac{[\text{VHC}] - [\mathcal{D}_n]}{[\mathcal{G}] - [\mathcal{D}_n]} - 1 = \frac{[\text{VHC}] - [\mathcal{G}]}{[\mathcal{G}] - [\mathcal{D}_n]}.$$

In all scenarios, VHC achieves at least 92.5% of the benefits of full flexibility. Moreover, in more than 2/3 of the scenarios, VHC achieves at least 95% of the benefits of full flexibility, and in 1/6 of the scenarios, the hub-and-chain is almost equivalent to full flexibility (i.e. an efficiency of at least 99%).

Figure 2.7.2: The hub-and-chain structure formed.

Furthermore, we see from the last two columns that on average, the hub-and-chain design improves on the long chain by 45%, and improves on the best design from the constraint sampling method by 25%. These are substantial improvements, which corroborate our earlier result that product demand variance plays a significant role in the performances of flexibility designs. Hence, it is not surprising that a method like the VHC, which takes into account demand variances, outperforms existing methods. In terms of computational time, the VHC method requires only at most 1% of the time required to run and compare 100 designs using the constraint sampling method.

Since we have no flexibility budget, one concern is the VHC method might

2.7 Simulation and Analysis

use too many flexibility links compared to the 40 links used in the long chain. A quick look at Table 2.3 reveals that the VHC method only uses, on average, about 2 more links. In fact, in some instances, the VHC even utilizes less links than the long chain, and still obtains substantial gains.

Table 2.3: The results of different methods (1)

Scenario	$[\mathcal{D}_n]$	$[\mathcal{C}_n]$	$[\mathcal{F}_n]$	[CS]	[VHC]	s	$ G_0 $	VHC	$f(\mu, \sigma, \text{VHC})$	Improvement over(%)	
										$\mathcal{C}_n$	CS
1	5279	5541	5664	5586	5651	4	2	43	96.6	42.0	21.2
2	4396	4606	4737	4649	4714	4	6	39	93.3	51.4	25.7
3	4659	4957	5053	4988	5036	5	4	43	95.7	26.5	14.6
4	4857	5007	5108	5060	5105	3	3	41	98.8	65.3	22.2
5	5839	6193	6309	6218	6301	4	1	45	98.3	30.5	21.9
6	4447	4699	4804	4728	4804	3	3	41	100	41.7	27.0
7	3749	3933	4046	3986	4040	5	3	42	98.0	58.2	22.8
8	5212	5496	5614	5547	5614	4	2	44	100	41.5	20.0
9	5936	6334	6446	6331	6428	3	1	42	96.5	23.6	24.6
10	4235	4465	4540	4436	4526	3	6	38	95.4	26.5	44.8
11	5956	6171	6295	6198	6272	4	4	41	93.2	47.0	30.6
12	4640	4891	4949	4878	4948	4	1	44	99.7	22.7	29.4
13	4833	5070	5202	5113	5193	5	3	43	97.6	51.9	28.6
14	4341	4513	4650	4596	4645	4	2	44	98.4	76.7	19.2
15	4770	5077	5129	5002	5121	3	2	41	97.8	14.3	51.3
16	4528	4753	4889	4817	4883	4	5	40	98.3	57.8	22.8
17	5080	5292	5394	5283	5382	2	4	38	96.2	42.5	48.8
18	4489	4704	4848	4741	4830	3	3	41	95.0	58.6	35.3
19	4224	4433	4561	4495	4540	4	4	41	93.8	51.2	16.6
20	5198	5492	5570	5473	5547	3	1	43	93.8	18.7	26.9
21	4888	5149	5237	5155	5226	4	2	43	96.8	29.5	26.6
22	4225	4430	4557	4497	4547	5	5	42	97.0	57.1	18.4
23	4834	5007	5081	5026	5073	3	2	42	96.8	38.2	24.5
24	4009	4136	4233	4198	4215	5	5	43	92.0	62.2	9.0
25	3817	4003	4070	3999	4062	4	3	43	96.8	31.7	34.6
26	3713	3907	4034	3954	4010	4	4	42	92.5	53.1	23.2
27	5301	5560	5646	5609	5643	3	2	42	99.1	32.0	11.0
28	4104	4293	4445	4370	4416	5	6	39	91.5	65.1	17.3
29	4122	4286	4398	4357	4397	5	4	43	99.6	67.7	17.0
30	3340	3440	3540	3510	3534	4	6	40	97.0	94.0	14.1

To verify the stability of our results, we randomly generate another 10,000 demand realizations. The results are reported in Table 2.4, which shows minor discrepancies from Table 2.3.

Figure 2.7.3: Structure efficiency of the three methods

Table 2.5 reports the result of Method 2. We observe that the maximum expected output of VHC is always obtained when 2 or 4 links are in the dedicated group. Also the performances of VHC turn out to be unimodal as the number of links in the dedicated groups increases. This demonstrates that as more and more links are excluded, the main structure loses the sources of excess demand and idle capacity though more links are left for the hub-and-chain structure. Therefore, the process of searching proper number of links to be excluded is to find out the balance of the trade-off between the loss of flexibility in dedicated group and a larger number of small chains in the hub-and-chain structure. Computationally, this means that once $[VHC]$ decreases for some $|G_0|$, we can terminate our iterative

2.7 Simulation and Analysis

Table 2.4: The Results of different methods (2)

Scenario	$[\mathcal{D}_n]$	$[\mathcal{C}_n]$	$[\mathcal{F}_n]$	[CS]	[VHC]	s	$ G_0 $	VHC	$f(\mu, \sigma, \text{VHC})$	Improvement over(%)	
										$\mathcal{C}_n$	CS
1	5277	5539	5660	5571	5648	4	2	43	96.9	41.5	26.2
2	4404	4612	4745	4646	4721	4	6	39	93.1	52.4	31.2
3	4659	4961	5051	4976	5039	5	4	43	96.9	25.7	19.6
4	4864	5008	5115	5054	5106	3	3	41	96.5	68.3	27.4
5	5836	6191	6306	6244	6301	4	1	45	99.0	31.1	14.0
6	4466	4697	4821	4712	4804	3	3	41	95.2	46.2	37.2
7	3747	3935	4040	3987	4040	5	3	42	99.7	56.1	22.1
8	5223	5497	5618	5544	5615	4	2	44	99.1	42.9	22.0
9	5935	6335	6454	6322	6429	3	1	42	95.3	23.4	27.6
10	4235	4464	4543	4439	4524	3	6	38	93.9	26.3	41.6
11	5960	6174	6296	6180	6275	4	4	41	93.8	47.5	43.4
12	4645	4890	4960	4878	4948	4	1	44	96.1	23.6	29.9
13	4823	5072	5196	5107	5196	5	3	43	100	49.8	31.5
14	4349	4515	4654	4611	4644	4	2	44	96.6	77.7	12.4
15	4772	5075	5131	5029	5120	3	2	41	97.2	15.0	35.5
16	4526	4751	4895	4831	4881	4	5	40	96.0	57.8	16.3
17	5077	5293	5385	5283	5384	2	4	38	99.7	42.2	48.9
18	4493	4707	4850	4745	4832	3	3	41	94.9	58.1	34.4
19	4218	4433	4549	4470	4540	4	4	41	97.1	49.2	27.6
20	5191	5492	5562	5471	5546	3	1	43	95.7	18.1	26.7
21	4887	5153	5241	5143	5230	4	2	43	97.0	28.6	33.7
22	4229	4428	4560	4488	4546	5	5	42	95.8	58.9	22.3
23	4842	5010	5080	5027	5075	3	2	42	97.9	38.9	25.9
24	4001	4133	4228	4167	4212	5	5	43	93.0	59.5	27.1
25	3813	4004	4069	4019	4062	4	3	43	97.4	30.7	21.1
26	3691	3900	4012	3959	4003	4	4	42	97.3	49.7	16.8
27	5299	5558	5645	5569	5641	3	2	42	98.8	31.9	26.5
28	4100	4293	4432	4381	4415	5	6	39	94.8	63.3	12.0
29	4122	4285	4395	4356	4395	5	4	43	100	66.7	16.3
30	3345	3440	3548	3514	3536	4	6	40	94.4	100.1	12.8

method. Although our budget is 40 links, the performances are very close to full flexibility and improve significantly over the long chain. In fact, in some scenarios, we even have not exhausted our budget yet. Figure 2.7.3 compares the efficiencies of the designs generated by the long chain, the constraint sampling method and VHC with 40 links. VHC results in designs with much higher efficiency than the other two methods. It leaves little room for further improvement. Finally, we also observe that the efficiency of the hub-and-chain design is less variable (more stable) than those of the other two methods.

2.7 Simulation and Analysis

Table 2.5: The Results of the second method with 40 links

Scenario	$[\mathcal{C}_n]$	$[\mathcal{F}_n]$	2*	4	6	8	10	12
1	5539	5660	<u>5636</u>	5636	5623	5593	5559	5510
2	4612	4745	4700	<u>4719</u>	4715	4705	4659	4640
3	4961	5051	5023	<u>5035</u>	5018	4992	4952	4903
4	5008	5115	<u>5102</u>	5100	5092	5064	5050	5021
5	6191	6306	6242	<u>6272</u>	6247	6203	6157	6103
6	4697	4821	4783	<u>4794</u>	4775	4744	4715	4683
7	3935	4040	4022	<u>4035</u>	4025	3988	3965	3931
8	5497	5618	<u>5603</u>	5598	5582	5535	5507	5463
9	6335	6454	6383	<u>6398</u>	6373	6327	6275	6217
10	4464	4543	4518	<u>4529</u>	4527	4518	4483	4442
11	6174	6296	6271	<u>6274</u>	6260	6244	6211	6180
12	4890	4960	<u>4936</u>	4923	4909	4879	4852	4822
13	5072	5196	5180	<u>5185</u>	5172	5152	5105	5073
14	4515	4654	<u>4640</u>	4635	4623	4595	4574	4545
15	5075	5131	<u>5109</u>	5105	5084	5035	5012	4969
16	4751	4895	4840	<u>4886</u>	4845	4839	4806	4770
17	5293	5385	5356	<u>5359</u>	5358	5330	5299	5270
18	4707	4850	4820	<u>4823</u>	4811	4777	4744	4695
19	4433	4549	4499	<u>4539</u>	4528	4483	4465	4439
20	5492	5562	<u>5541</u>	5533	5517	5489	5453	5414
21	5153	5241	<u>5225</u>	5212	5198	5162	5145	5111
22	4428	4560	4498	<u>4548</u>	4542	4508	4496	4458
23	5010	5080	<u>5071</u>	5064	5059	5049	5020	5017
24	4133	4228	4200	<u>4213</u>	4212	4201	4171	4161
25	4004	4069	4054	<u>4058</u>	4049	4030	3986	3975
26	3900	4012	3999	<u>4008</u>	3998	3987	3944	3935
27	5558	5645	<u>5634</u>	5631	5612	5590	5565	5530
28	4293	4432	4364	<u>4424</u>	4414	4405	4371	4344
29	4285	4395	<u>4391</u>	4391	4384	4370	4335	4313
30	3440	3548	3486	<u>3542</u>	3534	3521	3508	3478

*-Number of links in the dedicated groups

2.8 Extension to Unbalanced Case

The unbalanced system is referred to as that the demand mean of a product might not be the same as the capacity of its dedicated plant. In contrast to the balanced case, we have to take into account the mismatch of demand mean and capacity while dealing with the fluctuation of demand caused by the variance. [Chou et al. \(2011\)](#) investigate the unbalanced case and use the expansion heuristic to construct flexibility design. The expansion heuristic works very well to balance the demand mean and capacity. This has been demonstrated in JG's example and the levi graph. However, the development of the expansion heuristic in [Chou et al. \(2011\)](#) does not consider the variance. Furthermore, the algorithm given in [Chou et al. \(2011\)](#) will not naturally lead construction of process flexibility to the long chains, as shown in following example.

The vector of demand means for the 8 products is (600, 600, 650, 650, 150, 100, 110, 100) and all demand follow the normal distributions with CV 0.4 while the capacity vector is (500, 700, 600, 700, 100, 150, 100, 110,). Suppose we are given a budget of 8 links and the base structure is a dedicated system as shown in Figure [2.8.1\(a\)](#). By applying the expansion heuristic, the new structure is shown in Figure [2.8.1\(b\)](#). We can see that the new structure consists of 4 short chains. There is no longer any unbalanced demand mean or capacity, nonetheless the total performance is 2649 and the structure efficiency is only 57.19%. Even if we further increase the number of links to 10, the additional two links (5,2) and (7,6) only help raise the performance to 2674 with structure efficiency of 65.96%. If we construct the structure in the manner of Figure [2.8.1\(c\)](#), the performance will be much better.

2.8 Extension to Unbalanced Case

Figure 2.8.1: A counter example

The performance of the two longer chains with 8 links is 2739 while the structure efficiency is 88.77%. Two additional links (5, 4) and (4, 5) will push them to 2768 and 98.95%, respectively.

The example implies that the expansion heuristic alone does not perform very well in some cases. In contrast, the hub and chain structure works quite well even in the unbalanced case. This motivate us to use hub and chain as the base structure and apply the expansion heuristic, if necessary, to further balance demand mean and capacity. Since the condition that demand mean is equal to dedicated capacity no longer holds, we have to modify the GCV index to

$$\frac{\max_{i \in G_j} \sigma_i}{\min(\min_{i \in G_j} \mu_i, \min_{i \in G_j} c_i)}.$$

Algorithm 2 Unbalanced Case

Step 1: Pick products with unqualified variance.

- 1: Initialize $x \leftarrow 0$; $n_1 \leftarrow n$;
- 2: **while** $(x + \min_{i \in G_1} \sigma_i) / \sum_{i=1}^n \sigma_i < \theta_1$ **do**
- 3: $x \leftarrow x + \min_{i \in G_1} \sigma_i$; $G_1 \leftarrow G_1 \setminus \{\arg \min_{i \in G_1} \sigma_i\}$ and $G_d \leftarrow G_d \cup \{\arg \min_{i \in G_1} \sigma_i\}$; $n_1 \leftarrow n_1 - 1$;
- 4: **end while**

Step 2: Assort the remaining products.

- 1: Initialize $j \leftarrow 1$; $n_{j+1} \leftarrow 0$;
- 2: **while** $G_j \neq \emptyset$ **do**
- 3: **while** $\frac{\max_{i \in G_j} \sigma_i}{\min(\min_{i \in G_j} \mu_i, \min_{i \in G_j} c_i)} > \theta_2$ **do**
- 4: $l^* \leftarrow \arg \min_{i \in G_j} \mu_i$; $G_j \leftarrow G_j \setminus \{l^*\}$; $G_{j+1} \leftarrow G_{j+1} \cup \{l^*\}$; $n_{j+1} \leftarrow n_{j+1} + 1$;
- 5: **end while**
- 6: $j \leftarrow j + 1$; initialize $n_{j+1} \leftarrow 0$;
- 7: **end while**
- 8: $J \leftarrow j - 1$.

Step 3: Forge the hub-and-chain structure.

- 1: **for** $j = 1$ to J **do**
- 2: **for** $i = 1$ to $n_j - 1$ **do**
- 3: $\mathcal{G}(G_j(i), G_j(i)) \leftarrow 1$; $\mathcal{G}(G_j(i), G_j(i+1)) \leftarrow 1$;
- 4: **end for**
- 5: $\mathcal{G}(G_j(n_j), G_j(1)) \leftarrow 1$; $l_j^* \leftarrow \arg \max_{i \in G_j} \sigma_i$; $\mathcal{G}(G_1(l_1^*), G_j(l_j^*)) \leftarrow 1$;
- 6: $\mathcal{G}(G_j(l_j^*), G_1(l_1^*)) \leftarrow 1$;
- 7: **end for**

Step 4: Add links to the hub-and-chain structure through the expansion heuristic (Chou et al., 2011).

Remark 2.8.1. *Non-Symmetrical System:*

The algorithm is largely the same as the unbalanced case despite the GCV has been changed and the expansion heuristic is added as the last step. However, when number of plants and products is different ($n > m$), the algorithm should be modified to accommodate this situation. We denote the product set in G_j by G_{j1} and plant set by G_{j2} . Let I_k represent index set of products pertaining to plant k . We adopt L^2 norm $\|\sigma_{I_k}\|$ to measure the level of variability for the branch of products-plant, which in step 1 decides the exclusion of pairs into the dedicated group. And $\frac{\max_{i \in G_{j1}} \sigma_i}{\min_{k \in G_{j2}} \min(\min_{i \in I_k} \mu_i, c_k)}$ is the GCV.

2.8.1 Numerical Results

For our numerical implementation, we benchmark our VHC method against the long chain, the constraint sampling method, and the expansion heuristic. We consider two settings; namely, homogeneous variation and heterogeneous variation. In both settings, we consider system size of $n = 20$. For homogeneous variation, we simulate 30 scenarios of 20 normally distributed product demands with $\mu_i \in \{100, 101, \dots, 500\}$ and fixed $CV_i = 0.4$, for $i = 1, \dots, 20$. Capacities are randomly generated as $c_i \in \{[0.8\mu], [0.8\mu] + 1, \dots, [1.2\mu]\}$ for $i = 1, \dots, 20$. To run VHC, we calibrate the parameters at $\theta_1 = 0.05$, $\theta_2 = 0.05$, and $\theta_3 = 0.8$. We then allow $l \in \{0, 1, 2, 3, 4, 5, 6\}$ links to be added via the expansion heuristic to the flexibility design generated by Steps 1 - 3 of Algorithm 2. Hence, we have 210 scenarios for which we can compare the performance of VHC with constraint sampling and the expansion heuristic. For fair comparison, we ensure the same

2.8 Extension to Unbalanced Case

number of links used by all three methods. We then generate 1,000 demand realizations for each scenario and compare the performance of VHC versus the other two existing methods. We report that in all 210 scenarios, VHC performs better than constraint sampling, while in all but two scenarios, VHC outperforms the Expansion Heuristic. Moreover, in these two scenarios, VHC is less than 0.2% worse than the expansion heuristic.

For heterogeneous variation, we likewise simulate 30 scenarios of 20 normally distributed product demands, but this time with $\mu_i \in \{100, 101, \dots, 500\}$ and $CV_i \in [0, 0.5]$, for $i = 1, \dots, 20$. Similar to homogeneous variation, capacities are randomly generated as $c_i \in \{[0.8\mu], [0.8\mu] + 1, \dots, [1.2\mu]\}$ for $i = 1, \dots, 20$. The parameters used are $\theta_1 = 0.1$, $\theta_2 = 0.1$, and $\theta_3 = 0.8$. We then allow $l \in \{0, 1, 2, 3, 4, 5, 10\}$ links to be added via the expansion heuristic to the flexibility design generated by Steps 1 - 3 of Algorithm 2. Hence, we have 180 scenarios for which we can compare the performance of VHC with constraint sampling and the expansion heuristic. We report that in all 180 scenarios, VHC outperforms the expansion heuristic, while in all but two scenarios, VHC is better than constraint sampling. Even in these two scenarios, VHC is less than 0.5% worse than constraint sampling.

In order to compare with the long chain, our hub-and-chain design must utilize more or less 40 links. For the homogeneous variation setting, we use an average of 40.5 links for the 30 scenarios considered and find that VHC is always better than the long chain. Moreover, the smallest gaps between VHC and the long chain occur whenever the GCV of the entire system is low. This supports our earlier result that variance plays a substantial role in process flexibility design and that the

VHC method can be an advantageous method. For the heterogeneous variation setting, we utilize an average of 39.8 links and find that VHC is better than the long chain in 21 out of the 30 scenarios considered. In the 9 scenarios where the long chain performs better, the gaps are small and these instances also coincide with systems with low GCVs.

Finally, we observe that an average of 46.5 links for homogeneous variation and 45.8 links for heterogeneous variation are sufficient to achieve about 95% of the benefits of full flexibility. This level of performance cannot be obtained using the long chain which, on average, only provides 70-80% of the benefits of full flexibility. Furthermore, the use of 47 links is still only a fraction of the 400 links of full flexibility and less than the 60 links of the 3-chain (i.e. a flexibility design that appends n more flexible links to the long chain).

2.9 Distributional Robustness

In the previous sections, our analytical results are derived from the normal and the symmetric three-point distributions for reasons of tractability. And we then conduct the simulation on systems with the truncated normal distributions. Since all of the demand distributions are assumed to be symmetric and the VHC method indeed only needs the first and second moment information of the demand distributions, a question arises naturally whether the VHC method is distributionally robust. In other words, how do the flexibility designs generated by the VHC method perform if the demand is not necessarily symmetric but with the same mean and variance information as before?

Our goal herein is to numerically investigate the distributional robustness of VHC. To do that, we use the lognormal and the gamma distributions with the same 30 scenarios of mean and variance information as in Section 2.7. The lognormal and the gamma distributions appear in various papers in the literature (Berry et al., 2012; Burgin, 1975; Gallego, 1995). The reason for choosing the lognormal and the gamma distributions is that we can control the mean and variance of them freely. Note that many other distributions, like the exponential distribution, are not able to meet this requirement. The lognormal pdf is expressed as

$$p(x; m, v) = \frac{1}{x\sqrt{2\pi v}} e^{-\frac{(\ln x - m)^2}{2v}} \quad \text{for } x \geq 0 \text{ and } v > 0,$$

where m and $\sqrt{v}$ respectively represent the mean and the standard deviation of the variable x 's natural logarithm, i.e. the variable following the normal distribution with mean m and variance v . m and $\sqrt{v}$ are also referred to as the location parameter and the scale parameter, respectively. Given the mean μ and the standard deviation σ , m and v can be obtained as follows.

$$m = \ln \left(\frac{\mu^2}{\sqrt{\sigma^2 + \mu^2}} \right).$$

$$v = \ln \left(1 + \frac{\sigma^2}{m^2} \right).$$

The pdf of the gamma distribution is as follows.

$$g(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad \text{for } x \geq 0 \text{ and } \alpha, \beta > 0,$$

Table 2.6: The results of the lognormal distributions

Scenario	[VHC]	$[\mathcal{C}_n]$	$[\mathcal{F}_n]$	$[\mathcal{D}_n]$	$f(\mu, \sigma, \text{VHC})$	$f(\mu, \sigma, \mathcal{C}_n)$
1	5613	5508	5634	5238	94.6	68.2
2	4696	4584	4720	4371	93.2	61.2
3	5010	4935	5040	4635	92.7	74.3
4	5088	4988	5098	4829	96.3	58.9
5	6246	6158	6290	5800	91.1	73.0
6	4771	4670	4793	4426	94.1	66.5
7	4019	3911	4036	3727	94.3	59.5
8	5581	5470	5603	5187	94.8	68.0
9	6367	6299	6426	5890	89.0	76.3
10	4514	4440	4531	4208	94.7	71.9
11	6253	6141	6267	5916	96.0	64.1
12	4917	4869	4943	4614	92.0	77.4
13	5160	5044	5186	4801	93.1	63.2
14	4620	4487	4633	4312	95.8	54.4
15	5082	5049	5118	4740	90.4	81.8
16	4866	4723	4883	4502	95.5	58.1
17	5336	5263	5370	5047	89.4	66.8
18	4800	4682	4835	4464	90.6	58.7
19	4520	4406	4539	4188	94.7	62.1
20	5519	5463	5542	5157	94.1	79.5
21	5204	5123	5219	4854	95.8	73.7
22	4526	4406	4546	4201	94.3	59.6
23	5053	4985	5064	4803	95.9	70.0
24	4198	4114	4210	3974	95.0	59.4
25	4046	3983	4059	3792	95.2	71.8
26	3980	3880	4010	3677	91.0	61.0
27	5610	5536	5632	5279	93.6	72.7
28	4402	4261	4411	4058	97.3	57.4
29	4377	4266	4387	4103	96.6	57.3
30	3527	3422	3532	3317	97.6	48.7

where α is the shape parameter and β is the rate parameter. Here $\Gamma(\alpha)$ is the gamma function evaluated at α . Given the mean μ and the standard deviation σ of the gamma distribution, the shape parameter and the rate parameter are respectively expressed as

$$\alpha = \frac{\mu^2}{\sigma^2}$$

and

$$\beta = \frac{\mu}{\sigma^2}.$$

The systems to be studied are still balanced with 20 products and 20 plants. Again we let the budget of links be 40 in order to compare with the performances of the long chains. We use Method 2 for the balanced systems to generate the hub-and-chain structures. Obviously, these structures are only related to the mean

2.9 Distributional Robustness

and variance information and are independent of the demand distributions. So they are the same as those generated by Method 2 in Section 2.7 using 40 links. We randomly generate 10,000 demand instances for the lognormal and the gamma distributions, respectively. We omit the results of the VHC designs for the cases wherein the numbers of product-and-plant pairs in the dedicated groups differ from each other as shown in Table 2.5. Instead, we directly report the maximal performances of the VHC designs for the 30 mean and variance scenarios. The results are presented in Table 2.6 and 2.7.

Table 2.7: The results of the gamma distributions

Scenario	[VHC]	$[\mathcal{C}_n]$	$[\mathcal{F}_n]$	$[\mathcal{D}_n]$	$f(\mu, \sigma, \text{VHC})$	$f(\mu, \sigma, \mathcal{C}_n)$
1	5612	5501	5637	5222	94.1	67.3
2	4695	4575	4717	4352	94.0	61.0
3	5008	4923	5030	4614	94.6	74.2
4	5085	4981	5096	4821	95.9	58.2
5	6243	6153	6291	5787	90.5	72.5
6	4772	4665	4795	4414	94.2	66.0
7	4018	3905	4036	3715	94.1	58.9
8	5580	5460	5602	5172	94.7	67.0
9	6366	6288	6420	5871	90.1	75.8
10	4511	4435	4528	4195	95.0	72.1
11	6252	6139	6268	5909	95.4	63.9
12	4912	4866	4944	4605	90.6	77.0
13	5155	5034	5183	4783	93.0	62.8
14	4619	4480	4634	4302	95.6	53.7
15	5081	5046	5119	4729	90.2	81.2
16	4868	4711	4879	4484	97.2	57.6
17	5335	5259	5371	5037	89.1	66.2
18	4797	4675	4831	4448	91.3	59.4
19	4513	4399	4537	4175	93.3	61.7
20	5517	5460	5543	5147	93.4	79.1
21	5200	5120	5220	4845	94.6	73.4
22	4525	4402	4544	4189	94.6	60.0
23	5052	4984	5064	4800	95.5	69.7
24	4196	4112	4212	3969	93.1	58.7
25	4041	3976	4053	3781	95.5	71.6
26	3981	3871	4007	3661	92.4	60.8
27	5609	5526	5630	5266	94.3	71.6
28	4400	4261	4416	4054	95.5	57.3
29	4371	4262	4390	4097	93.7	56.4
30	3529	3416	3529	3309	99.8	48.4

For the 30 scenarios with demands under the lognormal distributions, we report that on average the efficiency of the VHC designs reaches 94.0% in contrast to 66.0% of the long chains. The worst efficiency of the VHC designs is 89.0% while that of the long chains is 48.7%. For the 30 scenarios with demands under the

gamma distributions, the average efficiency of the VHC designs and the long chains reach 93.9% and 65.5%, respectively. Meanwhile, the worst efficiency of the VHC designs is 89.1% in contrast to 48.4% of the long chains.

From the numerical results, we can see that the VHC designs are robust for the lognormal and the gamma distributions. More specifically, on average more than 93% of the benefits of full flexibility have been achieved by the VHC designs. However, the performances of the VHC design in each mean-and-variance scenario for the lognormal and the gamma distributions are less than that of the normal distributions due to the asymmetric demands. We observe this is consistent with the decreases in the performances of the long chains and full flexibility for the lognormal and the gamma distributions.

2.10 Conclusions

In this chapter we investigate the process flexibility design problem for balanced but nonidentical systems. Though the long chain has been proved very powerful flexibility design in the literature for identical systems, we show the long chain is not always superior to short chains for balanced but nonidentical systems. Also we notice that the constraint sampling method and the expansion heuristic in the literature do not incorporate the demand variance information into the process flexibility designs, at least not explicitly.

Then we show the effect of demand mean and variance on the performance of a flexibility structure. First of all, the potential of a system increases in the demand variance associated with the products in the system. It implies only systems

with high demand variation is worthwhile to invest in process flexibility. And also a product with lower demand variance contributes less to the whole system. Secondly, the design efficiency improves when the demand mean is increased. It indicates that a product with small demand mean could be a bottleneck of the efficiency of the design. Thirdly the long chain efficiency is bounded from above by that of another identical system, where the demand mean equals the maximal demand mean of the nonidentical long chain and the variance equals the minimal variance of the nonidentical one. Similarly it is bounded from below by an identical system with the minimal demand mean and the maximal variance of the nonidentical system. The group coefficient of variation, defined as the ratio of the maximal standard deviation and the minimal mean of a nonidentical long chain, can be applied to characterize the long chain efficiency.

Based on these insights, we then introduce our variance-based hub-and-chain heuristic to design flexibility structures. We show numerically that the VHC method can achieve at least 90% of the benefits of full flexibility even using the same number of links as the long chain. Moreover, our method is, on average, 45% better than the long chain, and 25% better than constraint sampling. Furthermore, our method uses only 1% of the computation time of the constraint sampling method.

We extend VHC to the case when product demands need not be matched with the capacities of their primary plants. We benchmark VHC against the constraint sampling method and the Expansion Heuristic. We find that our method produces better flexibility designs than both existing methods in 99% of all the scenarios considered. In fact, even in the scenarios where our method performs worse, we

are less than 0.5% worse off. The designs generated by VHC are also better than the long chain in most of the scenarios we tested.

Finally we investigate the distributional robustness of the VHC method. We show that the flexibility designs generated by the VHC method can cope with demands following the lognormal and the gamma distributions very well. For the lognormal and the gamma distributions, the flexibility design efficiency on average can achieve more than 93%.

3

Multistage Process Flexibility

3.1 Introduction

[Chou et al. \(2010\)](#) show that in a single stage system there exists a sparse structure which can achieve $1 - \epsilon$ of the benefits of full flexibility. This essay targets to show this is also true for a multistage system. Meanwhile in the appendix of [Jordan and Graves \(1991\)](#), they point out that the dual of their model (single stage) is a 0-1 integer program. We are also curious about whether such a result is true for the multistage case and, if true, what future research it can lead to?

The multistage process flexibility problem has seldom been studied in the literature, in part because of its complexity. The only paper dedicated to research on the multistage problem is [Graves and Tomlin \(2003\)](#). They find that the stage-spanning bottleneck and the floating bottleneck deteriorate the expected output of multistage flexibility structures. They also provide an index to measure the efficiency of a flexibility structure. The guidelines provided for designing the

multistage flexibility basically is applying J-G guidelines (Jordan and Graves, 1995) to each stage in a decentralized way. Zheng (2007) also briefly studies a two-stage process flexibility problem using the constraint sampling method.

We model the problem in a way similar to Graves and Tomlin (2003) except our objective is to maximize the output. The interstage logistics routes are also ignored. This assumption is consistent with the reality of some industries. For instance, suppliers of automakers often build their plants adjacent to auto assembly lines.

In the first part of this essay we consider the multistage system with known distribution. We extend the existence of a sparse structure for the single-stage system to the multistage case. Then we apply the constraint sampling method to the multistage system. In the second part of this essay we consider multistage systems with unknown demand distributions. By introducing a complete positive cross moment model developed by Natarajan et al. (2011) we solve the worst performance evaluation problem for such systems. We then discuss the flexibility design problem in the multistage setting.

3.2 The Model

Consider a K -stage supply chain with I products and J facilities in total. Stage $k, k = 1, \dots, K$, has J_k different plants. Each product needs processing in each stage. The flexibility of the manufacturing system at stage k is represented by a bipartite graph $G_k = (A, B_k, \mathcal{G}_k)$, where A and B_k denote the product and plant vertices in stage k and $\mathcal{G}_k \subseteq A \times B_k$ consists of the set of links from A

to B_k . The link $(i, j) \in \mathcal{G}_k$ represents the capability of processing product i in plant j in stage k . Let $\mathcal{G} = \{\mathcal{G}_1, \dots, \mathcal{G}_K\}$ denote the multistage flexibility design consisting of all stages. Then we use bipartite graph $G = (A, \bigcup B_k, \mathcal{G})$ to denote the multistage flexibility system. $|\mathcal{G}_k|$ is defined as the cardinality of $\mathcal{G}_k$, i.e. the number of links connecting the plants in stage k to the products. Likewise, let $|\mathcal{G}| = \sum_{k=1}^K |\mathcal{G}_k|$ denote the cardinality of $\mathcal{G}$, i.e. the number of links in total for the multistage system G . For simplicity, we use N to denote it. The j th plant in stage k has capacity c_j^k where $j = 1, \dots, J_k$. Let D_i , $i = 1, \dots, I$, denote the random demand for product i . Let $\mathbf{D} = \{D_1, \dots, D_I\}$ represent the random demand vector. Subsequently let d_i denote an instance of demand for product i . The *performance* of a flexibility design is the expected output over all random instances. And hereafter we use the terms *performance* and *expected output* interchangeably. Given an instance of demand $\mathbf{d} = \{d_1, \dots, d_I\}$ and a flexibility design $\mathcal{G}$, the maximum output of the multistage system G is given as follows.

$$Z(\mathbf{d}, \mathcal{G}) = \max \sum_{i=1}^I y_i, \quad (3.1)$$

$$\text{s.t.} \quad y_i - \sum_{(i,j) \in \mathcal{G}_k} x_{ij}^k \leq 0 \quad i = 1, \dots, I, k = 1, \dots, K \quad (3.2)$$

$$y_i \leq d_i \quad i = 1, \dots, I \quad (3.3)$$

$$\sum_{(i,j) \in \mathcal{G}_k} x_{ij}^k \leq c_j^k \quad j = 1, \dots, J_k, k = 1, \dots, K \quad (3.4)$$

$$x_{ij}^k, y_i \geq 0 \quad (3.5)$$

where x_{ij}^k is the number of product i processed by plant j at stage k and y_i is the eventual output of product i from system G .

Under random demand $\mathbf{D}$, the maximal output $Z(\mathbf{D}, \mathcal{G})$ is also a random variable. The expected output of system G with flexibility design $\mathcal{G}$ is defined as $E_{\mathbf{D}}[Z(\mathbf{D}, \mathcal{G})]$. We use $\mathbf{b}$ to denote the vector consisting of the right-hand side values of the constraints in problem $Z(\mathbf{D}, \mathcal{G})$. Let $\mathbf{b}_r$ represent an element of $\mathbf{b}$, $r = 1, \dots, (K+1)I+J$. Thus $\mathbf{b}$ is in the form of $(0, \dots, 0, D_1, \dots, D_I, c_1^1, \dots, c_{J_1}^1, \dots, c_j^k, \dots, c_{J_K}^K)^T$, where $j = 1, \dots, J_k$ and $k = 1, \dots, K$. Since for $Z(\mathbf{D}, \mathcal{G})$ only the demand elements of $\mathbf{b}$ are random, $\mathbf{b}$ is partially random. Let b denote a realization of $\mathbf{b}$ with element b_r , $r = 1, \dots, (K+1)I+J$. The problem of interest is to find a flexibility design $\mathcal{G}$ with cardinality N that maximizes the expected output $E_{\mathbf{D}}[Z(\mathbf{D}, \mathcal{G})]$.

Before we proceed to our main result, we introduce the following lemma. Let $\{x_{ij}^{k*}(\mathbf{D}), y_i^*(\mathbf{D})\}$ and $\{x_{ij}^{k*}(\mathbf{d}), y_i^*(\mathbf{d})\}$ denote the optimal solutions in $Z(\mathbf{D}, \mathcal{G})$ and $Z(\mathbf{d}, \mathcal{G})$, respectively.

Lemma 3.2.1. *For any optimal solution $\{x_{ij}^{k*}(\mathbf{D}), y_i^*(\mathbf{D})\}$ of $Z(\mathbf{D}, \mathcal{G})$,*

$$\min_{k=1}^K \left\{ \sum_{i=1}^I \sum_{j=1}^{J_k} E_{\mathbf{D}}[x_{ij}^{k*}(\mathbf{D})] \right\} \geq \sum_{i=1}^I E_{\mathbf{D}}[y_i^*(\mathbf{D})].$$

Proof For a demand instance $\mathbf{d}$, we have $\sum_{i=1}^I \sum_{j=1}^{J_k} x_{ij}^{k*}(\mathbf{d}) \geq \sum_{i=1}^I y_i^*(\mathbf{d})$, $\forall k = 1, \dots, K$. Suppose the demand $\mathbf{D}$ follows a probability distribution with

the cumulative distribution function $F(\mathbf{D})$ and positive support U^I . Then

$$\int_{U^I} \sum_{i=1}^I \sum_{j=1}^{J_k} x_{ij}^{k*}(\mathbf{D}) dF(\mathbf{D}) \geq \int_{U^I} \sum_{i=1}^I y_i^*(\mathbf{D}) dF(\mathbf{D}), \quad \forall k = 1, \dots, K,$$

which is equivalent to

$$\min_{k=1}^K \left\{ \sum_{i=1}^I \sum_{j=1}^{J_k} \mathbb{E}_{\mathbf{D}}[x_{ij}^{k*}(\mathbf{D})] \right\} \geq \sum_{i=1}^I \mathbb{E}_{\mathbf{D}}[y_i^*(\mathbf{D})].$$

We have finished the proof. $\square$

Lemma 3.2.1 indicates that the expected maximum output of a multistage system will not exceed the expected output of any stage. In other words, the eventual output of the multistage system is determined by the bottleneck stages. Graves and Tomlin (2003) show that the bottleneck stage changes with the demand instance (floating bottleneck) and the eventual output is affected by the multistage flexibility design (stage spanning bottleneck) as well.

Let $\mathcal{F}$ represent the full flexibility design for a multistage system. And the maximal output $Z(\mathbf{D}, \mathcal{F})$ for this system is

$$Z(\mathbf{D}, \mathcal{F}) = \min_{k=1}^K \min \left\{ \sum_{i=1}^I D_i, \sum_{j=1}^{J_k} c_j^k \right\},$$

and an optimal solution for problem $Z(\mathbf{D}, \mathcal{F})$ is

$$\begin{aligned} x_{ij}^{k*}(\mathbf{D}) &= \frac{D_i c_j^k}{\max\{\sum_{i=1}^I D_i, \sum_{j=1}^{J_k} c_j^k\}} \\ &= \frac{D_i}{\sum_{i=1}^I D_i} \frac{c_j^k}{\sum_{j=1}^{J_k} c_j^k} \min\{\sum_{i=1}^I D_i, \sum_{j=1}^{J_k} c_j^k\}, \end{aligned}$$

and

$$\begin{aligned} \min_{k=1}^K \sum_{j=1}^{J_k} x_{ij}^{k*}(\mathbf{D}) &= \frac{D_i}{\sum_{i=1}^I D_i} \min_{k=1}^K \min\{\sum_{i=1}^I D_i, \sum_{j=1}^{J_k} c_j^k\} \\ &= \frac{D_i}{\sum_{i=1}^I D_i} Z(\mathbf{D}, \mathcal{F}), \end{aligned}$$

Clearly,

$$\sum_{i=1}^I \min_{k=1}^K \sum_{j=1}^{J_k} x_{ij}^{k*}(\mathbf{D}) = Z(\mathbf{D}, \mathcal{F}).$$

If we denote the coefficient matrix for x_{ij}^k as A and the coefficient matrix for variable y_i as B in problem $Z(\mathbf{D}, \mathcal{F})$, the multistage process flexibility problem $Z(\mathbf{D}, \mathcal{F})$ can be expressed in the following form:

$$(P1) : Z(\mathbf{b}) = \max\left\{\sum_{i=1}^{n_2} y_i : Ax + By \leq \mathbf{b}; x_j \geq 0, y_i \geq 0 \ j = 1, \dots, n_1\right\} \quad (3.6)$$

where A and B are $m \times n_1$ and $m \times n_2$ matrices, respectively. The reason for using $\mathbf{b}$ instead of $\mathbf{D}$ is for convenience and also because the demand variables are included in $\mathbf{b}$. And (P1) will be a more appropriate representation of the problem if in $Z(\mathbf{D}, \mathcal{G})$ the capacity is not necessarily deterministic. However, for the case wherein only demands are random, the random variables in $\mathbf{b}$ are exactly

the same as in $\mathbf{D}$ and therefore the operators $E_{\mathbf{D}}$ and $E_{\mathbf{b}}$ have the same meaning. Note for the specific problem $Z(\mathbf{D}, \mathcal{F})$, $m = (K + 1)I + J$, $n_1 = IJ$ and $n_2 = I$. Notice that matrix B has the following form:

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots \\ 0 & \vdots & 0 & \cdots & \cdots \\ 0 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & \ddots & 0 & \cdots \\ & \mathbf{e} & & & \\ & \mathbf{0} & & & \end{bmatrix}.$$

Matrix A has the form of :

$$\begin{bmatrix} -1 & \cdots & -1 & & & \ddots & & \\ \ddots & & & -1 & \cdots & -1 & & 0 \\ & 0 & & & & & -1 & \cdots & -1 \\ & & \ddots & & & & & & \ddots \\ & & & \mathbf{0} & & & & & \\ & 1 & & 1 & & 1 & & & \\ & & 1 & & 1 & & 1 & & \\ 0 & & 1 & 0 & & 1 & 0 & & 1 & 0 \\ & & & \ddots & & \ddots & & \ddots & \end{bmatrix}.$$

The first KI rows of matrices A and B consist of the coefficients of the constraints $y_i - \sum_{(i,j) \in \mathcal{G}_k} x_{ij}^k \leq 0$. The following I rows consist of the coefficients of constraints $y_i \leq d_i$. And the last J rows consist of the coefficients of the constraints $\sum_{(i,j) \in \mathcal{G}_k} x_{ij}^k \leq c_j^k$.

The dual of (P1) is:

$$(D1) : Z(\mathbf{b}) = \min \left\{ \sum_{j=1}^m \mathbf{b}_j w_j : A^T w \geq 0; B^T w \geq \mathbf{1}; w_j \geq 0 \right\} \quad (3.7)$$

Following the form of A and B , clearly $\mathbf{b}_j = 0$ for $j = 1, \dots, KI$. In addition,

in each row of B^T , there are $K + 1$ variables of which coefficients are equal to 1 while other coefficients are equal to 0. That is to say, in the i th row of A^T , coefficients for w_j with $j = (i - 1)I + 1, \dots, iI$ and $j = KI + i$ are equal to 1. And in each row of A^T , there are only two elements not equal to 0. The first one is equal to -1 and the second one is equal to 1. For example in the i th row, $i \leq J_1$, of A^T , the i th element is equal to -1 and the $(K + 1)I + i$ th element is equal to 1.

Let $\{x_i^*(\mathbf{b}), y_j^*(\mathbf{b})\}$, $i = 1, \dots, n_1$ and $j = 1, \dots, n_2$, denote the optimal solution to (P1). If we sample N constraints from $A^T w \geq 0$ by attributing a sampling probability of

$$p_i = \frac{\mathbf{E}_{\mathbf{b}}[x_i^*(\mathbf{b})]}{\sum_{i=1}^{n_1} \mathbf{E}_{\mathbf{b}}[x_i^*(\mathbf{b})]}, i = 1, \dots, n_1$$

to the i th constraint of $A^T w \geq 0$ and denote the constraints sampled by S , which represents the links from demands to plants, then we have a problem:

$$\text{D}(S) : Z(\mathbf{b}, S) = \min\left\{\sum_{j=1}^m \mathbf{b}_j w_j : A_S^T w \geq 0; B^T w \geq \mathbf{1}; w_j \geq 0\right\} \quad (3.8)$$

The dual of $\text{D}(S)$ is

$$\text{P}(S) : Z(\mathbf{b}, S) = \max\left\{\sum_{i=1}^{n_2} y_i : A_S x_S + B y \leq \mathbf{b}; x_S \geq 0; y_i \geq 0\right\} \quad (3.9)$$

Jordan and Graves (1991) show that the dual of a single-stage problem is an integer program. They use contradiction to prove the result. Here we prove it for the multistage setting.

Lemma 3.2.2. *In $\text{D}(S)$ there exists an optimal solution such that $w_j^*(b, S) = 0$ or 1, $\forall j = 1, \dots, m$ and a realization b .*

Proof We use the shadow price to prove the result. According to the complementary theorem, the corresponding complementary dual variable in $D(S)$ of a nonbinding constraint in $P(S)$ equals 0. We first prove it in two cases for the constraints $y_i \leq d_i, \forall i = 1, \dots, I$ in $P(S)$. The shadow price of increasing d_i by 1 is either 1 or 0. If there is excess capacity, then the increase will be 1 and otherwise will be 0. Likewise, the shadow price of decreasing d_i by 1 is either -1 or 0.

If the shadow price of increasing d_i by 1 is 1, then the dual variable $w_{KI+i} \geq 1$ because the feasible region of $D(S)$ is not affected and $w_{KI+i} < 1$ contradicts the strong dual theorem, i.e., the dual objective value must be greater than or equal to the primal objective value.

If the shadow price of increasing d_i by 1 is 0, then $y_i < d_i + 1$ which is no longer binding. Then let d_i decrease by 1. There can be two subcases as follows. Shadow price is 0 or -1. If it is 0, that means the capacity vacated by product i is taken by other products. In this case $w_{KI+i} = 0$ otherwise it will contradict the dual theorem. If the shadow price is -1, this means that all demands that can utilize the plant capacity for d_i have been fulfilled. Hence $w_{KI+i} = 1$.

The similar argument can be applied to the dual variables corresponding to constraints $\sum_{(i,j) \in \mathcal{G}_k} x_{ij}^k \leq c_j^k$ and $y_i - \sum_{(i,j) \in \mathcal{G}_k} x_{ij}^k \leq 0$. $\square$

Remark 3.2.1. *Since the coefficients of the dual problem correspond to demands or capacities, we can solve the dual problem by computer first and find the bottleneck facilities through the optimal solution of the dual. For example, if a decision variable with demand coefficient is equal to 1, then it infers that this demand is fully fulfilled. If a decision variable with demand coefficient is equal to 0, then we can infer that there is a bottleneck for this demand somewhere. If a decision*

variable with capacity coefficient is equal to 1, we can infer the corresponding facility is a bottleneck and increasing its capacity will improve the performance of the whole system. If a decision variable with capacity coefficient is equal to 0, we can infer the corresponding facility is not a bottleneck.

Property 1. $x_i^*(\mathbf{b}) \leq \lambda E_{\mathbf{b}}[x_i^*(\mathbf{b})]$ almost surely for some constant $\lambda \geq 0$ (independent of n), and for all $i = 1, \dots, n_1$.

Theorem 3.2.1. (Chou et al., 2010) For a system with $K = 1$, suppose that Property 1 holds for (P1). Then, there exists a set S with cardinality $N = O(\lambda m / \epsilon)$, such that

$$E_{\mathbf{b}}[Z(\mathbf{b}, S)] \geq (1 - \epsilon) E_{\mathbf{b}}[Z(\mathbf{b})].$$

Next we want to verify that there also exists a sparse support set S with cardinality N such that $E_{\mathbf{b}}[Z(\mathbf{b}, S)] \geq (1 - \epsilon) E_{\mathbf{b}}[Z(\mathbf{b})]$ for a multistage system. That is to say, we only consider the case $K > 1$.

Theorem 3.2.2. Suppose Property 1 holds for (P1). Then there exists a set S with cardinality $N = O(K \lambda m / \epsilon)$, such that $E_{\mathbf{b}}[Z(\mathbf{b}, S)] \geq (1 - \epsilon) E_{\mathbf{b}}[Z(\mathbf{b})]$.

Proof. Given a fixed b , first we let $y^*(b)$ and $w^*(b)$ denote the optimal solution in (P1) and (D1), respectively. Likewise, we let $y^*(b, S)$ and $w^*(b, S)$ denote the

optimal solution in $P(S)$ and $D(S)$, respectively.

$$\begin{aligned}
Z(b, S) &= \sum_{j=1}^m b_j w_j^*(b, S) \\
&= \sum_{j=1}^m b_j w_j^*(b, S) + Z(b) - \sum_{i=1}^{n_2} y_i^*(b) \\
&\geq Z(b) + \sum_{j=1}^m \left(\sum_{i=1}^{n_2} B_{ji} y_i^*(b) + \sum_{i=1}^{n_1} A_{ji} x_i^*(b) \right) w_j^*(b, S) - \sum_{i=1}^{n_2} y_i^*(b) \\
&= Z(b) + \sum_{i=1}^{n_1} x_i^*(b) \left(\sum_{j=1}^m A_{ji} w_j^*(b, S) \right) + \sum_{i=1}^{n_2} y_i^*(b) \left(\sum_{j=1}^m B_{ji} w_j^*(b, S) - 1 \right) \\
&\geq Z(b) + \sum_{i=1}^{n_1} x_i^*(b) \left(\sum_{j=1}^m A_{ji} w_j^*(b, S) \right).
\end{aligned}$$

The last step is because we do not sample constraints from $B^T w \geq \mathbf{e}$ and thus for any given b we always have $\sum_{j=1}^m B_{ji} w_j^*(b, S) \geq 1$

Due to the particular structure of matrix A the term $\sum_{j=1}^m A_{ji} w_j^*(b, S)$ can be represented in the form of $w_{p(i)}^*(b, S) - w_{q(i)}^*(b, S)$, where the subscripts $p(i), q(i)$ are dependent on i . Since $w_j^*(b, S) \geq 0$,

$$\sum_{j=1}^m A_{ji} w_j^*(b, S) \geq 0$$

does not hold for all $w_j^*(b, S)$.

We also define

$$E_S[\alpha(S)] = \int_S \alpha(S) dF(S)$$

where $\alpha(S)$ is a function on space S , and $F(S)$ is a cumulative distribution

function on S . Then

$$\begin{aligned}
 \mathbb{E}_S [Z(b, S)] &\geq Z(b) + \mathbb{E}_S \left[\sum_{i=1}^{n_1} x_i^*(b) \left(\sum_{j=1}^m A_{ji} w_j^*(b, S) + 1 - 1 \right) \right] \\
 &\geq Z(b) - \sum_{i=1}^{n_1} x_i^*(b) P \left(\sum_{j=1}^m A_{ji} w_j^*(b, S) \leq 0 \right) \\
 &\geq Z(b) - \sum_{i=1}^{n_1} \lambda \mathbb{E}_{\mathbf{b}} [x_i^*(\mathbf{b})] P \left(\sum_{j=1}^m A_{ji} w_j^*(b, S) \leq 0 \right) \\
 &\quad \text{(by Property 1)} \\
 &\geq Z(b) - \lambda \left(\sum_{i=1}^{n_1} \mathbb{E}_{\mathbf{b}} [x_i^*(\mathbf{b})] \right) \left[\sum_{i=1}^{n_1} \frac{\beta \mathbb{E}_{\mathbf{b}} [x_i^*(\mathbf{b})]}{\sum_{l=1}^{n_1} \beta \mathbb{E}_{\mathbf{b}} [x_l^*(\mathbf{b})]} P \left(\sum_{j=1}^m A_{ji} w_j^*(b, S) \leq 0 \right) \right] \\
 &\geq Z(b) - \lambda \left(\sum_{i=1}^{n_1} \mathbb{E}_{\mathbf{b}} [x_i^*(\mathbf{b})] \right) \left[\sum_{i=1}^{n_1} \frac{\mathbb{E}_{\mathbf{b}} [x_i^*(\mathbf{b})]}{\sum_{l=1}^{n_1} \mathbb{E}_{\mathbf{b}} [x_l^*(\mathbf{b})]} P \left(\sum_{j=1}^m A_{ji} w_j^*(b, S) \leq 0 \right) \right] \\
 &\geq Z(b) - \frac{\lambda m}{N+1} \left(\sum_{i=1}^{n_1} \mathbb{E}_{\mathbf{b}} [x_i^*(\mathbf{b})] \right) \\
 &\quad \text{From (Calafiore and Campi, 2005)} \\
 &\geq Z(b) - \frac{K \lambda m}{N+1} \sum_{i=1}^{n_2} \mathbb{E}_{\mathbf{b}} [y_i^*(\mathbf{b})] \\
 &\quad \text{(by Lemma 3.2.1)} \\
 &= Z(b) - \frac{K \lambda m}{N+1} \sum_{i=1}^{n_2} \mathbb{E}_{\mathbf{b}} [y_i^*(\mathbf{b})]
 \end{aligned}$$

The second step is derived from Lemma 3.2.2. Since $w_j^*(b, S) = 0$ or $1 \quad \forall j$, $\sum_{j=1}^m A_{ji} w_j^*(b, S) + 1 \geq 0$.

Clearly $\sum_{i=1}^{n_2} \mathbb{E}_{\mathbf{b}} [y_i^*(\mathbf{b})] = \mathbb{E}_{\mathbf{b}} [Z(\mathbf{b})]$. And take expectation over $\mathbf{b}$, we obtain

$$\mathbb{E}_{\mathbf{b}, S} [Z(\mathbf{b}, S)] \geq (1 - \epsilon) \mathbb{E}_{\mathbf{b}} [Z(\mathbf{b})].$$

Hence there exists a set S with cardinality $N = O(K\lambda m/\epsilon)$, such that $E_{\mathbf{b}}[Z(\mathbf{b}, S)] \geq (1 - \epsilon)E_{\mathbf{b}}[Z(\mathbf{b})]$. $\square$

3.3 Numerical Results

We verify the existence of the sparse structure numerically using the constraint sampling method. The purpose of this numerical study is also to show that the number of links per product and stage needed to achieve a specific level of performance increases in the number of stages. To simplify the problem, we only test balanced and identical systems with 10 products and 10 plants in each stage. The demand following the truncated normal distribution conditional on $2\sigma < \mu$ have mean $\mu = 100$ and $CV = 0.4$. The plant capacity $c = 100$. We assume the dedicated structure as the base design. That is, each product has a unique primary plant in the first place. We use the constraint sampling method, which can be referred to [Chou et al. \(2010\)](#), to sample structures.

The algorithm consists of three main steps (for details see [Chou et al. \(2010\)](#)). First, we apply Monte Carlo simulation to estimating $E_{\mathbf{D}}[x_{ij}^{k*}(\mathbf{D})]$ under full flexibility design by generating 1,000 demand instances and then we compute the link probability $p_{ij}^k = \frac{E_{\mathbf{D}}[x_{ij}^{k*}(\mathbf{D})]}{\sum_{i=1}^I \sum_{j=1}^{J_k} \sum_{k=1}^K E_{\mathbf{D}}[x_{ij}^{k*}(\mathbf{D})]}$ based on $E_{\mathbf{D}}[x_{ij}^{k*}(\mathbf{D})]$. We use $\tilde{p}_{ij}^k$ to designate the estimated p_{ij}^k . That is, the larger the expected flow that a link bears, the more chance it will stand to be sampled. Second, we sample the links according to the estimated link probability $\tilde{p}_{ij}^k$ until the number of links sampled reaches our budget N , i.e. the number of links allowed to add. Here $\tilde{p}_{ij}^k$ means the probability that is obtained in step 1. By repeating step 2, we sample $100K$

structures. Because the stage-spanning and floating bottlenecks deteriorate the performance of a multistage systems as K increases (Graves and Tomlin, 2003), we let the number of structures sampled increase in K here so that we can have more structures to compare and select. Third, we generate 500 demand instances to evaluate the $100K$ structures and pick the one with the maximal average output. Clearly, the algorithm becomes more and more time-consuming as K increases. In addition, given the random nature of the constraint sampling method, there exist scenarios where the links sampled cluster in some stages, which make the performance even worse.

In case that links cluster together, instead we attribute equal number of links to each stage in step 2. Since we do not take into consideration the interstage logistics routes, the structure for each stage can be sampled independently. In other words, we first sample the structure $\mathcal{G}_1$ for the first stage and then $\mathcal{G}_2$ for the second stage, and so on. In each stage, the constraint sampling method is implemented just as a single stage system. The constraint sampling algorithm can be implemented in this way partly due to the systems that we are investigating in this section are balanced and identical. That is, stages are repetitions of each other. The combined structure $\mathcal{G} = \{\mathcal{G}_1, \dots, \mathcal{G}_K\}$ becomes the flexibility design for the multistage system.

For a K stage system, the number of links to be added starts from $10K$ and increases with an increment of 10. We test systems with 2, 3, 4, and 5 stages, respectively. We summarize our result in Figure 3.3.1

Figure 3.3.1 shows clearly that the systems with more stages have to employ more links per product on average than those with less stages in order to achieve

Figure 3.3.1: Average shortfalls of multistage systems

the same level of shortfalls. Despite that difference, the performances of all multistage systems converge to full flexibility as the number of links increases. We can see that when links per product per stage on average reaches 3, the shortfall for all systems actually is quite small and nearly equal to that of full flexibility. Notice full flexibility needs 10 links per product. So the structures sampled by constraint sampling are still sparse enough.

3.4 Worst Performance Evaluation with Unknown Distribution

Knowing the exact demand distribution is usually not the case in practice. Many companies only have some historical demand data on hand but do not know the exact forms of the demand distributions. However, it is not difficult to obtain the information of the first and second moments usually. In such manufacturing systems, we face the difficulty in evaluating the performance of a flexibility structure since simulation by sampling demand realizations is no longer possible.

To our best knowledge no one considers the case of uncertain capacity in the process flexibility literature. It is true that capacity could be random variables in reality and such case has been studied in the literature, like random yield. However, it is extremely difficult to evaluate the performance of such a flexible system with both demand and capacity uncertainty.

In this section, our goal is to present a method to calculate the lower bound of the performance of a flexible system assuming that the first and second moments of the demand distributions are known. That is, we are going to calculate the bound for the worst performance of all demand distributions with the same first and second moments. We refer to such a problem as *worst performance evaluation* problem. The method can be extended to the case wherein the capacities are random and the information of the first and second moments is known.

Recall that we have proved in Lemma 3.2.2 that for a given structure and a demand instance, $D(S)$ is a mixed 0-1 linear program. Obviously, the performance can be obtained by solving a mixed 0-1 linear program under objective uncertainty.

3.4 Worst Performance Evaluation with Unknown Distribution

We find the connection of such a problem with the ones in network reliability and project management. As to compute the expected maximal objective value has been a basic focus in network reliability and project management, we are going to adopt the approach recently developed by [Natarajan et al. \(2011\)](#) to address the worst case expected output of the problem below.

$$\begin{aligned}
 (S) \quad Z(\mathbf{b}, S) &= \min \mathbf{b}^T \mathbf{w} \\
 \text{s.t.} \quad &A_S^T \mathbf{w} \geq \mathbf{0} \\
 &B^T \mathbf{w} \geq \mathbf{1} \\
 &w_j \in \{0, 1\} \quad \forall j = 1, \dots, m
 \end{aligned}$$

3.4.1 Completely Positive Representation

The problem of interest in the network reliability literature is in the form of :

$$\begin{aligned}
 Z(\tilde{\mathbf{c}}) &= \max \tilde{\mathbf{c}}^T \mathbf{x} \\
 \text{s.t.} \quad &\mathbf{a}_i^T \mathbf{x} = b_i \quad \forall i = 1, \dots, m \\
 &\mathbf{x} \geq \mathbf{0} \\
 &x_j \in \{0, 1\} \quad \forall j \in \mathcal{B} \subseteq \{1, \dots, n\}
 \end{aligned}$$

where $\tilde{\mathbf{c}}$ is the random objective coefficient vector with known first moment vector $\boldsymbol{\mu} = \mathbf{E}[\tilde{\mathbf{c}}]$ and second moment matrix $\boldsymbol{\Sigma} = \mathbf{E}[\tilde{\mathbf{c}}\tilde{\mathbf{c}}^T]$. The object is to compute

3.4 Worst Performance Evaluation with Unknown Distribution

the expected maximal objective value, $E[Z(\tilde{\mathbf{c}})]$. [Natarajan et al. \(2011\)](#) formulate this type of problem as follows

$$(P) : \quad Z_P = \sup_{\tilde{\mathbf{c}} \sim (\boldsymbol{\mu}, \boldsymbol{\Sigma})^+} E \left[\max_{\mathbf{x} \in \mathcal{L} \cap \{0,1\}^{|\mathcal{B}|}} \tilde{\mathbf{c}}^T \mathbf{x} \right] \quad (3.10)$$

where

$$\mathcal{L} := \{ \mathbf{x} \geq \mathbf{0} \mid \mathbf{a}_i^T \mathbf{x} = b_i \quad \forall i = 1, \dots, m \},$$

With three key assumptions:

1. The set of distributions of $\tilde{\mathbf{c}}$, with finite mean vector $\boldsymbol{\mu}$ and finite moment matrix $\boldsymbol{\Sigma}$, is nonempty.
2. $x \in \mathcal{L} \Rightarrow x_j \leq 1, \forall j \in \mathcal{B}$
3. The feasible region $\mathcal{L} \cap \{0, 1\}^{|\mathcal{B}|}$ is nonempty and bounded.

[Natarajan et al. \(2011\)](#) then construct a completely positive program (C) which

3.4 Worst Performance Evaluation with Unknown Distribution

is equivalent to (P).

$$(C) : \quad Z_C = \max I_n \bullet Y \quad (3.11)$$

$$\text{s.t.} \quad \mathbf{a}_i^T \mathbf{p} = b_i \quad \forall i = 1, \dots, m \quad (3.12)$$

$$\mathbf{a}_i^T X \mathbf{a}_i = b_i^2 \quad \forall i = 1, \dots, m \quad (3.13)$$

$$X_{jj} = p_j \quad \forall j \in \mathcal{B} \subseteq \{1, \dots, n\} \quad (3.14)$$

$$\begin{bmatrix} 1 & \boldsymbol{\mu}^T & \mathbf{p}^T \\ \boldsymbol{\mu} & \Sigma & Y^T \\ \mathbf{p} & Y & X \end{bmatrix} \succeq_{cp} 0 \quad (3.15)$$

where

$$\mathbf{p} \doteq \mathbf{E}[\mathbf{x}(\tilde{\mathbf{c}})],$$

$$Y \doteq \mathbf{E}[\mathbf{x}(\tilde{\mathbf{c}})\tilde{\mathbf{c}}^T],$$

$$X \doteq \mathbf{E}[\mathbf{x}(\tilde{\mathbf{c}})\mathbf{x}(\tilde{\mathbf{c}})],$$

The above formulation is referred to as *Completely Positive Cross Moment Model* (CPCMM).

[Natarajan et al. \(2011\)](#) show that $Z_P \leq Z_C$ from the construction of the model. That is Z_C is an upper bound of Z_P . Interestingly and importantly, then they manage to prove the equivalence of Z_P and Z_C .

Theorem 3.4.1. ([Natarajan et al. \(2011\)](#)) Under Assumptions 1, 2, and 3, $Z_P = Z_C$

Notice that the equivalence of Z_P and Z_C needs equality constraints in the

3.4 Worst Performance Evaluation with Unknown Distribution

formulation of $Z(\tilde{\mathbf{c}})$. If the constraints in $\mathcal{L}$ are not tight, the corresponding Z_C , a completely positive model which is the same as (C) except that the constraints 3.12 and 3.13 are not tight, only serves as an upper bound of Z_P with inequality constraints. One way to find the equivalent completely positive model for such a problem (P) with inequality constraints is to add slack variables in the inequality constraints to convert them to equality constraints. Then we have a new program (P') and we can construct a completely positive model (C') for it. However, this method will result in a larger completely positive matrix in (C') in comparison to the one in (C).

3.4.2 CPPCMM for Process Flexibility

Notable is the fact that though the above (P) program is in the *max* form, [Natarajan et al. \(2011\)](#) claim that it can be naturally extended to the *min* form from the proofs of the theorem in that paper, which is the form of our performance evaluation problem.

Here we present our worst performance evaluation problem. Suppose we have the first moment vector of $\mathbf{b}$, $\boldsymbol{\mu} = \mathbf{E}[\mathbf{b}]$ and the second moment matrix $\boldsymbol{\Sigma} = \mathbf{E}[\mathbf{b}\mathbf{b}^T]$ in a flexibility design. Then the worst performance of this flexibility design among all of the possible distributions with $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ is:

$$(P2) : \quad Z_P(\mathbf{b}, S) = \inf_{\mathbf{b} \sim (\boldsymbol{\mu}, \boldsymbol{\Sigma})^+} \mathbf{E} \left[\min_{w \in \mathcal{L} \cap \{0,1\}^{|\mathcal{B}|}} \mathbf{b}^T w \right] \quad (3.16)$$

where

$$\mathcal{L} := \{w \geq 0 \mid A^T w \geq 0, B^T w \geq \mathbf{1}\},$$

3.4 Worst Performance Evaluation with Unknown Distribution

Let $\mathbf{r}$ denote the right-hand side vector of (S) and Let $w(\mathbf{b})$ denote an optimal solution to problem (S). It is clear that $w(\mathbf{b})$ is also random. Let c_j , $j = 1, \dots, J$ denote the capacity elements in $\mathbf{b}$. c_j in fact is the plant capacity in lexicographical order. In other words, each c_j , $j = 1, \dots, J$ corresponds to some c_{pq}^k , $p = 1, \dots, I$, $k = 1, \dots, K$, $q = 1, \dots, J_k$. Applying the same processes in [Natarajan et al. \(2011\)](#) we construct a completely positive model for (P2).

$$(C2) : \quad Z_C(\mathbf{b}, S) = \min I_{(K+1)I+J} \bullet Y \quad (3.17)$$

$$\text{s.t.} \quad [A_S \ B]^T \mathbf{p} \geq \mathbf{r} \quad \forall i = 1, \dots, m \quad (3.18)$$

$$[A_S \ B]^T X [A_S \ B] \geq \mathbf{r}^2 \quad \forall i = 1, \dots, m \quad (3.19)$$

$$X_{jj} = p_j \quad \forall j \in \mathcal{B} \subseteq \{1, \dots, (K+1)I+J\} \quad (3.20)$$

$$\begin{bmatrix} 1 & \boldsymbol{\mu}^T & \mathbf{p}^T \\ \boldsymbol{\mu} & \boldsymbol{\Sigma} & Y^T \\ \mathbf{p} & Y & X \end{bmatrix} \succeq_{cp} 0 \quad (3.21)$$

where

$$\mathbf{p} \doteq \mathbf{E}[w(\mathbf{b})],$$

$$Y \doteq \mathbf{E}[w(\mathbf{b})\mathbf{b}^T],$$

$$X \doteq \mathbf{E}[w(\mathbf{b})w(\mathbf{b})],$$

The size of the completely positive matrix in (C2) is $[2(K+1)I+2J+1] \times [2(K+1)I+2J+1]$, which could be very large when the flexible system consists

3.4 Worst Performance Evaluation with Unknown Distribution

of many products and stages. It is desirable to reduce the size of the completely positive matrix if possible. [Natarajan et al. \(2011\)](#) discuss the dimension reduction issue and show that the dimension can be reduced by the number of the constant coefficients in the objective function of (P). Back to our problem (P2), there are KI coefficients which are 0 while the J plant capacity coefficients could be constant if we assume there is no randomness for them. Suppose the demand vector $\mathbf{D} \sim (\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and assume capacities are deterministic. Following the steps in [Natarajan et al. \(2011\)](#), we reduce the dimension to $[(K+2)I + J + 1] \times [(K+2)I + J + 1]$. Consequently, we have the following CPCMM:

$$\begin{aligned}
 (CC2) : \quad Z_{CC}(\mathbf{b}, S) &= \min \sum_{i=1}^I Y_{ii} + \sum_{j=1}^J c_j p_j \\
 \text{s.t.} \quad &[A_s \ B]^T \mathbf{p} \geq \mathbf{r} \forall i = 1, \dots, m \\
 &[A \ B]^T X [A_s \ B] \geq \mathbf{r}^2 \forall i = 1, \dots, m \\
 &X_{jj} = p_j \quad \forall j \in \mathcal{B} \subseteq \{1, \dots, (K+1)I + J\} \\
 &\begin{bmatrix} 1 & \boldsymbol{\mu}_1^T & \mathbf{p}^T \\ \boldsymbol{\mu}_1 & \boldsymbol{\Sigma}_1 & Y^T \\ \mathbf{p} & Y & X \end{bmatrix} \succeq_{cp} 0
 \end{aligned}$$

The CPCMM model is a good representation of the problem (P2). Especially the information of the first and second moments is contained in the completely positive matrix. More importantly, both the capacities and demands are allowed to be correlated to each other. From this perspective, CPCMM is more attrac-

3.4 Worst Performance Evaluation with Unknown Distribution

tive than Marginal Moment Model (MMM) (Bertsimas et al., 2004, 2006), which ignores the correlation. However, converting the initial problem to CPCMM does not reduce the difficulty of the problem. To solve the CPCMM, Natarajan et al. (2011) further relax the completely positive constraints from $\succeq_{cp} 0$ to doubly nonnegative constraints $\succeq 0$ and ≥ 0 . We denote the relaxed model of (C2) as (CR2) and the relaxation of (CC2) as (CCR2). Let $Z_{CR}(\mathbf{b}, S)$ and $Z_{CCR}(\mathbf{b}, S)$ denote the optimal objective value of (CR2) and (CCR2), respectively. We have the following result.

Theorem 3.4.2. *Under Assumptions 1, 2, and 3, $Z_P(\mathbf{b}, S) \geq Z_{CC}(\mathbf{b}, S) \geq Z_{CCR}(\mathbf{b}, S)$.*

Proof Since our problem of interest is in the setting of process flexibility, the nonempty coefficients distributions always have nonnegative support, finite $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Hence Assumption 1 is fulfilled. It is trivial to verify Assumptions 2 and 3.

Following the construction steps in Natarajan et al. (2011), we have $Z_P(\mathbf{b}, S) \geq Z_{CC}(\mathbf{b}, S)$. And because the feasible region of $Z_{CCR}(\mathbf{b}, S)$ is relaxed from $Z_{CC}(\mathbf{b}, S)$, it follows that $Z_{CC}(\mathbf{b}, S) \geq Z_{CCR}(\mathbf{b}, S)$. $\square$

(CCR2) is a semidefinite program that can be solved by standard solvers in polynomial time but provide weaker bounds. We next use (CCR2) to numerically generate the bounds for the worst performances.

3.4.3 Numerical Test

We test 5×5 , 10×10 and 20×20 identical 2-chain system with $\mu = c = 10$ and $\sigma = 3, 4, 5$, respectively. To compare, we also use the truncated normal distribu-

3.4 Worst Performance Evaluation with Unknown Distribution

tion with mean μ and standard deviation σ , and $2\sigma \leq \mu$ as a benchmark. To simplify the numerical test, the demands are identical and independent of each other. That is, the covariances are 0. Table (3.1) showcases the expected outputs calculated by CPCMM and Table (3.2) shows the expected outputs of the benchmark systems with the truncated normal distributions. We can see that when the system size is small (for example 5×5), the expected output by CPCMM is closer to the benchmark expected output. This is consistent with the observation in the literature that the 2-chain efficiency (Chou et al. (2010)) is decreasing in the system size. In other words, whatever the true demand distribution is, 2-chain can achieve almost the same benefits as full flexibility when the system size is small. It implies that the expected output of CPCMM, which is the lower bound for all distributions, for a small system is also close to the benchmark - the normal distribution. The numerical test is conducted on a Lenovo X61 notebook with Intel Core 2 Duo Merom T7500 2.2 G, RAM 2GB, Microsoft Windows XP Professional SP3. The SDP is solved using CVX in MatLab environment, a package for specifying and solving convex programs (Grant and Boyd, 2008, 2014).

Remark 3.4.1. *In the CPCMM model, if a structure $S_1 \subseteq S_2$, then $Z(\tilde{c}, S_1) \leq Z(\tilde{c}, S_2)$. In other words, the objective value of CPCMM is non-decreasing in the number of arcs. This is because as we incrementally add arcs the feasible region of CPCMM becomes smaller and smaller. This totally resolves the worry that CPCMM is not able to discern the small increment in the number of links.*

3.4 Worst Performance Evaluation with Unknown Distribution

CV	Expected Output			CPU time		
	0.3	0.4	0.5	0.3	0.4	0.5
5×5 chaining	43	40	37.5	2.1	2.1	2.1
10×10 chaining	85	80	75	16.3	15.3	15.9
20×20 chaining	170	160	150	1539.4	1558.6	1546.0

Table 3.1: Performances and CPU time of the SDP

CV	Expected Output		
	0.3	0.4	0.5
5×5 chaining	48	47	46
10×10 chaining	97	95	93
20×20 chaining	194	191	187

Table 3.2: Benchmark performances by simulation

3.4.4 Flexibility Structure Design

The constraint sampling algorithm is developed by (Chou et al., 2010) under the assumption that demand distributions are known. Steps 1 and 3 of the algorithm need generating demand instances based on the distribution information. Without the exact demand distribution information, of course we can still use historical data to both design and evaluate a structure. But the structure might be biased. Another disadvantage of constrain sampling is the computing time. The algorithm is time consuming particularly in the evaluation part. The computing time increases as the number of structures or demand realizations sampled increases. It is still acceptable in one stage problem in contrast to multistage case. For example, whereas we only use 500 demand instances in Section 3.2, the computing time already amounts to around one hour for a five stage 10 products balanced and identical system.

Given the information of the first and second moments, the VHC method proposed in Chapter 1 can be applied to design flexibility structure for single stage problem. The procedure will be the same as in Chapter 1 since VHC method only needs group coefficient of variation to assort the products and plants. However, so far we have not extended it to multistage systems.

So if it is a single stage problem, we recommend using VHC method to design the flexibility structure with the following steps:

1. Use the historical data to estimate the first and second moments
2. Generate a hub-and-chain structure using VHC method
3. Evaluate previously sampled structure by CPCMM method.

If it is a multistage problem with unknown demand distributions, we have to use the constraint sampling method.

1. Use the historical data to estimate the first and second moments
2. Use the historical demand instances to do constraint sampling to get a specified number of structures
3. Evaluate previously sampled structures by CPCMM method and choose the best structure.

3.5 Conclusion

In this chapter we have studied the multistage process flexibility problem. We have shown, similar to the single stage case, the existence of a sparse process flexibility

structure that can achieve $1 - \epsilon$ benefits of full flexibility for the multistage system. However, the number of links needed is related to the number of stages. This is verified by our numerical test using the constraint sampling method.

We then find that the dual of our model is a 0-1 integer program. We verify that the dual of our model satisfies the assumptions of a CPCMM method in the literature. Thus it solves the worst performance evaluation problem for systems with incomplete demand distribution information. The performance evaluated by CPCMM serves as a lower bound for the performances of all possible demand distributions with the first and second moment information.

4

Production Planning Using Real Options

4.1 Introduction

The Internet has significantly changed the landscape of business in the past two decades. Online sales have flourished in many countries. Nowadays B2C platforms such as taobao and ebay, have millions of registered retailers and manufacturers. The manufacturers like Lenovo also have constructed their own internet sales systems on their website (<http://shop.lenovo.com.cn/>) to complement the traditional brick-and-mortar retail stores. As the traditional communication obstacles between the manufacturer and her customers no longer exist, interactions in real time now are possible. On one hand, promotions, release of new products can easily reach customers through online sale channels and email subscriptions. On the other hand, feedback, ratings about the products from customers also are

easily available to the manufacturer. For instance on the aforementioned Lenovo online shop, customers can contact the sales force directly through instant messaging systems. The significantly lowered communication cost and fast speed of information transmission encourage the manufacturer to acquire demand information from customers through the Internet, especially when she needs to plan the capacity for a new product before its release. This reduces the overage and underage costs in the future. More importantly, innovative business models have continuously been developed in the new business landscape.

On October 12 2012, Lenovo launched an ultrabook Ideapad Yoga 13 in China, which is equipped with a touchscreen and can impressively fold 360 degree into four modes. Instead of accepting orders directly after the debut, Lenovo set the period from October 12 to 31 as a pre-order season both on shop.lenovo.com.cn and Tmall.com. Customers who paid 100 yuan as a nonrefundable reservation fee during the pre-order season were given an option of purchasing one Yoga 13 after October 31 at a price 6599 yuan in comparison with the formal launch price 6999 yuan for customers who purchased after October 31. Lenovo also guaranteed that they would refund the price difference if any discount offered in 60 days after the date of order placement. Subsequently, customers made decisions to exercise the option or otherwise abandon their rights with 100 lost in the following November. The option is not transferable.

Unlike the traditional advance selling which secures the demand in the advance selling season by giving customers a discount or premium, Lenovo offered a real option, consisting of an option price of 100 yuan and a strike price of 6599 yuan, that allows customers to mitigate their consumption risk in the future. Notice

that consumers are uncertain about their valuations on the product until the uncertainty resolved in the spot market, e.g. November in Lenovo's example. From Lenovo's perspective, although the option cannot secure all option buyers in the spot market, Lenovo obtains benefit which is twofold as follows. On one hand, the exercise price or strike price (6599 yuan) is lower than the spot price (6999 yuan). The promotion would encourage customers to buy. Also Lenovo secures the nonrefundable option price already whatever the customers' moves are. In this regard, option selling serves as an instrument for revenue management. On the other hand, more importantly Lenovo can observe the reaction of the market to the product and then possess better demand information about the spot market, especially when the demand of informed customers (defined as customers aware of option selling) is correlated to uninformed customers. The information, as Lenovo is a manufacturing company, helps Lenovo to better plan her production or capacity reservation from OEM, if the production is outsourced.

Options, in contrast to advance selling, can discriminate market segments. We will show that manufacturers can offer different options to the market and induce customers in a specific segment to buy an option tailored for them without incurring cross-buying behavior. This finding implies that using options can save marketing expenses on discriminating segments such as premiums, verifying customer identity for identity-based incentives etc.

Lenovo has to make the decision of option pricing and production planning jointly, which has three dimensions, the option price, strike price and order quantity and has not been explored in the academic community. By introducing the option scheme, advance selling in fact is a special option with 0 strike price. Gal-

lego and Şahin (2010) point out that in airline industry partially refundable fares, which have long been offered by carriers, are actually options. However, in the revenue management setting, carriers' capacity, say the number of flight seats, is fixed. In this chapter, we aim to study the joint decision making of option pricing and production planning and to provide analytical justification for the option-selling scheme. Although we use manufacturing companies as our primary research background, our research can be applied to other industries as long as customers face uncertain valuations and the industry players need to do capacity planning.

Our main contributions are as follows. First, we show under what conditions option selling is better than the newsvendor or advance selling. Second, we identify the optimal order and option policies, respectively. In particular, there are three policies for a market with two segments. We characterize the optimal prices and production quantity for each policy. Furthermore, we compare among the three policies and identify the conditions for dominance. Third, we study the effect of correlated informed and uninformed customers on the optimal prices and the production quantity. We find in some circumstances providing all customers with the option-selling information is not optimal for the manufacturer.

The remainder of this chapter is organized as follows. We review the literature in Section 2. A fluid model focusing on the single option scheme is introduced and analyzed in Section 3. In Section 4 we investigate the case of offering two options to two segments under the fluid model. We then switch to a binomial model emphasizing analysis on informed and uninformed customers with the same valuation distribution in Section 5. Section 6 concludes this chapter.

4.2 Literature Review

Unlike the finance literature focusing on pricing of options traded on open options exchanges, most of options related literature in operations management is about designing business-to-business option contracts, the bilateral contracts signed between the supplier and the manufacturer or the manufacturer and the retailer. [Ritchken and Tapiero \(1986\)](#) consider using contingent contracts (options) in inventory planning to reduce risk by determining the number of options to buy at the beginning of the selling period, assuming the option and strike prices are exogenously given. [Cachon and Lariviere \(2001\)](#) study a contract including firm commitment and options of contingent purchase in the future offered by the manufacturer to supplier. The supplier may have the same information about the demand as the manufacturer or not. In addition, they consider two compliance regimes which are the forced compliance and the voluntary compliance. The forced compliance means the supplier must provide sufficient capacity to cover the maximum possible order quantity while the voluntary compliance need not. The forced compliance could be interpreted as the penalty cost is infinite and this case is similar to one policy in this essay called the ordering to upper bound. [Zhao et al. \(2010\)](#) show the option contract of capacity reservation can coordinate the supply chain through a game theory approach. The manufacturer in fact also executes the forced compliance policy due to the market power of her retailers. [Cheng et al. \(2003\)](#) study a supplier who offers call or put options to her buyer as an supplement to the committed order. Both of the order decision for the manufacturer and pricing of the option for the supplier are examined. [Barnes-Schuster](#)

[et al. \(2002\)](#) analyze a similar model with two periods. The manufacturer can place order at the beginning of each period and buy options at the beginning of period 1. Options may be exercised in period 2. They also allow the demands in the two periods are correlated so that the demand information in period 2 can be updated after the demand in period 1 is materialized. In this essay we also assume the informed and uninformed market segments are correlated. However the difference is in our essay the demand of informed customers only could be observed if they are induced to buy options in period 1 otherwise they postpone their purchases to period 2. Besides, the supplier has two production modes in [Barnes-Schuster et al. \(2002\)](#). The supplier produces the committed orders placed at the beginning of period 1 and 2 using a cheaper production mode. In contrast, the supplier undertakes additional production using an expensive production mode after options are exercised in period 2. In our essay, we also study an expedite model which is equivalent to the expensive production mode here. The supplier incurs more production cost due to rescheduling the production or overtime pay etc. This model can be justified by the electronic industry. For example customers can pay an extra amount of money to expedite the deliveries of a product.

[Wu et al. \(2002\)](#) study a two-stage Stackelberg game in which a seller offers options of capacity reservation to her buyer. Each option gives the buyer the right of reserving one unit capacity for the selling season. The buyer in the selling season chooses the number of options to exercise and also determines the quantity to purchase from the spot market. Using backward method the buyer determines the capacity reserved based on the option price and the strike price. At last the seller chooses the optimal option price and strike price. The spot price is an

exogenous random variable. The buyer's objective is to maximize her utility. A problem in the same setting is studied by Kleindorfer and Wu (2003) and they provide a good review on the related literature as well. Then Wu and Kleindorfer (2005) extend it to a multi-seller and single-buyer setting. Golovachkina and Bradley (2002) also study a similar problem without the utility framework. Fu et al. (2010) consider a procurement portfolio problem that the buyer has three sources of supply, fixed price contract, options and spot market. The spot price is also a random variable. An algorithm is developed to find out the best portfolio.

Other studies about options in supply chain include option contract with downward sloping demand (Burnetas and Ritchken, 2005), a risk averse newsvendor (Chen and Parlar, 2007), a retailer using option contracts with service level commitment (Chen and Shen, 2012), and multisupplier competing for a single buyer Martínez-de-Albéniz and Simchi-Levi (2009).

Gallego and Şahin (2010) view a partially refundable ticket as a real option sold from a carrier to a customer and in this regard the fully refundable and nonrefundable fares are special cases of the option. Facing uncertain travel valuations, the partially refundable fares are more attractive since they provide a new way to balance between the consumption and price risk. In consistence with the revenue management literature, the capacity is assumed to be fixed. They show the optimal option prices and timing to issue as well as the social optimality. Besides, the fluid model is justified using the Poisson distribution. They show the solution obtained by the fluid model is also asymptotically optimal for a related stochastic problem.

Another stream of research related to this essay is advance selling from re-

tailers to customers. [Weng and Parlar \(1999\)](#) examine a problem using prior-sale discount to attract customers, who have a specific probability of purchase, to be early committed, thereby reducing the variance associated with the binomial size of market. [Tang et al. \(2004\)](#) study an advance-selling problem with two brands which are substitutable. Assuming the market sizes of the two brands are correlated, an discount advance sale for brand A updates the demand information of both brand A and B. An extension of [Tang et al. \(2004\)](#) is [McCardle et al. \(2004\)](#). The previous two papers do not adopt the willingness-to-pay scheme for the customers. Instead, the fraction of customers who purchase in advance is determined by a function of the discount factor. [Prasad et al. \(2011\)](#) consider an advance-selling problem with correlated informed and uninformed customers. All customers have the same willingness-to-pay distribution such that the total number of buyers follows the binomial distribution, which can be approximated by the normal distribution if the market size is large enough. After observing the number of customers who commit early purchase, the demand information for the uninformed segment is subsequently updated. In this sense, this essay is similar to [Prasad et al. \(2011\)](#). However, we will show that the problem in [Prasad et al. \(2011\)](#) is a special case of the general option selling and that option selling dominates advance selling.

[Xie and Shugan \(2001\)](#) study a retailer's best strategy of advance selling with customers who are uncertain about the future consumption states. [Yu et al. \(2011\)](#) study how different types of customer valuation interdependence models impact the seller's strategy and profits. In this essay we adopt both of the deterministic model (fluid model) and heterogeneous valuation model (binomial model) but we

allow the binomial model to be approximated by the normal distribution. [Cho and Tang \(2013\)](#) study three selling strategies including advance selling, regular selling and dynamic selling of a manufacturer who produces and sells a seasonal product in a supply chain setting. They examine the policies from the manufacturer's perspective and the retailer's perspective.

4.3 Model Settings - Single Option

Consider a manufacturer sells a product and faces i market segments over two periods. The market size M_i of the i th segment follows distribution $G_i(\cdot)$. Although we use the normal distribution $N(\mu_i, \sigma_i^2)$ in our analysis, most of our results are applicable to other distributions. Risk-neutral customers in the i th market segment have an identical valuation distribution $F_i(\cdot)$ and are independent of each other. $F_i(\cdot)$ is known to both of the manufacturer and customers. To simplify the problem we assume the standard deviation of the customer valuation or willingness-to-pay V_i is τ identically across all market segments while $E[V_i] = \nu_i$ and $\nu_i \geq \nu_{i+1}$. The manufacturer announces x, p, p_s at the beginning of period 1 (see Figure 4.3.1). In period 1, after comparing the expected consumer surplus of buying in period 1 with period 2, customers can either pay price x_i to buy an option, which gives customers the right to buy one unit of the product at price p_i in period 2, or postpone their purchases to period 2 at price p_s (see Figure 4.3.2). However, customers choosing to buy in period 2 face a stockout rate η . Since customers' decisions are made based on expectations and the customer valuation distributions are identical, customers in segment i either all buy a same

4.3 Model Settings - Single Option

option or none of them will buy. After updating the market size of segment i , if any, the manufacturer makes a decision on the production quantity or capacity reservation at the end of period 1. For the sake of simplicity, we use order quantity q to represent the production quantity or the capacity reserved. Products are produced in period 2 and delivered to customers at the end of period 2. At the beginning of period 2, the customer valuations on the product are resolved. Then customers who have purchased option x_i in period 1 choose to exercise the option if their realized valuation v is no less than the strike price p_i . Meanwhile spot customers choose to buy if their valuation v is no less than the spot price p_s . We assume any excess inventory at the end of period 2 has no salvage value. The strategy of the manufacturer is to induce customers to buy options in period 1 so that she can obtain better information about the market size.

Figure 4.3.1: Manufacturer's decision process

4.3 Model Settings - Single Option

x	option price
p	strike price
p_s	price at spot market
V_i	willingness-to-pay of market i ; $V_i \sim N(\nu_i, \tau^2)$
v_i	realization of the willingness-to-pay of market i
$F_i(\cdot)$	cdf V_i
M_i	i th maket size, $M_i \sim N(\mu_i, \sigma_i^2)$
m_i	realization of the i th market size
$G_i(\cdot)$	cdf of M_i
c_a	penalty cost for lost sale of advance market
c_s	penalty cost for lost sale of spot market $c_s \leq c_a$
c_m	manufacturing cost per unit of product
η	stockout rate
N	amount of options sold out
q	production quantity planned
$\phi(\cdot)$	density function of the standard normal distribution
$\Phi(\cdot)$	cdf of the standard normal distribution
z	safety factor equal to $\Phi^{-1}\left(\frac{p_s+c_s-c_m}{p_s+c_s}\right)$
z_a	safety factor for option buying equal to $\Phi^{-1}\left(\frac{p+c_a-c_m}{p+c_a}\right)$

Table 4.1: Notations

Throughout this chapter random variables are denoted by upper case letters while their realizations are denoted by lower case letters. Symbols with hat are related to the demand in period 2, namely the spot market. And symbols with tide overhead are related to the newsvendor scheme. We abuse the notations when there is no comparison between some sections.

4.3.1 Single Market Model

In this section we consider the case of single market segment. Through the remaining of this section, we will ignore the subscript i for the sake of simplicity.

Customers choose to buy options only when $\int_p^{+\infty} (v - p)f(v)dv - x \geq (1 -$

Figure 4.3.2: Customer's decision process

$\eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv$. Otherwise, the manufacturer still faces a newsvendor problem since no customer buys in period 1. Notice that $\int_p^{+\infty} (v - p)f(v)dv - x = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv$ is the indifference value curve of customers (see Figure 4.3.3). (x, p) on and below the curve is able to induce customers to buy this option in the advance market while (x, p) above the curve fails. Let $x(p)$ denote a function of p and $x(p) = \int_p^{+\infty} (v - p)f(v)dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv$ which is the solution of x given p for the indifference value curve. So $(x(p), p)$ is on the indifference value curve. If there exist multiple valuation distributions, we will use subscripts to discriminate them. Notice because of the stockout rate η , interestingly x could be negative as p approaches toward infinity. This indicates that the manufacturer in fact will choose to pay the customers to obtain demand

information in some circumstances. In this chapter, we first focus on the case $x, p \geq 0$.

Proposition 4.3.1. $\int_p^{+\infty} (v - p)f(v)dv - x$ is decreasing and convex in x, p .

$$V \sim N(100, 400); p_s = 120, \eta = 0.05.$$

Figure 4.3.3: Indifference curve

The manufacturer's expected profit in the option case can be formulated in the following way:

4.3 Model Settings - Single Option

$$\Pi = \max_{x,p} E_M [\hat{\pi}(x, p, M)], \quad (4.1)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p)f(v)dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv, \quad (4.2)$$

$$x, p \geq 0. \quad (4.3)$$

where $\hat{\pi}(x, p, m) = \max_q mx + p \min(q, m\bar{F}(p)) - c_a(m\bar{F}(p) - q)^+ - c_m q$. Since customers are identical and independent of each other, the manufacturer observes the real market size as long as all of the customers are induced to buy in period 1. Knowing the market size m , we can see that the optimal order quantity $q^* = m\bar{F}(p)$ results in a profit of $\hat{\pi}(x, p, m) = mx + m\bar{F}(p)(p - c_m)$. Thus the optimal expected profit is given by

$$\Pi = \max_{x,p} E_M [Mx + M\bar{F}(p)(p - c_m)], \quad (4.4)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p)f(v)dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv,$$

$$x, p \geq 0.$$

Proposition 4.3.2. *The expected revenue $\mu x + \mu p\bar{F}(p)$ is decreasing with respect to p . The expected profit is increasing for $p \in [0, c_m]$ and decreasing for $p \in (c_m, +\infty)$.*

Based on this proposition the optimal option is obtained as follows,

Theorem 4.3.1. *The optimal solution to problem (4.1) is $p^* = c_m$, $x^* = \int_{c_m}^{+\infty} (v - c_m)f(v)dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv$. The optimal expected profit is μx^* . And*

4.3 Model Settings - Single Option

the optimal order quantity $q^* = m\bar{F}(p^*)$ given m .

Proof Taking expectation with respect to M does not impact the choice of x^* , p^* . Thus problem (4.4) becomes

$$\begin{aligned} & \max_{x,p} \mu x + \mu \bar{F}(p)(p - c_m), \\ \text{s.t.} \quad & \int_p^{+\infty} (v - p)f(v)dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv, \\ & 0 \leq p \leq p_s, x \geq 0. \end{aligned} \tag{4.5}$$

Besides, x^* , p^* always make (4.2) binding, otherwise the seller can obtain higher profit by increasing x . Replace x in (4.5) with $\int_p^{+\infty} (v - p)f(v)dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv$. Then the derivative with respect to p is given by $f(p)(c_m - p)$. That is to say, the objective is increasing in p when $p \leq c_m$ while decreasing when $p > c_m$. Therefore, $p^* = c_m$ and $x^* = \int_{c_m}^{+\infty} (v - c_m)f(v)dv + (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv$ $\square$

Figure 4.3.4 illustrates the optimal solution in the single market case.

As the derivative of the manufacturer's expected profit with respect to p is increasing when $p < c_m$ and decreasing when $p > c_m$, one of the options $(x(0), 0)$ and $(0, p_s)$ will be the lower bound of profit for all feasible options.

A question arises when we investigate the option strategy. Does it dominate the newsvendor strategy? Obviously, if the manufacturer chooses the newsvendor model, her expected profit will be

$$\Pi_N = \max_{q \geq 0} p_s E_M [\min(q, M\bar{F}(p_s))] - c_s E_M [M\bar{F}(p_s) - q]^+ - c_m q$$

Figure 4.3.4: Optimal decision for single market

Π_N can be easily transformed to $\max_{q \geq 0} (p_s - c_m) \mu \bar{F}(p_s) - L(q)$, where $L(q) = E_{\hat{M}} \left[(p_s + c_s - c_m)(\hat{M} - q)^+ + c_m(q - \hat{M})^+ \right]$ and $\hat{M} = M \bar{F}(p_s)$. It can be shown its optimal solution is $q_N^* = G^{-1}(\frac{p_s + c_s - c_m}{p_s + c_s}) \bar{F}(p_s) = (\mu + z\sigma) \bar{F}(p_s)$, where $z = \Phi^{-1}(\frac{p_s + c_s - c_m}{p_s + c_s})$. And the optimal profit is $\pi_N = [(p_s - c_m)\mu - (p_s + c_s)\sigma\phi(z)] \bar{F}(p_s)$

Theorem 4.3.2. *In the single market case, $\Pi > \Pi_N$.*

Proof Clearly, $(0, p_s)$ is always a feasible option that satisfies condition (4.2). Given $(0, p_s)$, all customers will purchase in period 1. Therefore, it allows the manufacturer to know exactly the realized demand information m and consequently to order a quantity $m \bar{F}(p_s)$ which incurs neither overage nor underage

cost. Also the manufacturer collects p_s from an option buyer which is the same as the spot customer. By adopting option selling, the manufacturer's long-run profit will be $(p_s - c_m)\mu\bar{F}(p_s)$. Interestingly, we can see this is the first part of $\Pi_N = \max_{q \geq 0} (p_s - c_m)\mu\bar{F}(p_s) - L(q)$. Since the sum of underage and overage cost $L(q) > 0$ definitely, we have $\Pi > \Pi_N$. $\square$

Theorem (4.3.2) claims that the real option model dominates the newsvendor model. Moreover, as shown by the proof process, even for a nonoptimal but feasible option, we already have the result $\pi > \pi_n$. In other words, the optimal option given in Theorem (4.3.1) will be even better. The manufacturer should adopt the real option method once necessary conditions satisfied (e.g. online channel)

Notice that both of the optimal option price and strike price are positive. Prasad et al. (2011) show that advance selling through discount in the advance season could improve the profit over the newsvendor model under the condition that the difference between the mean of valuations for a product and the consumer surplus of buying in the spot market is greater than a threshold. Because a discount in the advance season, $(x, 0)$, is only a special case of our model but the optimal option is not in the form of $(x, 0)$, we believe that option selling with positive strike price can gain more profit than discount.

4.3.1.1 Advance Selling

Advance selling is very common in practice. As a special case of the option policy, it secures all the revenue in the advance-selling season. The manufacturer's expected profit is $(x(0) - c_m)\mu$. As $x(0)$ is not related to c_m , once c_m exceeds

$x(0)$ the manufacturer cannot make any profit. So for products with $c_m > x(0)$ especially $c_m > \nu$, the manufacturer should not adopt advance selling.

The advantage of the option policy is it can adjust according to the production cost. Whatever c_m is, the manufacturer can obtain a marginal profit equal to the option price and offset the production cost by charging a strike price c_m .

4.3.2 Single Option with Two Market Segments

Now we consider the case where the manufacturer faces two market segments. The distribution $F_i(\cdot)$ of the willingness-to-pay V_i in market segment i is normal. First we define four price thresholds,

$$\bar{p}_1 = \{p \mid \int_p^{+\infty} (v - p) f_1(v) dv = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv\}, \quad (4.6)$$

$$\bar{p}_2 = \{p \mid \int_p^{+\infty} (v - p) f_2(v) dv = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv\}, \quad (4.7)$$

$$\bar{p} = \{p \mid \int_p^{+\infty} (v - p) (f_1(v) - f_2(v)) dv = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) (f_1(v) - f_2(v)) dv\}. \quad (4.8)$$

$\bar{p}_1$ and $\bar{p}_2$ are the intersections of the indifference curves of segment 1 and 2 with the horizontal axis, respectively. $\bar{p}$ is the intersection of the two indifference curves.

Proposition 4.3.3. *Suppose $\nu_1 > \nu_2$, when $p \leq \bar{p}$, if $\int_p^{+\infty} (v - p) f_2(v) dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv$, then $\int_p^{+\infty} (v - p) f_1(v) dv - x > (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$. Equivalently, if $\int_p^{+\infty} (v - p) f_1(v) dv - x \leq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$, then $\int_p^{+\infty} (v - p) f_2(v) dv - x < (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv$. When $p > \bar{p}$, the*

4.3 Model Settings - Single Option

situation is reversed if $\int_p^{+\infty} (v-p)f_1(v)dv - x \geq (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_1(v)dv$, then $\int_p^{+\infty} (v-p)f_2(v)dv - x > (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_2(v)dv$.

Proof Given $\int_p^{+\infty} (v-p)f_2(v)dv - x \geq (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_2(v)dv$,

$$\begin{aligned} \int_p^{+\infty} (v-p)f_1(v)dv - x &\geq \int_p^{+\infty} (v-p)f_1(v)dv + (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_2(v)dv \\ &\quad - \int_p^{+\infty} (v-p)f_2(v)dv, \\ &= \int_p^{+\infty} (v-p)(f_1(v) - f_2(v))dv - (1-\eta) \int_{p_s}^{+\infty} (v-p_s)(f_1(v) \\ &\quad - f_2(v))dv + (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_1(v)dv, \\ &> (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_1(v)dv, \end{aligned}$$

The last step is because $\int_p^{+\infty} (v-p)(f_1(v) - f_2(v))dv$ is a decreasing function of p in the interval $[0, p_s]$. For the case $p > p_s$, we can use the same method to prove that $\int_p^{+\infty} (v-p)f_1(v)dv - x \geq (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_1(v)dv$ yields $\int_p^{+\infty} (v-p)f_2(v)dv - x > (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f_2(v)dv$. $\square$

Proposition 4.3.3 shows that the two curves intersect at the point $\bar{p}$. On the left of $\bar{p}$, the indifference curve of segment 1 should be over the curve of segment 2 while on the right of $\bar{p}$ the curve of segment 2 over the curve of segment 1. This result is shown in Figure (4.3.5). Meanwhile, The proof of Proposition 4.3.3 implies that the intersection of the two curves is below the horizontal axis when $\bar{p} > \bar{p}_1$ while above the horizontal axis when $\bar{p} < \bar{p}_2$.

Based on the fact that when $p > \bar{p}$ the indifference curve of segment 2 is above

that of segment 1, and $(x_1(p_s), p_s)$ is above the indifference curve of segment 2, we can show by contradiction that

Corollary 4.3.1. $p_s < \bar{p}$.

Given $\nu_1 > \nu_2$ and $\bar{p}_1 < \bar{p}$, the two indifference value curves do not cross each other above the horizontal axis. As shown in problem (4.1), all optimal options for any segment are on the boundary of (4.2). It implies that the optimal option for segment 1 is not feasible for segment 2. In other words, customers in segment 2 will not buy an option which is optimal for segment 1.

Likewise, if $\bar{p}_2 > \bar{p}$, the intersection of the two indifference curves are above the horizontal axis. (x, p) on the indifference curve with $p < \bar{p}$ is not able to induce segment 2 to buy. And (x, p) , optimal for segment 2 with $p > \bar{p}$, fails to induce segment 1 to buy. We summarize this in the following corollary.

Corollary 4.3.2. (*Boundary condition*) Suppose $\nu_1 > \nu_2$, when $p \leq \bar{p}$ options that make $\int_p^{+\infty} (v - p)f_1(v)dv - x = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f_1(v)dv$ are not able to induce customers in segment 2 to buy. Likewise, when $p > \bar{p}$ options satisfying $\int_p^{+\infty} (v - p)f_2(v)dv - x = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f_1(v)dv$ are not able to induce customers in segment 2 to buy

Proof suppose $\int_{p^*}^{+\infty} (v - p^*)f_1(v)dv - x^* = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f_1(v)dv$. We can show $\int_{p^*}^{+\infty} (v - p^*)f_2(v)dv - x^* < (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f_2(v)dv$ is equivalent to $\int_{p^*}^{+\infty} (v - p^*)(f_1(v) - f_2(v))dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)(f_1(v) - f_2(v))dv > 0$, which has been shown in Proposition (4.3.3). $\square$

Next we focus our analysis on the case wherein $p \leq p_s$, although whenever possible we will also analyze the case where $p > p_s$. In fact, we believe $p \leq p_s$ is true for most manufacturing settings, since the valuation span on a product

$$V_1 \sim N(120, 400), V_2 \sim N(100, 400); p_s = 120, \eta = 0.05.$$

Figure 4.3.5: Indifference valuation curves of two segments

will not be as large as a flight ticket. Customers very possibly are reluctant to pay more than the spot price to secure a product and meanwhile it is difficult to prevent customers from switching to the spot market by cheating on their identities if $p > p_s$. Despite these facts, the case where $p > p_s$ is still possible and useful for niche market, for example luxury products. We leave the case where $p > p_s$ to future research. Based on Proposition (4.3.3), there are three cases if the manufacturer decides to use real options.

4.3.2.1 Case 1.

If the manufacturer fails to incentivize the customers to buy options, the situation will be just the same as the newsvendor problem.

$$\Pi_N = \max_{q \geq 0} E_{\hat{M}} \left[p_s \min(q, \hat{M}) - c_s(\hat{M} - q)^+ \right] - c_m q, \quad (4.9)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p) f_1(v) dv - x < (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv, \quad (4.10)$$

$$\int_p^{+\infty} (v - p) f_2(v) dv - x < (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv, \quad (4.11)$$

$$x, p \geq 0. \quad (4.12)$$

where $\hat{M} = M_1 \bar{F}_1(p_s) + M_2 \bar{F}_2(p_s) \sim N(\mu_1 \bar{F}_1(p_s) + \mu_2 \bar{F}_2(p_s), \sigma_1^2 \bar{F}_1(p_s)^2 + \sigma_2^2 \bar{F}_2(p_s)^2)$.

So the manufacturer's expected profit is

$$\max_{q \geq 0} (p_s - c_m)(\mu_1 \bar{F}_1(p_s) + \mu_2 \bar{F}_2(p_s)) - L(q), \quad (4.13)$$

where $L(q) = E_{\hat{M}} \left[(p_s + c_s - c_m)(\hat{M} - q)^+ + c_m(q - \hat{M})^+ \right]$.

4.3.2.2 Case 2.

Only customers in one segment buy the option. As we can see from Figure (4.3.5), an option in the area encircled by the indifference curves and the vertical axis is only able to induce segment 1 to buy. In comparison, the area encircled by the indifference curves and the horizontal axis represents that only segment 2 is induced to buy. Let us first examine the situation where only customers in segment 1 buy the option. Now the manufacturer obtains the information about the size of segment 1 yet does not know the exact size of segment 2.

$$\Pi_1 = \max_{x,p} E_{M_1} [\hat{\pi}_1(x, p, M_1)], \quad (4.14)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p) f_1(v) dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv, \quad (4.15)$$

$$\int_p^{+\infty} (v - p) f_2(v) dv - x < (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv, \quad (4.16)$$

$$x, p \geq 0,$$

where

$$\begin{aligned} \hat{\pi}_1(x, p, m_1) = & \max_{q \geq 0} m_1 x - c_m q + p \min(q, m_1 \bar{F}_1(p)) - c_a(m_1 \bar{F}_1(p) - q)^+ \\ & + E_{M_2}[p_s \min((q - m_1 \bar{F}_1(p))^+, M_2 \bar{F}_2(p_s)) \\ & - c_s(M_2 \bar{F}_2(p_s) - (q - m_1 \bar{F}_1(p))^+)^+]. \end{aligned}$$

4.3 Model Settings - Single Option

It can be shown by contradiction that $q \geq m_1 \bar{F}_1(p)$ such that

$$\hat{\pi}_1(x, p, m_1) = \max_{\hat{q} \geq 0} E_{M_2} \left[p_s \min(\hat{q}, \hat{M}) - c_s(\hat{M} - \hat{q})^+ \right] - c_m \hat{q} + K, \quad (4.17)$$

where $K = (p - c_m)m_1 \bar{F}_1(p) + m_1 x$ and $\hat{q} = q - m_1 \bar{F}_1(p)$, $\hat{M} = M_2 \bar{F}_2(p_s)$.

So the optimal expected profit can be transformed to

$$\Pi_1 = \max_{\hat{q} \geq 0} (p - c_m)\mu_1 \bar{F}_1(p) + \mu_1 x_1(p) + (p_s - c_m)\mu_2 \bar{F}_2(p_s) - L(\hat{q}) \quad (4.18)$$

where $L(\hat{q}) = E_{\hat{M}} \left[(p_s + c_s - c_m)(\hat{M} - \hat{q})^+ + c_m(\hat{q} - \hat{M})^+ \right]$.

Obviously, $\hat{\pi}_1(x, p, m_1)$ has been divided into two separable subproblems, i.e. the real option and the newsvendor for segment 1 and 2, respectively. Therefore we have:

Proposition 4.3.4. *If only customers in segment 1 are induced to buy the option, $p^* = c_m$ and $x_1^* = \int_{c_m}^{+\infty} (v - c_m)f_1(v)dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f_1(v)dv$. $q^* = m_1 \bar{F}_1(p^*) + G_2^{-1}\left(\frac{p_s + c_s - c_m}{p_s + c_s}\right) \bar{F}_2(p_s)$,*

Figure 4.3.6 illustrates the optimal solution to this case.

Next we investigate the case wherein only customers in segment 2 buy the

Figure 4.3.6: Optimal decision when only high segment buy

option. We can only achieve this situation when $p > \bar{p}$ and $\bar{p} < \bar{p}_2$.

$$\Pi_1 = \max_{x,p} E_{M_2} [\hat{\pi}_1(x, p, M_2)], \quad (4.19)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p) f_1(v) dv - x < (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv, \quad (4.20)$$

$$\int_p^{+\infty} (v - p) f_2(v) dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv, \quad (4.21)$$

$$x, p \geq 0,$$

Similarly we can show

$$\hat{\pi}_1(x, p, m_2) = \max_{\hat{q} \geq 0} E_{M_1} \left[p_s \min(\hat{q}, \hat{M}) - c_s(\hat{M} - \hat{q})^+ \right] - c_m \hat{q} + K, \quad (4.22)$$

where $K = (p - c_m)m_2\bar{F}_2(p) + m_2x$ and $\hat{q} = q - m_2\bar{F}_2(p)$, $\hat{M} = M_1\bar{F}_1(p_s)$. The optimal solution can be obtained by setting constraint 4.21 binding. As shown in the previous section, K is decreasing in p when $p > c_m$. And it is trivial to show $\bar{p} > p_s$. Therefore, the optimal solution in this case should be $p = \bar{p}$ and $x = \int_{\bar{p}}^{+\infty} (v - \bar{p})f_2(v)dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f_2(v)dv$. But this optimum also can induce segment 1 to buy. Thus we add an infinitesimal increment to $\bar{p}$ such that only segment 2 buy this option. The optimum now is $(x(\bar{p} + \epsilon), \bar{p} + \epsilon)$. Since we concentrate on $p \leq p_s$ case. We will not extensively analyze this case.

Intuitively, the option provided reduces the uncertainty about the demand information. In the single-market and single-option situation, we show that offering an option can achieve better profits. And in the single-option and two-segment case, the problem is divided into two separate subproblems. The first one is a single-market and single-option problem while the second one is a newsvendor problem associated with less uncertainty. So the profit extracted from each segment will be improved. We show this result in the following theorem.

In addition to the profit improvement, providing an option is a good way to discriminate the market segments. In practice, a manufacturer relies on means like identity discount, membership, service or facility to do price discrimination, which need extra support from operation department. Our option method will automatically create a rate fence to prevent the higher segment from buying in

the spot market.

Theorem 4.3.3. *For a market with two segments, the profit by offering single option to the market as in Case 2, $\Pi_1 > \Pi_N$.*

Proof Since we abuse the notations in this chapter. we have to use several new notations temporarily in this proof in order to compare several cases.

The expected newsvendor profit $\max_{q \geq 0} (p_s - c_m)(\mu_1 \bar{F}_1(p_s) + \mu_2 \bar{F}_2(p_s)) - L(q)$, where $L(q) = E_{\hat{M}}[(p_s + c_s - c_m)(\hat{M} - q)^+ + c_m(q - \hat{M})^+]$. The first part of π_N is already fixed. $L(q)$, the sum of overage and underage cost, solely depends on the variance of $\hat{M}$, which is $\sigma_1^2 \bar{F}_1(p_s)^2 + \sigma_2^2 \bar{F}_2(p_s)^2$ and we denote it as σ_a^2 temporarily in this proof.

In Case 2, firstly when $\bar{p} > \bar{p}_1$ it is only possible that segment 1 buy the option while segment 2 not. the manufacturer's profit is $\Pi_1 = \max_{\hat{q} \geq 0} (p - c_m)\mu_1 \bar{F}_1(p) + \mu_1 x + (p_s - c_m)(\mu_2 \bar{F}_2(p)) - L(\hat{q})$ where $L(\hat{q}) = E_{\hat{M}}[(p_s + c_s - c_m)(\hat{M} - \hat{q})^+ + c_m(\hat{q} - \hat{M})^+]$. Notice that $(0, \bar{p}_1)$ is a feasible solution for this situation, in other words only segment 1 buying the option and $\bar{p}_1 > p_s$. So $\Pi_1 > \max_{\hat{q} \geq 0} (\bar{p}_1 - c_m)\mu_1 \bar{F}_1(p) + (p_s - c_m)(\mu_2 \bar{F}_2(p)) - L(\hat{q}) > \max_{\hat{q} \geq 0} (p_s - c_m)(\mu_1 \bar{F}_1(p) + \mu_2 \bar{F}_2(p)) - L(\hat{q})$. The variance of $\hat{M}$ in this case is $\sigma_2^2 \bar{F}_2(p_s)^2$ which we denote it as σ_{b1}^2 . Clearly σ_{b1}^2 is less than σ_a^2 . So in this subcase $\pi > \pi_N$. Secondly when $\bar{p} < \bar{p}_2$ we also have $(0, \bar{p} - \epsilon)$ and $(0, \bar{p} + \epsilon)$ as feasible options, although the former induces segment 1 to buy while the latter induces segment 2. And $\bar{p} \pm \epsilon > p_s$. So $\Pi_1 > (p_s - c_m)(\mu_1 \bar{F}_1(p) + \mu_2 \bar{F}_2(p)) - \max_{\hat{q} \geq 0} L(\hat{q})$. Similarly we can show $\sigma_{b2}^2 = \sigma_1^2 \bar{F}_1(p_s)^2$. Clearly the variances in both subcases are smaller than in the newsvendor case. So $\Pi_1 > \Pi_N$.

□

4.3.2.3 Case 3.

All customers buy the option.

$$\Pi_2 = \max_{x,p} E_N [\hat{\pi}_2(x, p, N)], \quad (4.23)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p) f_1(v) dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv, \quad (4.24)$$

$$\int_p^{+\infty} (v - p) f_2(v) dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv, \quad (4.25)$$

$$x, p \geq 0,$$

where $N = M_1 + M_2$. Let n denote the realization of N and $\hat{N} = (M_1|n)\bar{F}_1(p) + (M_2|n)\bar{F}_2(p)$ and $(M_1|n) + (M_2|n) = n$. $\hat{N}$ is the amount of customers who exercise the option in period 2. Based on Proposition 4.3.3, condition (4.24) is redundant when $p \leq \bar{p}$ while condition (4.25) is redundant when $p > \bar{p}$.

We know

$$\hat{\pi}_2(x, p, n) = \max_{q \geq 0} E_{\hat{N}} \left[p \min(q, \hat{N}) - c_a(\hat{N} - q)^+ \right] - c_m q + n x(p). \quad (4.26)$$

where x could be $x_2(p)$ if $p \leq \bar{p}$ or $x_1(p)$ if $p > \bar{p}$.

As all customers in the market buy the option, the manufacturer observes n , the size of the whole market, but loses the exact information about the size of segment 1. As a result she has to guess the proportion of customers in segment 1 or 2. This prompts us to think about the tradeoff between Cases 2 and 3, in other words, choosing accurate information about a portion of subsegments but

less accurate about the whole market or more accurate information about the whole market but less accurate about subsegments.

(4.26) is a newsvendor problem with $\hat{N}$ as the random variable. To solve it, we should find out the conditional distribution of $\hat{N}$. This is a classic signal extraction problem of Kalman filtering problem. It can be shown that $E[M_i|n] = \mu_i + (n - \mu_1 - \mu_2) \frac{\sigma_i^2}{\sigma_1^2 + \sigma_2^2}$, $Var(M_i|n) = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$ and $M_i|n$ is still normal, $i = 1, 2$. Then, $\hat{N} \sim N(\sum E[M_i|n] \bar{F}_i(p), \sigma_{\hat{N}}^2)$ where $\sigma_{\hat{N}} = \frac{\sigma_1 \sigma_2 (\bar{F}_1(p) - \bar{F}_2(p))}{\sqrt{\sigma_1^2 + \sigma_2^2}}$. $\hat{N}$ is truncated within $[n\bar{F}_2(p), n\bar{F}_1(p)]$. Since $\sigma_{\hat{N}}$ is not related to M_1 and M_2 , $E[\sigma_{\hat{N}}] = \sigma_{\hat{N}}$.

Consequently the expected optimal profit for the manufacturer is

$$\Pi_2 = \max_{q, x, p \geq 0} (p - c_m)(\mu_1 \bar{F}_1(p) + \mu_2 \bar{F}_2(p)) + (\mu_1 + \mu_2)x(p) - L(q), \quad (4.27)$$

where $L(q) = E_N \left[E_{\hat{N}} \left[(p + c_a - c_m)(\hat{N} - q)^+ + c_m(q - \hat{N})^+ \right] \right] = (p + c_a)\phi(z_a)\sigma_{\hat{N}}$.

The optimal solution can be obtained by the first order condition of Π_2 with respect to p , but there is no explicit closed form.

We restrict our analysis to $p \leq p_s$. Notice that if $c_a \rightarrow c_s$, the underage and overage costs are less than the counterparts of the newsvendor when the manufacturer uses option $(0, p_s)$, because the variance of Case 3, $\frac{\sigma_1^2 \sigma_2^2 (\bar{F}_1(p) - \bar{F}_2(p))^2}{\sigma_1^2 + \sigma_2^2}$, is less than $\sigma_1^2 \bar{F}_1(p_s)^2 + \sigma_2^2 \bar{F}_2(p_s)^2$ of the newsvendor. So the optimal expected profit is still greater than that of the newsvendor in this situation.

Theorem 4.3.4. $\Pi_2 > \Pi_1$ if there exists p such that $\mu_1 \Delta_{M_1}(p) < \mu_2 \Delta_{M_2}(p) + k\sigma_2 \bar{F}_2(p_s) - k_a \sigma_{\hat{N}}$. where $\Delta_{M_1}(p) = x_1(c_m) - x_2(p) - (p - c_m)\bar{F}_1(p)$ and $\Delta_{M_2}(p) = (p - c_m)\bar{F}_2(p) + x_2(p) - (p_s - c_m)\bar{F}_2(p_s)$, $k = (p_s + c_s)\phi(z)$, $k_a = (p + c_a)\phi(z_a)$.

Theorem (4.3.4) is derived from the comparison between (4.18) and (4.27).

4.4 Two Options with Two Market Segments

$\Delta_{M_1}(p)$ is the marginal revenue of switching segment 1 from option selling in Case 2 to the one in Case 3. $\Delta_{M_2}(p)$ is the marginal revenue of switching segment 2 from the newsvendor in Case 2 to option selling in Case 3. And $k\sigma_2\bar{F}_2(p_s) - k_a\sigma_{\hat{N}}$ is the change of underage and overage costs from Case 2 to Case 3.

However, it is hard to compare the expected profit of Case 3 with the former two cases generally speaking. The cost $L(q)$ in Case 3 by and large depends on p and c_a . As we have not imposed any restrictions on c_a , we are not able to claim Case 3 is better than Case 1 or Case 2.

Consider the situation where c_a is formidable, like the case where the option is a legal contract and any breach of the contract may cause lawsuit. Then $c_a \gg c_s$. The manufacturer has to order $n\bar{F}_1(p)$ in order to avoid any underage cost. It can be shown that the expected overage cost $c_m(q - \hat{N})^+$ equals $c_m\mu_2(\bar{F}_1(p) - \bar{F}_2(p))$.

Interestingly, the profit here seems increasing in the proportion of segment 2 if we fix $\mu_1 + \mu_2 = 1$ and other conditions hold. However, substitute $1 - \mu_1$ for μ_2 and we can see that for any p an increase in μ_1 improves the expected profit. In other words, we always hope the high segment be larger. This is also true for Case 2.

4.4 Two Options with Two Market Segments

As shown in the previous section, offering only one option to the whole market is not enough to be capable of extracting the exact demand information from each segment. In Case 2, the manufacturer still knows little about the demand in segment 2. When either segment 2 accounts for relatively small portion of

4.4 Two Options with Two Market Segments

the market or the demand variation is small, the manufacturer would prefer Case 2. In Case 3, the manufacturer has to lower the option price enough to induce segment 2 to buy the option. This might result in a big loss of profit from segment 1. As a result the manufacturer prefers Case 3 only when segment 1 is not the main revenue source.

Thinking of a situation where demand in either segment has high variation, it will be attractive to know exactly the demand in both segments due to the high underage and overage costs generated by demand variation. When offering only one option might pose the manufacturer in the dilemma of choosing between Case 2 and 3 in the previous section, the manufacturer should try to induce each segment to buy an option tailored for them. In other words, the manufacturer can provide the customers with a set of options. However, to discriminate the two segments, the manufacturer has to ensure that there is no cross buying.

$$\Pi_3 = \max_{x_1, p_1, x_2, p_2} E_{M_1, M_2} [\hat{\pi}_3(x_1, p_1, x_2, p_2, M_1, M_2)], \quad (4.28)$$

$$\text{s.t.} \quad \int_{p_1}^{+\infty} (v - p_1) f_1(v) dv - x_1 \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv, \quad (4.29)$$

$$\int_{p_2}^{+\infty} (v - p_2) f_2(v) dv - x_2 \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv, \quad (4.30)$$

$$\int_{p_1}^{+\infty} (v - p_1) f_1(v) dv - x_1 \geq \int_{p_2}^{+\infty} (v - p_2) f_1(v) dv - x_2, \quad (4.31)$$

$$\int_{p_2}^{+\infty} (v - p_2) f_2(v) dv - x_2 \geq \int_{p_1}^{+\infty} (v - p_1) f_2(v) dv - x_1, \quad (4.32)$$

$$x_1, p_1, x_2, p_2 \geq 0,$$

4.4 Two Options with Two Market Segments

Problem (4.28) has a form that is very similar to the mechanism design problem in game theory (see Rasmusen (2001)). (4.29) and (4.30) are the *participation* constraints which guarantee both segments buy in advance while (4.29) and (4.30) are *self-selection* constraints which guarantee no cross buying.

Suppose the manufacturer successfully induces customers in each segment to buy options in advance, that are different from those of the other segment. Then the manufacturer knows the segment sizes m_1 and m_2 . Therefore, $\hat{\pi}_3(x_1, p_1, x_2, p_2, M_1, M_2)$ becomes $\hat{\pi}_3(x_1, p_1, x_2, p_2, m_1, m_2)$ which has the following form:

$$\begin{aligned} \hat{\pi}_3(x_1, p_1, x_2, p_2, m_1, m_2) = & \max_q m_1 x_1 + m_2 x_2 + p_1 \min(q, m_1 \bar{F}_1(p_1)) \\ & + p_2 \min((q - m_1 \bar{F}_1(p_1))^+, m_2 \bar{F}_2(p_2)) - c_m q \\ & - c_a (q - m_1 \bar{F}_1(p_1) - m_2 \bar{F}_2(p_2))^+ \end{aligned} \quad (4.33)$$

With the segment sizes known and using the same argument as in the single-market and single-option case, the manufacturer will choose an optimal order quantity that does not incur any overage or underage cost. That is $q = m_1 \bar{F}_1(p_1) + m_2 \bar{F}_2(p_2)$. Then $\hat{\pi}_3(x_1, p_1, x_2, p_2, m_1, m_2) = m_1 x_1 + m_1 \bar{F}_1(p_1)(p_1 - c_m) + m_2 x_2 + m_2 \bar{F}_2(p_2)(p_2 - c_m)$. Since in any demand scenario the segment sizes are visible for the manufacturer, taking expectation over M_1 and M_2 does not affect the form of $\hat{\pi}(x_1, p_1, x_2, p_2, m_1, m_2)$. We only need to replace m_i by μ_i . Consequently, the objective function becomes

$$\Pi_3 = \max_{x_1, p_1, x_2, p_2} \mu_1 x_1 + \mu_1 \bar{F}_1(p_1)(p_1 - c_m) + \mu_2 x_2 + \mu_2 \bar{F}_2(p_2)(p_2 - c_m). \quad (4.34)$$

4.4 Two Options with Two Market Segments

Next let us look at the constraints of problem (4.28) and find out the feasible regions.

Proposition 4.4.1. $\int_{p_2}^{p_1} F_1(v)dv \geq x_1 + p_1 - x_2 - p_2 \geq \int_{p_2}^{p_1} F_2(v)dv$

Proof Constraints (4.31) and (4.32) can be transformed in the following way,

$$\begin{aligned} \int_{p_1}^{+\infty} v f_1(v)dv - x_1 - p_1 + p_1 F_1(p_1) &\geq \int_{p_2}^{+\infty} v f_1(v)dv - x_2 - p_2 + p_2 F_1(p_2), \\ \int_{p_2}^{+\infty} v f_2(v)dv - x_2 - p_2 + p_2 F_2(p_2) &\geq \int_{p_1}^{+\infty} v f_2(v)dv - x_1 - p_1 + p_1 F_2(p_1), \end{aligned}$$

Transpose terms and we get

$$\begin{aligned} - \int_{p_2}^{p_1} v f_1(v)dv + p_1 F_1(p_1) - p_2 F_1(p_2) &\geq x_1 + p_1 - x_2 - p_2, \\ x_1 + p_1 - x_2 - p_2 &\geq - \int_{p_2}^{p_1} v f_2(v)dv + p_1 F_2(p_1) - p_2 F_2(p_2), \end{aligned}$$

By integrating by parts, $-\int_{p_2}^{p_1} v f_1(v)dv = p_2 F_1(p_2) - p_1 F_1(p_1) + \int_{p_2}^{p_1} F_1(v)dv$ and $-\int_{p_2}^{p_1} v f_2(v)dv = p_2 F_2(p_2) - p_1 F_2(p_1) + \int_{p_2}^{p_1} F_2(v)dv$, we get the result. $\square$

Since we do not consider the time value of money in our model, Proposition (4.4.1) simply shows that the difference of payment between segment 1 and 2 should be within an interval to prevent customers from cross buying. Next we show the first order conditions and the possible optimal solutions of problem (4.28).

Proposition 4.4.2. *To obtain the optimal solution to problem (4.28), there are four cases, 1) all constraints binding except (4.31), 2) all constraints binding except (4.32), 3) constraints (4.29) and (4.32) binding while (4.30) and (4.31) not, or 4)*

4.4 Two Options with Two Market Segments

vice versa.

Proof Clearly constraints (4.31) and (4.32) should not be binding simultaneously otherwise it becomes the single-option case. In addition to that, one of constraints (4.29) and (4.30) must be binding. If neither of them is binding, then manipulating x_i , p_i can improve the objective value. Thus we exclude these cases.

Let us first look at the necessary conditions we can obtain from Case 1. If only (4.29), (4.30) and (4.31) are binding. Substitute $\int_{p_1}^{+\infty} (v - p_1) f_1(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$ for x_1 and $\int_{p_2}^{+\infty} (v - p_2) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv$ for x_2 in (4.31) and (4.32), respectively. That yields $\int_{p_2}^{+\infty} (v - p_2) (f_1(v) - f_2(v)) dv = (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) (f_1(v) - f_2(v)) dv$, which means $p_2 = \bar{p}$ and $\int_{p_1}^{+\infty} (v - p_1) (f_1(v) - f_2(v)) dv > (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) (f_1(v) - f_2(v)) dv$, which is equivalent to $p_1 < \bar{p}$.

Similarly, for Case 2 we have $x_1 = \int_{p_1}^{+\infty} (v - p_1) f_1(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$ and $x_2 = \int_{p_2}^{+\infty} (v - p_2) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv$ as well as $p_2 < \bar{p}$, $p_1 = \bar{p}$.

Next we prove why only Cases 3 and 4 are possible. Suppose we have only constraints (4.29) and (4.31) are binding and other two constraints (4.30) and (4.32) are strict inequality. Then without affecting x_1, p_1 , increasing x_2 brings more profits to the manufacturer. Thus it is impossible that (4.29) and (4.31) are binding while the other two not. By the same token, we can also prove that having (4.30) and (4.31) binding while (4.29) and (4.32) not is another scenario to get optimal solution. $\square$

Given the four cases of feasible regions in Proposition 4.4.2, we can show the corresponding first order conditions and consequently get the optimal solution.

Case 1, we already know $p_2 = \bar{p}$. Replace x_1 in (4.34) by $\int_{p_1}^{+\infty} (v - p_1) f_1(v) dv -$

4.4 Two Options with Two Market Segments

$(1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$. The first order condition in p_1 is

$$-\mu_1 f_1(p_1)(p_1 - c_m) = 0.$$

Thus the optimal solution for Case 1 is $p_1^* = c_m$, $x_1^* = \int_{c_m}^{+\infty} (v - c_m) f_1(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$ and $p_2^* = \bar{p}$, $x_2^* = \int_{\bar{p}}^{+\infty} (v - \bar{p}) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv$. However, we do not know whether x_2^* is positive or not. Besides, as $p_2^* = \bar{p}$, we can rule out the this case if we limit to $p_i \leq p_s$.

Case 2, by the same method as in Case 1, we have the first order condition $-\mu_2 f_2(p_2)(p_2 - c_m) = 0$. The optimal solution is $p_1^* = \bar{p}$, $x_1^* = \int_{\bar{p}}^{+\infty} (v - \bar{p}) f_1(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$ and $p_2^* = c_m$, $x_2^* = \int_{c_m}^{+\infty} (v - c_m) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv$. Very clearly the point (x_2^*, p_2^*) is below the indifference curve of segment 1. Therefore the customer surplus at (x_1^*, p_1^*) cannot be greater than or equal to at (x_2^*, p_2^*) . In other words, the constraint (4.31) is violated. As a consequence, the optimal options yielded in Case 2 are not feasible for problem (4.28).

Case 3, constraints (4.29) and (4.32) are binding, substitute $\int_{p_2}^{+\infty} (v - p_2) f_2(v) dv - \int_{p_1}^{+\infty} (v - p_1) f_2(v) dv + x_1$ for x_2 and $\int_{p_1}^{+\infty} (v - p_1) f_1(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$ for x_1 in (4.34). Then we have the first order conditions for (4.34),

$$-\mu_1 f_1(p_1)(p_1 - c_m) + \mu_2 (\bar{F}_2(p_1) - \bar{F}_1(p_1)) = 0, \quad (4.35)$$

$$-\mu_2 f_2(p_2)(p_2 - c_m) = 0. \quad (4.36)$$

Let p_1^* denote the solution to (4.35). Since $\bar{F}_2(p_1) - \bar{F}_1(p_1) < 0$, p_1^* must be

4.4 Two Options with Two Market Segments

less than c_m to satisfy (4.35). As a result the optimal solution of (4.34) is p_1^* , $x_1^* = \int_{p_1^*}^{+\infty} (v - p_1^*) f_1(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$, $p_2^* = c_m$, $x_2^* = \int_{c_m}^{+\infty} (v - c_m) f_2(v) dv - \int_{p_1^*}^{+\infty} (v - p_1^*) f_2(v) dv + x_1^*$.

When $p_1 = 0$ the left side of constraint (4.35), $-\mu_1 f_1(p_1)(p_1 - c_m) + \mu_2(\bar{F}_2(p_1) - \bar{F}_1(p_1))$, is positive while negative when $p_1 = c_m$. We can also show that the left side is continuous in p_1 . So according to the intermediate value theorem, there exists $0 < p_1^* < c_m$ such that $-\mu_1 f_1(p_1^*)(p_1^* - c_m) + \mu_2(\bar{F}_2(p_1^*) - \bar{F}_1(p_1^*)) = 0$.

However, we can show $x_2^* = \int_{c_m}^{+\infty} (v - c_m) f_2(v) dv - \int_{p_1^*}^{+\infty} (v - p_1^*) f_2(v) dv + x_1^* = \int_{c_m}^{+\infty} (v - c_m) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv - (\int_{p_1^*}^{+\infty} (v - p_1^*) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv) + x_1^* = x_2(c_m) + x_1(p_1^*) - x_2(p_2^*) > x_2(c_m)$. Obviously, this violates the constraint (4.30). So Case 3 cannot be the optimal solution.

Case 4, constraints (4.30) and (4.31) binding, similarly we have the first order conditions for (4.34),

$$-\mu_1 f_1(p_1)(p_1 - c_m) = 0, \quad (4.37)$$

$$-\mu_2 f_2(p_2)(p_2 - c_m) + \mu_1(\bar{F}_1(p_2) - \bar{F}_2(p_2)) = 0. \quad (4.38)$$

Denote the strike price of the optimal option for segment 2 by p_2^* and p_2^* is obtained by solving (4.38). Then

4.4 Two Options with Two Market Segments

Theorem 4.4.1.

$$x_2^* = \int_{p_2^*}^{+\infty} (v - p_2^*) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv, \quad (4.39)$$

$$p_1^* = c_m, \quad (4.40)$$

$$x_1^* = \int_{c_m}^{+\infty} (v - c_m) f_1(v) dv - \int_{p_2^*}^{+\infty} (v - p_2^*) f_1(v) dv + x_2^*. \quad (4.41)$$

Figure 4.4.1: Optimal decision for the two-market and two-option case

It is trivial to show that x_1^* is equal to $x_1(c_m) - (x_1(p_2^*) - x_2(p_2^*))$. Hence $x_2(c_m) < x_1^* < x_1(c_m)$. Obviously $\bar{F}_1(p_2) - \bar{F}_2(p_2) > 0$, therefore $p_2^* > c_m$ to satisfy condition (4.38). At the same time, $-\mu_2 f_2(p_2)(p_2 - c_m) + \mu_1(\bar{F}_1(p_2) - \bar{F}_2(p_2))$ is

4.4 Two Options with Two Market Segments

continuous in p_2 . If the value of $-\mu_2 f_2(p_2)(p_2 - c_m) + \mu_1(\bar{F}_1(p_2) - \bar{F}_2(p_2))$ at p_s is negative. p_2^* must be less than p_s and satisfy (4.38). In contrast, If the value at p_s is nonnegative. p_2^* can be p_s when either $-\mu_2 f_2(p_2)(p_2 - c_m) + \mu_1(\bar{F}_1(p_2) - \bar{F}_2(p_2))$ is positive $\forall p_2 \in (c_m, p_s)$, or the profit at points satisfying (4.38) within (c_m, p_s) is still less than at p_s . Whether $-\mu_2 f_2(p_2)(p_2 - c_m) + \mu_1(\bar{F}_1(p_2) - \bar{F}_2(p_2))$ is negative or not depends on many factors. The ratio of μ_1, μ_2 , the density function of the willingness-to-pay of segment 2 as well as the positions of ν_1 and ν_2 all affect its value.

Corollary 4.4.1. $x_2(c_m) < x_1^* < x_1(c_m), p_2^* > c_m$.

4.4.1 Comparison with The Newsvendor and Single Option

As Cases 2 and 3 are not possible to yield optimal solution, we only need go to Cases 1 and 4 to obtain the optimum. Furthermore, if we assume $p \leq p_s$, Case 1 is also no longer feasible since the two indifference curves never intersect on the left of p_s . The solution worked out in Case 4 then will be optimal for problem (4.28). With the optimum known, we can do some interesting comparisons with the previous schemes. Summarized in the following theorem, we show that the two-option policy is better than the newsvendor in terms of profit.

Theorem 4.4.2. *Given $\nu_1 > \nu_2$, offering two options, which are capable of discriminating the two market segments, dominates the the newsvendor scheme (4.9).*

Proof Suppose the manufacturer offers two options with $(x_2(p_s), p_s)$ for segment 2 and (x_1, c_m) for segment 1, where we use $x_1 = \int_{c_m}^{+\infty} (v - c_m) f_1(v) dv -$

4.4 Two Options with Two Market Segments

$\int_{p_s}^{+\infty} (v - p_s) f_1(v) dv + x_2(p_s)$ which derives from constraint (4.31) when it is binding. Next we show $x_2(c_m) < x_1 < x_1(c_m)$.

Manipulate x_1 to $\int_{c_m}^{+\infty} (v - c_m) f_2(v) dv + \int_{c_m}^{+\infty} (v - c_m) f_1(v) dv - \int_{c_m}^{+\infty} (v - c_m) f_2(v) dv - \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv + \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv = x_2(c_m) + \int_{c_m}^{+\infty} (v - c_m) (f_1(v) - f_2(v)) dv - \int_{p_s}^{+\infty} (v - p_s) (f_1(v) - f_2(v)) dv$. As $\int_p^{+\infty} (v - p) (f_1(v) - f_2(v)) dv$ is decreasing in p and $\nu_1 > \nu_2$, we obtain $x_1 > x_2(c_m)$. Meanwhile we can show $x_1 = \int_{c_m}^{+\infty} (v - c_m) f_1(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv + (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv - \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv + \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv = x_1(c_m) + \eta \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv - \eta \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$. Since $\eta \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv < \eta \int_{p_s}^{+\infty} (v - p_s) f_1(v) dv$, we have $x_1 < x_1(c_m)$.

$x_2(c_m) < x_1 < x_1(c_m)$ indicates that (x_1, c_m) is on the vertical line segment connecting $(x_2(c_m), c_m)$ and $(x_1(c_m), c_m)$ and located between the two indifference curves. As for customers in segment 1 the consumer surplus of the points in the area between the indifference curves is greater than the points on indifference curve of segment 1, constraint (4.29) is satisfied. Clearly, the two options offered fulfill the left constraints (4.30, 4.31, 4.32). So the two options offered are feasible solution for problem (4.28). Next we show the profit by offering these two options is already greater than that of the newsvendor scheme.

Substitute these two options into the objective function. Notice x_1 can be decomposed into $(p_s - c_m) \bar{F}_1(p_s) + \int_{c_m}^{p_s} v f_1(v) dv - c_m \int_{c_m}^{p_s} f_1(v) dv + \eta \int_{p_s}^{+\infty} (v - p_s) f_2(v) dv$ and $\int_{c_m}^{p_s} v f_1(v) dv > c_m \int_{c_m}^{p_s} f_1(v) dv$. So $\mu_1 x_1 > \mu_1 \bar{F}_1(p_s - c_m)$. Together with $x_2(p_s) > 0$ then we have $\Pi_3 > (\mu_1 \bar{F}_1(p_s) + \mu_2 \bar{F}_2(p_s))(p_s - c_m) > \Pi_N$ for the two market segments setting. $\square$

Although the two-option policy is better than the newsvendor, it is not trivial

4.4 Two Options with Two Market Segments

to claim its dominance over the single-option policy. Let us only consider the situation $p \leq p_s$ in which solutions from Case 1 is no longer feasible. We first compare the two-option policy with Case 2 of the single-option strategy. $\Pi_3 - \Pi_1 = -\mu_1(x_1(p_2^*) - x_2(p_2^*)) + [\mu_2 x_2(p_2^*) + \mu_2 \bar{F}_2(p_2^*)(p_2^* - c_m) - \mu_2 \bar{F}_2(p_s)(p_s - c_m) + L(\hat{q})]$, where $(x_2(p_2^*), p_2^*)$ is the optimal option for segment 2 in two-option scenario, $\hat{q}$ is the order quantity for Case 2 of the single-option scenario. We can see that $\Pi_3 - \Pi_1$ is broken down into two parts. The first part, $-\mu_1(x_1(p_2^*) - x_2(p_2^*))$, is the profit given up from segment 1 to prevent the customers in segment 1 from buying the option for segment 2 and the rest part is the profit improvement by offering the option $(x_2(p_2^*), p_2^*)$ to segment 2 instead of the newsvendor scheme. It can be shown that $(x_2(p_2^*), p_2^*)$ gains more profits than the newsvendor scheme from segment 2. So the two-option policy generates more profits from segment 2 while deteriorates the profit from segment 1 in comparison with Case 1 of the single-option policy. As long as the profit increase from segment 2 can offset the loss of segment 1, the manufacturer will offer two separate options to the market. However the superiority of the two-option policy depends on the size of the two markets as well as the level of demand variation. We summarize this result in the proposition as follows:

Theorem 4.4.3. *The two-option strategy dominates the single-option strategy (Case 2 of the single-option policy) if and only if $\frac{\mu_1}{\mu_2} \leq \frac{x_2(p_2^*) + \Delta(p_2^*) + k\theta_2}{x_1(p_2^*) - x_2(p_2^*)}$, where $\Delta(p_2^*) = \bar{F}_2(p_2^*)(p_2^* - c_m) - \bar{F}_2(p_s)(p_s - c_m)$, $k = (p_s + c_s)\phi(\Phi^{-1}(\frac{p_s + c_s - c_m}{p_s + c_s}))$ and $\theta_2 = \frac{\sigma_2}{\mu_2}$, p_2^* is the optimal strike price for segment 2 in Case 4 of the two-option scenario.*

Theorem 4.4.3 indicates the larger the CV of the size of segment 2, the easier to

4.5 Binomial Model with Informed and Uninformed Customers

fulfill the condition. Besides, the larger the ratio of the expected size of segment 1 to segment 2 is, the more difficult the two-option strategy dominates, because the loss from high segment exceeds the benefit from low segment. Notice that $\Delta(p_2^*)$ is the unit marginal benefit in period 2 and $x_2(p_2^*)$ is the revenue in period 1 from segment 2. Also $k\theta_2$ is the unit revenue saved from demand uncertainty reduction. So $x_2(p_2^*) + \Delta_2(p_2^*) + k\theta_2$ is the unit marginal benefit while the denominator $x_1(p_2^*) - x_2(p_2^*)$ is the unit marginal cost of adopting the two-option policy.

Since we do not have the closed-form optimal solution to Case 3 of the single-option policy, it is unclear whether the two-option policy dominates it. However, according to Theorem 4.4.3, if the condition is fulfilled, then we only need to compare the two-option policy to Case 3 of the single-option policy and otherwise we have to compare Case 2 and 3 of the single-option policy. The manufacturer who wants to adopt option selling should always choose between Case 2 and 3 of the single-option policy and the two-option policy.

4.5 Binomial Model with Informed and Uninformed Customers

In the fluid model once the market size is obtained in period 1, the number of customers who exercise the option is deterministic. The manufacturer only needs to focus on pricing options. In this section we continue to explore the option problem yet in different settings. Prasad et al. (2011) study advance selling through discount faced by a newsvendor retailer. In their paper customers with the identical

4.5 Binomial Model with Informed and Uninformed Customers

willingness-to-pay distribution, $F(\cdot)$, are divided into two segments, informed and uninformed customers. Informed customers could obtain the advance-selling information from TV, advertise, etc. and consequently make the decisions of buying in advance or postponing the purchases to the spot market in the advance-selling season while uninformed customers only buy in the spot market as they miss the advance-selling information. The decision of advance purchase is made by comparing the expected utility of buying in advance with that of buying in the spot market. Since customers have the same willingness-to-pay distribution, all informed customers will choose to buy in advance as long as the expected utility of advance buying is greater than buying in the spot market. However, in the spot market the situation is quite different. As an uninformed customer arrives, she accepts the product at the spot price p_s only if her realized willingness to pay is no less than p_s . Given n uninformed customers in the spot market, the number of customers who eventually buy the product follows the binomial distribution, but nonetheless the authors assume the market is large enough such that the binomial distribution can be well approximated by a normal distribution with mean $n\bar{F}(p_s)$ and variance $n\bar{F}(p_s)F(p_s)$. The probability of buying is $\bar{F}(p_s)$ in contrast to $F(p_s)$ of not buying. The sizes of the informed and uninformed segments follow the normal distributions, $M_i \sim N(\mu_i, \sigma_i^2)$ and $M_u \sim N(\mu_u, \sigma_u^2)$ and are correlated with correlation coefficient ρ .

The rationale of the advantage of discount over the pure newsvendor in Prasad et al. (2011) is the demand uncertainty of the informed segment is eliminated by inducing them to buy and the risk from the uninformed segment is reduced as well because the two segments are correlated so that the manufacturer can update

4.5 Binomial Model with Informed and Uninformed Customers

the demand information about the uninformed segment based on the true size of the informed segment. Besides, in order to attract informed customers to buy in advance, the manufacturer chooses to lower the price in period 1, thereby reducing the marginal profit. But we can see that the total number of customers choosing to buy the product in period 1 and 2 has been increased. When the marginal cost is less than the marginal profit the manufacturer should sell in advance.

Clearly, advance selling is a special case of the option strategy, of which the strike price is 0. In the discount scenario the manufacturer has collected the revenue from the customers buying in advance and consequently are fully committed to delivering the product to them at the end of period 2. So the manufacturer only needs to order a quantity equal to the number of customers buying in advance plus the quantity maximizing the profit from uninformed customers. The later quantity can be obtained by solving a newsvendor problem. In contrast, in the general case where the strike prices are positive the manufacturer faces a much more difficult situation. Because customers buying in advance can still choose not to exercise the option, the order quantity must be able to accommodate customers to exercise the option as well as customers in the uninformed segment who eventually buy the product. The risk pooling effect such that the total order quantity is less than the sum of two separate orders from solving two newsvendor problems tangles with the different lost sale costs of the advance and spot market customers.

Here we follow the similar settings shown above except we still use $\int_p^{+\infty} (v - p)f(v)dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv$ as the condition to justify the success of inducing informed customers to buy in advance. The whole informed segment

4.5 Binomial Model with Informed and Uninformed Customers

will choose to buy the option once the previous condition satisfied as in period 1 the comparison of consumer surpluses is based on expectations. And we continue to assume that customers are risk neutral. Lost sale costs for the option buyer and spot buyer are c_a and c_s , respectively, with $c_a > c_s$.

4.5.1 All Customers Are Informed

If none of the customers are aware of the advance sale, then the problem is simply a newsvendor problem. So we first consider the case where all customers are informed. For simplicity, we omit the uninformed segment M_u which is null here. Customers arrive in period 1 and if all of them buy the option the manufacturer should manage to deliver the products in period 2 and meanwhile maximize her profit. A penalty cost c_a will be incurred for each failure of delivery,

Suppose the number of customers buying the option i.e. the realized market size, is m_i . In period 2 the number of customers to exercise the option, $\hat{M}_i$ follows the normal distribution $N(m_i\bar{F}(p), m_i\bar{F}(p)F(p))$ conditional on $M_i = m_i$. The manufacturer's decision making can be formulated as follows:

$$\Pi_i = \max_{x,p} E_{M_i} \left[\hat{\pi}(x, p, \hat{M}_i) \right], \quad (4.42)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p)f(v)dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s)f(v)dv, \quad (4.43)$$

$$x, p \geq 0, \quad (4.44)$$

where $\hat{\pi}(x, p, (\hat{M}_i|M_i = m_i)) = m_i x + (p - c_m)m_i\bar{F}(p) - \min_q L(q)$ and $L(q) = E_{\hat{M}_i|m_i}[(p + c_a - c_m)(\hat{M}_i - q)^+ + c_m(q - \hat{M}_i)^+]$. Let z_a denote $\Phi^{-1}(\frac{p+c_a-c_m}{p+c_a})$ and

4.5 Binomial Model with Informed and Uninformed Customers

$\hat{\mu}_i \equiv m_i \bar{F}(p)$, $\hat{\sigma}_i \equiv \sqrt{m_i \bar{F}(p) F(p)}$. Conditional on the market size m_i , the optimal order quantity is

$$q^* = \hat{\mu}_i + z_a \hat{\sigma}_i \quad (4.45)$$

and $L(q^*) = (p + c_a) \phi(z_a) \hat{\sigma}_i$.

The optimal option can be obtained when (4.43) is binding. The manufacturer's profit is

$$\Pi_i = \max_{x,p} \mu_i x + (p - c_m) \mu_i \bar{F}(p) - (p + c_a) \phi(z_a) E_{M_i}[\hat{\sigma}_i] \quad (4.46)$$

where $E_{M_i}[\hat{\sigma}_i] \leq \sqrt{\mu_i \bar{F}(p) F(p)}$ according to Jensen's inequality. The first order condition with regard to p is

$$-\mu_i f(p)(p - c_m) - \phi(z_a) E_{M_i}[\hat{\sigma}_i] + \frac{c_m z_a}{p + c_a} E_{M_i}[\hat{\sigma}_i] - \phi(z_a)(p + c_a) E'_{M_i}[\hat{\sigma}_i] = 0 \quad (4.47)$$

$$\phi'(z_a) = -z \phi(z_a) \frac{c_m}{(p + c_a)^2} \frac{1}{\phi(z_a)} = \frac{-c_m z_a}{(p + c_a)^2}$$

The optimal solution can be derived from either the first order condition or the extreme points $(x(0), 0)$ or $(x(p_s), p_s)$, since we are not able to verify the convexity of the objective function.

Suppose $c_m \leq \nu$. Note $\mu_i x + (p - c_m) \mu_i \bar{F}(p)$ is decreasing but $(p + c_a) \phi(z_a) E_{M_i}[\hat{\sigma}_i]$ is increasing in p for $\nu > p > c_m$. So the optimal p is outside interval (c_m, ν) . In contrast, if $c_m \geq \nu$, $\mu_i x + (p - c_m) \mu_i \bar{F}(p)$ is increasing while $(p + c_a) \phi(z_a) E_{M_i}[\hat{\sigma}_i]$ is decreasing for $c_m > p > \nu$. So the optimal p is outside (ν, c_m) .

Proposition 4.5.1. $p^* \leq c_m$ or $p^* \geq \nu$ if $c_m \leq \nu$; $p^* \geq c_m$ or $p^* \leq \nu$ if $c_m > \nu$.

4.5 Binomial Model with Informed and Uninformed Customers

4.5.1.1 When Lost Sale Cost is Formidable

Given option (x, p) and market size m_i , if $c_a \rightarrow +\infty$, then $z_a \rightarrow +\infty$. So the optimal order quantity should be $\hat{\mu}_i + z_a \hat{\sigma}_i = m_i$. In other words, the manufacturer cannot afford any underage cost and thus she always orders the quantity equal to the upper bound of demand. We call this policy the *ordering to upper bound*. In practice, this happens when the product is quite valuable and any failure of delivery may result in huge loss for the customer. In addition to that, the product may be not easily substitutable. This case has been studied in Cachon and Lariviere (2001) as the Forced compliance regime where the manufacturer must reserve enough capacity to cope with any possible demand in the future.

The profit for the manufacturer under the ordering to upper bound policy becomes $\Pi_i = \mu_i x + p \mu_i \bar{F}(p) - c_m \mu_i$. Substitute $\int_p^{+\infty} (v-p)f(v)dv - (1-\eta) \int_{p_s}^{+\infty} (v-p_s)f(v)dv$ for x and take the derivative with respect to p . Then we have $-\mu_i p f(p)$, which is nonpositive for $p \in [0, p_s]$. So clearly the optimal option should be in the form of $(x(0), 0)$, which is exactly the discount scenario.

Proposition 4.5.2. *Advance selling, namely $(x(0), 0)$, is optimal under the ordering to upper bound policy*

Advance selling is equivalent to the optimal option under the ordering to upper bound policy. This is quite interesting and it allows us to compare advance selling with the general option policy.

Given any (x, p) and $c_a < +\infty$, the manufacturer would like to choose an optimal order quantity $q^* \leq m_i$ after observing m_i and calculation based on the newsvendor policy. The ordering to upper bound is like a worst case that the

4.5 Binomial Model with Informed and Uninformed Customers

manufacturer will prefer. This indicates the manufacturer's profit by ordering q^* in this specific situation is greater than or equal to by ordering m_i . So for any (x, p) ordering according to the newsvendor is better than the ordering to upper bound. Furthermore, the manufacturer's optimal profit is obtained by comparing all feasible options including $(x(0), 0)$. It indicates that the manufacturer's optimal profit for $c_a < +\infty$ is greater than or equal to $\mu_i x(0) - c_m \mu_i$. Hence the discount scenario serves as a lower bound of the manufacturer's profit. We summarize this in the following proposition.

Proposition 4.5.3. *For any lost sale cost c_a , the lower bound of the manufacturer's profit by using any option (x, p) that satisfies condition (4.43) is $\mu_i x + p \mu_i \bar{F}(p) - c_m \mu_i$. Furthermore the optimal profit is greater than or equal to that of advance selling.*

Actually when the lost sale cost is tremendously high, the manufacturer does not want to incur any underage cost. However, she still tries to minimize the uncertainty in period 2. So she has to provide an option that secures customers in period 2 and keep the uncertainty as low as possible. Intuitively, this is just like a fair contract in which both parties have equal rights and responsibility. The manufacturer guarantees the delivery of the product while the customer promises full payment in advance. But Proposition (4.5.3) suggests that the manufacturer always seek options with $p > 0$ as her choice whenever $c_a < +\infty$.

4.5.1.2 The Expedite Model

In some circumstances the product is not customized, like commodity, agricultural products, etc. The manufacturer in fact can buy standard products in open market

4.5 Binomial Model with Informed and Uninformed Customers

to cope with shortage. Also for other products, like computers, she can pay extra cost to expedite the production process. The expedite model has been studied in the supply chain contract literature. Suppose the unit cost of purchase in open market or expediting production per unit in period 2 is p_e and $p_e > p_s$. Then $c_s = p_e - p_s$. Similarly, $c_a = p_e - p$.

The manufacturer's profit now is

$$\Pi_i = \max_{x,p} \mu_i x + (p - c_m) \mu_i \bar{F}(p) - p_e \phi(z_a) E_{M_i}[\hat{\sigma}_i], \quad (4.48)$$

where $z_a = \Phi^{-1}\left(\frac{p_e - c_m}{p_e}\right)$. The first order condition is

$$-(p - c_m) \mu_i f(p) - p_e E'_{M_i}[\hat{\sigma}_i] = 0$$

4.5.1.3 Comparing with The Newsvendor

The profit for the newsvendor scheme is

$$\Pi_N = (p_s - c_m) \mu_i \bar{F}(p_s) - (p_s + c_s) \phi(z) \tilde{\sigma}_i, \quad (4.49)$$

where $z = \Phi^{-1}\left(\frac{p_s + c_s - c_m}{p_s + c_s}\right)$ and $\tilde{\sigma}_i = \sqrt{\mu_i \bar{F}(p_s) F(p_s) + \sigma_i^2 \bar{F}^2(p_s)}$.

Obviously, in the setting of the expedite model $(x(p_s), p_s)$ is a feasible solution for the option strategy. And its corresponding profit will be greater than or equal to $(p_s - c_m) \mu_i \bar{F}(p_s) - p_e \phi(z_a) \sqrt{\mu_i \bar{F}(p_s) F(p_s)}$. The profit for the newsvendor scheme is $(p_s - c_m) \mu_i \bar{F}(p_s) - p_e \phi(z) \sqrt{\mu_i \bar{F}(p_s) F(p_s) + \sigma_i^2 \bar{F}^2(p_s)}$. Here $z = z_a$. Notice the variance of demand in the newsvendor, $\mu_i \bar{F}(p_s) F(p_s) + \sigma_i^2 \bar{F}^2(p_s)$, is

4.5 Binomial Model with Informed and Uninformed Customers

greater than $\mu_i \bar{F}(p_s)F(p_s)$ in the option strategy. So the option strategy always dominates the newsvendor.

Proposition 4.5.4. *Under the expedite setting, the manufacturer is better off in terms of expected profit by adopting options than the newsvendor when all customer are informed of option selling.*

However for the discount scenario some conditions are necessary to ensure its dominance over the newsvendor. Prasad et al. (2011) provide a threshold for the difference between the utility of advance buying and spot buying as the condition of advance selling for the manufacturer. Interestingly we have a similar result as follows,

The optimal profit of the manufacturer in the discount scenario is $\mu_i x(0) - c_m \mu_i$. $\Pi_i > \Pi_N$ is equivalent to $\mu_i x(0) - c_m \mu_i > (p_s - c_m) \mu_i \bar{F}(p_s) - (p_s + c_s) \phi(z) \tilde{\sigma}_i$. Cancel μ_i and yield $x(0) > p_s \bar{F}(p_s) + c_m F(p_s) - (p_s + c_s) \phi(z) \tilde{\sigma}_i / \mu_i$. We know $x(0) = \nu - (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f(v) dv$ and ν is the expected valuation. Let ν_w denote $(1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f(v) dv$ which is the expected consumer surplus for delaying purchase to period 2. Define a threshold $\Delta \nu_i = p_s \bar{F}(p_s) + c_m F(p_s) - (p_s + c_s) \phi(z) \tilde{\sigma}_i / \mu_i$. So we have

Proposition 4.5.5. *Advance selling dominates the newsvendor if and only if $\nu - \nu_w > \Delta \nu_i$ when all customers are aware of advance selling.*

This is a stricter condition to guarantee that the general option strategy dominates the newsvendor. In the general case it is difficult to have the closed form of the optimal option for the normal distribution. After obtaining the optimum p^* from (4.47), we need to compare the expected profit of the option strategy to

4.5 Binomial Model with Informed and Uninformed Customers

the newsvendor directly.

$$\mu_i x(p^*) + (p^* - c_m) \mu_i \bar{F}(p^*) - (p_s - c_m) \mu_i \bar{F}(p_s) \geq (p^* + c_a) \phi(z_a) E_{M_i}[\hat{\sigma}_i] - (p_s + c_s) \phi(z) \tilde{\sigma}_i, \quad (4.50)$$

Which is equivalent to

$$(p^* + c_a) \phi(z_a) \leq [\mu_i x(p^*) + (p^* - c_m) \mu_i \bar{F}(p^*) - (p_s - c_m) \mu_i \bar{F}(p_s) + (p_s + c_s) \phi(z) \tilde{\sigma}_i] / E_{M_i}[\hat{\sigma}_i]. \quad (4.51)$$

The right side of (4.50) is the difference between costs while the left side is the difference between the expected revenue collected from customers of the two schemes. If the right term is nonpositive, then of course option selling is better. If the right term is positive, the option strategy dominates the newsvendor when the revenue difference is greater than the cost difference.

4.5.2 Not All Customers Are Informed

In this section we discuss the situation where there exist informed and uninformed segments simultaneously. Informed customers arrive in period 1 and compare their expected consumer surplus of advance buying with that of spot buying. At the end of period 1 the manufacturer observes the number of customers of advance buying. In period 2 uninformed customers arrive in the spot market. The valuations about the product are realized for both advance buying and uninformed customers. The former decide whether to exercise the option while the latter choose to buy or not. Suppose that the manufacturer always fulfill the advance customers' order first, which implies $p + c_a \geq p_s + c_s$. Then the manufacturer's profit can be formulated

4.5 Binomial Model with Informed and Uninformed Customers

as follows:

$$\Pi_u = \max_{x,p} E_{M_i} \left[\hat{\pi}(x, p, \hat{M}_i, \hat{M}_u) \right], \quad (4.52)$$

$$\text{s.t.} \quad \int_p^{+\infty} (v - p) f(v) dv - x \geq (1 - \eta) \int_{p_s}^{+\infty} (v - p_s) f(v) dv, \quad (4.53)$$

$$x, p \geq 0, \quad (4.54)$$

$\hat{M}_i$ and $\hat{M}_u$ are conditional on the realized $M_i = m_i$ in period 1. We suppose the market size $\hat{M}_i$ and $\hat{M}_u$ are large enough such that they can be approximated by the normal distributions $N(m_i \bar{F}(p), m_i \bar{F}(p) F(p))$ and $N(\mu_u \bar{F}(p_s), \mu_u \bar{F}(p_s) F(p_s) + \sigma_u^2 \bar{F}^2(p_s))$. Similar to [Prasad et al. \(2011\)](#), the manufacturer actually can update the demand information at the end of period 1 because M_i has been realized as m_i and the two segments are correlated with correlation coefficient ρ . Given m_i , after the update the market size of the uninformed segment follows a normal distribution with

$$\hat{\mu}_u = (\mu_u + \rho(m_i - \mu_i) \frac{\sigma_u}{\sigma_i}) \bar{F}(p_s), \quad (4.55)$$

$$\hat{\sigma}_u = [(\mu_u + \rho(m_i - \mu_i) \frac{\sigma_u}{\sigma_i}) \bar{F}(p_s) F(p_s) + \sigma_u^2 (1 - \rho^2) \bar{F}^2(p_s)]^{1/2} \quad (4.56)$$

$E_{M_i}[\hat{\mu}_u] = \mu_u \bar{F}(p_s)$ and again according to Jensen's inequality $E_{M_i}[\hat{\sigma}_u] \leq [\mu_u \bar{F}(p_s) F(p_s) + \sigma_u^2 (1 - \rho^2) \bar{F}^2(p_s)]^{1/2}$. Let

$$\hat{\mu} \equiv \hat{\mu}_i + \hat{\mu}_u, \quad (4.57)$$

$$\hat{\sigma} \equiv (\hat{\sigma}_i^2 + \hat{\sigma}_u^2)^{1/2}. \quad (4.58)$$

4.5 Binomial Model with Informed and Uninformed Customers

Consequently the manufacturer faces a newsvendor problem with two demand classes and one class has higher priority over the other one. The manufacturer's profit is

$$\hat{\pi}(x, p, \hat{M}_i, \hat{M}_u) = m_i x + \max_q E_{\hat{M}_i, \hat{M}_u} [p \min(q, \hat{M}_i) - c_a(\hat{M}_i - q)^+ \quad (4.59)$$

$$\begin{aligned} &+ p_s \min((q - \hat{M}_i)^+, \hat{M}_u) - c_s(\hat{M}_u - (q - \hat{M}_i)^+)^+ - c_m q] \\ &= m_i x + (p - c_m)\hat{\mu}_i + (p_s - c_m)\hat{\mu}_u - \min_{q \geq 0} L(q) \end{aligned} \quad (4.60)$$

where $L(q) = E_{\hat{M}_i, \hat{M}_u} [(p + c_a - c_m)(\hat{M}_i - q)^+ + (p_s + c_s - c_m)(\hat{M}_u - (q - \hat{M}_i)^+)^+ + c_m(q - \hat{M}_i - \hat{M}_u)^+]$ and $\hat{M}_i, \hat{M}_u$ are conditional on $M_i = m_i$. The last step is because $q = \hat{M}_i + \hat{M}_u - (\hat{M}_i - q)^+ - (\hat{M}_u - (q - \hat{M}_i)^+)^+ - (q - \hat{M}_i - \hat{M}_u)^+$

[Sen and Zhang \(1999\)](#) study a newsvendor problem with two demand classes similar to our case. Let $\underline{q} \equiv \hat{\mu}_i + \hat{\mu}_u + \Phi^{-1}(1 - \frac{c_m}{p_s + c_s})\hat{\sigma}$ and $\bar{q} \equiv \hat{\mu}_i + \hat{\mu}_u + \Phi^{-1}(1 - \frac{c_m}{p + c_a})\hat{\sigma}$.

Proposition 4.5.6. *The optimal order quantity for (4.60) satisfies $\hat{\mu}_i + \hat{\mu}_u + \Phi^{-1}(1 - \frac{c_m}{p_s + c_s})\hat{\sigma} \leq q^* \leq \hat{\mu}_i + \hat{\mu}_u + \Phi^{-1}(1 - \frac{c_m}{p + c_a})\hat{\sigma}$ conditional on $M_i = m_i$.*

Proof $L(q)$ can be transformed to $E_{\hat{M}_i, \hat{M}_u} [(p + c_a - p_s - c_s)(\hat{M}_i - q)^+ + (p_s + c_s - c_m)(\hat{M}_u + \hat{M}_i - q)^+ + c_m(q - \hat{M}_i - \hat{M}_u)^+]$. Take the derivative of $L(q)$ w.r.t. q , $L'(q) = (p + c_a - p_s - c_s)Pr(\hat{M}_i > q) + (p_s + c_s - c_m)Pr(\hat{M}_u + \hat{M}_i > q) - c_m Pr(q > \hat{M}_i + \hat{M}_u)$. Notice that $(p_s + c_s - c_m)Pr(\hat{M}_u + \hat{M}_i > \underline{q}) - c_m Pr(\underline{q} > \hat{M}_i + \hat{M}_u) = 0$ and $(p + c_a - p_s - c_s)Pr(\hat{M}_i > q) \geq 0$. Hence it is necessary to have $q \geq \underline{q}$ to make $L'(q) = 0$.

Similarly $L(q) = E_{\hat{M}_i, \hat{M}_u} [-(p + c_a - p_s - c_s)(\hat{M}_u - (q - \hat{M}_i)^+)^+ + (p + c_a - c_m)(\hat{M}_u + \hat{M}_i - q)^+ + c_m(q - \hat{M}_i - \hat{M}_u)^+]$. Again $L'(q) = -(p + c_a - p_s - c_s)Pr(\hat{M}_u >$

4.5 Binomial Model with Informed and Uninformed Customers

$(q - \hat{M}_i)^+ + (p + c_a - c_m)Pr(\hat{M}_u + \hat{M}_i > q) - c_m Pr(q > \hat{M}_i + \hat{M}_u)$. Then $L'(\bar{q}) \leq 0$. Only when $q \leq \bar{q}$ it is possible to make $L'(q) = 0$. $\square$

Proposition (4.5.6) gives the upper and lower bounds for the optimal order quantity. We can see that $p + c_a - p_s - c_s$ determines the width of the interval between the bounds which can be very loose if $p + c_a \gg p_s + c_s$. Based on the proof of Proposition (4.5.6) we find :

Proposition 4.5.7. *Having observed $M_i = m_i$, the optimal order quantity q^* can be obtained from the following condition:*

$$(p + c_a - p_s - c_s)Pr(\hat{M}_i > q) + (p_s + c_s)Pr(\hat{M}_u + \hat{M}_i > q) = c_m. \quad (4.61)$$

Proof From the proof of Proposition (4.5.6) $L'(q) = -(p + c_a - p_s - c_s)Pr(\hat{M}_i > q) - (p_s + c_s - c_m)Pr(\hat{M}_u + \hat{M}_i > q) + c_m Pr(q > \hat{M}_i + \hat{M}_u)$. Then $L''(q) = (p + c_a - p_s - c_s)\hat{g}_i(q) + (p_s + c_s - c_m)\hat{g}(q) + c_m\hat{g}(q) \geq 0$, where $\hat{g}_i$ and $\hat{g}$ are the density functions for $\hat{M}_i$ and $\hat{M}_i + \hat{M}_u$, respectively. So $L(q)$ is convex and consequently the optimal order quantity satisfies $L'(q) = 0$. For $Pr(q > \hat{M}_i + \hat{M}_u) = 1 - Pr(\hat{M}_u + \hat{M}_i > q)$, we have $(p + c_a - p_s - c_s)Pr(\hat{M}_i > q) + (p_s + c_s)Pr(\hat{M}_u + \hat{M}_i > q) = c_m$. $\square$

Note the left side term of (4.61) is the marginal revenue of adding one unit of product and the right side is the marginal cost. Only when $q < \hat{M}_i$ the marginal revenue of $p + c_a - p_s - c_s$ aggregates otherwise $p_s + c_s$ is in effect.

4.5.2.1 When Lost Sale Cost is Formidable

Normally after the demand of informed customers is realized and observed in period 1, the manufacturer will order q^* products which is derived from a standard

4.5 Binomial Model with Informed and Uninformed Customers

newsvendor problem as in (4.60). We call this the ordering to the newsvendor policy.

However, the manufacturer will choose to order up to at least m_i when the lost sale cost for advance buying customers is unaffordable. We still call it the ordering to upper bound policy which is same as in the situation where all of the customers are informed. Although it stems from the situation of $c_a \rightarrow +\infty$, we extend it to general c_a . Under this policy (4.60) becomes $\hat{\pi}(x, p, \hat{M}_i, \hat{M}_u) = m_i x + (p - c_m)\hat{\mu}_i + (p_s - c_m)\hat{\mu}_u - \min_{q \geq m_i} L(q)$ and

$$L(q) = E_{\hat{M}_i, \hat{M}_u} [(p_s + c_s - c_m)(\hat{M}_u + \hat{M}_i - q)^+ + c_m(q - \hat{M}_i - \hat{M}_u)^+]. \quad (4.62)$$

Notice it requires $q \geq m_i$ and the term $p + c_a - p_s - c_s$ has disappeared. The optimal order quantity now is a piecewise function as follows

$$q^* = \begin{cases} m_i, & \text{if } \underline{q} < m_i \\ \underline{q}, & \text{if } \underline{q} \geq m_i. \end{cases} \quad (4.63)$$

Given $M_i = m_i$ under the ordering to the newsvendor policy the optimal order quantity is always greater than $\underline{q}$. On one hand when $\underline{q} \geq m_i$ the term $(\hat{M}_i - q)^+$ equals 0 so that the optimal order quantity actually is $\underline{q}$. In such cases, the ordering to the newsvendor policy is equivalent to the ordering to upper bound. On the other hand, when $\underline{q} < m_i$ the optimal order quantity is also less than m_i (It can be shown by contradiction that $\underline{q} < m_i < q^*$ is impossible). In these cases, the order quantity is less than in the ordering to upper bound. Therefore

4.5 Binomial Model with Informed and Uninformed Customers

the overall profit by applying the ordering to the newsvendor is greater than the ordering to upper bound for a specific (x, p) . Since advance selling is a special case of the option strategies associated with the ordering to upper bound policy, it is dominated by the general option strategy with the ordering to the newsvendor policy. We have a proposition similar to Proposition (4.5.3)

Proposition 4.5.8. *Given lost sale cost c_a , the lower bound of the manufacturer's profit by using any option (x, p) that satisfies condition (4.53) and the ordering to the newsvendor is $\mu_i x + (p - c_m)E[\hat{\mu}_i] + (p_s - c_m)E_{M_i}[\hat{\mu}_u] - E_{M_i}[L(q^*)|M_i = m_i]$, where $E[\hat{\mu}_i] = \mu_i \bar{F}(p)$, $E[\hat{\mu}_u] = \mu_u \bar{F}(p_s)$ and $L(q^*)$ is defined in (4.62) and (4.63). Furthermore the optimal profit of option selling is no less than that of advance selling.*

This proposition suggests that the manufacturer should always seek the optimal option which is not necessarily to be $(x(0), 0)$.

Now another question arises. Is advance selling optimal under the ordering to upper bound policy? We have shown that when all customers are informed $(x(0), 0)$ is optimal among all options under the ordering to upper bound policy. The following proposition deals with this question. Let z_{m_i} denote $\frac{m_i - \hat{\mu}_i - \hat{\mu}_u}{\hat{\sigma}}$ and z_{M_i} follows by substituting M_i for m_i . Define an indicator function using z_{m_i} as follows

$$\delta(m_i) = \begin{cases} 1, & \text{if } z \geq z_{m_i}, \\ 0, & \text{otherwise.} \end{cases} \quad (4.64)$$

Under the ordering to upper bound policy $\min_{q \geq m_i} L(q) = (p_s + c_s)\phi(z)\hat{\sigma}\delta(m_i) + \hat{\sigma}[c_m z_{m_i} + (p_s + c_s)(\phi(z_{m_i}) - (1 - \Phi(z_{m_i}))z_{m_i}](1 - \delta(m_i))$. Let k_{m_i} denote $[c_m z_{m_i} +$

4.5 Binomial Model with Informed and Uninformed Customers

$(p_s + c_s)(\phi(z_{m_i}) - (1 - \Phi(z_{m_i}))z_{m_i})(1 - \delta(m_i))$. So the expected profit will be

$$\begin{aligned} \Pi_u &= \mu_i x + (p - c_m)\mu_i \bar{F}(p) + (p_s - c_m)\mu_u \bar{F}(P_s) - \int_0^{+\infty} \min_{q \geq m_i} L(q) g_i(m_i) dm_i \\ &\geq \mu_i x + (p - c_m)\mu_i \bar{F}(p) + (p_s - c_m)\mu_u \bar{F}(P_s) \\ &\quad - (p_s + c_s)\phi(z) \left[(E[\hat{\sigma}_i^2] + E[\hat{\sigma}_u^2]) Pr(z \geq z_{M_i}) \right]^{1/2} - [E[k_{m_i}^2 \hat{\sigma}^2] Pr(z < z_{M_i})]^{1/2}, \end{aligned} \quad (4.65)$$

where the last step is an application of Cauchy-Schwarz inequality.

Although we have Proposition (4.5.2) when all customers are informed of option selling, it is not the case when not all customers are informed. In the former case because the manufacturer's production cost is fixed given m_i and the revenue is decreasing in p , the manufacturer has to choose $p = 0$ to maximize her revenue and in turn the profit. On the contrary, in the latter case the two segments share common stocks, though informed customers enjoy higher priority. Any stock left after the distribution to informed customers can still be used to fulfill uninformed customers. As a consequence the risk pooling effect reduces the total cost in comparison with two separate segments even without taking into account the demand information update for the uninformed segment.

Proposition 4.5.9. *Given c_a , advance selling with price $x(0)$ is optimal among all options under the ordering to upper bound policy, if $Pr(z \geq z_{M_i}) = O(\epsilon)$.*

Proof Suppose we observe m_i in period 1, the manufacturer's profit can be written as follows $m_i x + p \hat{\mu}_i - c_m m_i + \max_{q \geq 0} E[p_s \min(\hat{M}_u, q + m_i - \hat{M}_i) - c_s(\hat{M}_u - q - m_i + \hat{M}_i)^+ - c_m q]$. Here q represents the extra order quantity exceeding m_i . For $\min(\hat{M}_u, q + m_i - \hat{M}_i) = \hat{M}_u - (\hat{M}_u - q - m_i + \hat{M}_i)^+$, we write the profit as

4.5 Binomial Model with Informed and Uninformed Customers

$m_i x + p\hat{\mu}_i - c_m m_i + p_s \hat{\mu}_u - \min_{q \geq 0} E[(p_s + c_s)(\hat{M}_u - q - m_i + \hat{M}_i)^+ + c_m q]$. It is helpful to think of $\hat{M}_u + \hat{M}_i - m_i$ as a new random variable. $\min_{q \geq 0} E[(p_s + c_s)(\hat{M}_u - q - m_i + \hat{M}_i)^+ + c_m q]$ is a newsvendor problem with zero salvage value.

Then the expected profit is $\mu_i x + p\mu_i \bar{F}(p) - c_m \mu_i + p_s \mu_u \bar{F}(p_s) - E_{M_i}[\min_{q \geq 0} E[(p_s + c_s)(\hat{M}_u - q - m_i + \hat{M}_i)^+ + c_m q]]$. Because $Pr(z \geq z_{M_i}) = O(\epsilon)$, the last term is negligible. And the profit will be $\mu_i x + p\mu_i \bar{F}(p) - c_m m_i + p_s \hat{\mu}_u - O(\epsilon)$. As revenue $\mu_i x + p\mu_i \bar{F}(p) - c_m \mu_i \leq \mu_i(x(0) - c_m)$. Therefore option $(x(0), 0)$ achieves the maximal expected profit and it is exactly the advance selling policy. $\square$

A situation can be constructed to satisfy the condition in Proposition (4.5.9) using the uniform, truncated normal, or normal distributions. For example, M_u and M_i are random variables ranged mostly within $[a, b]$ and $[c, d]$, respectively. That is to say $Pr(M_u > b, M_u < a) = O(\epsilon)$ and $Pr(M_i > d, M_i < c) = O(\epsilon)$. Meanwhile $c \gg b$ and $\sigma_i \gg \sigma_u$. These assumptions guarantee that we still can make $z_{m_i} > z$ even when we observe $M_i = c$ in period 1.

Again we can construct a situation by setting $Pr(M_u < a) = \epsilon$ and $Pr(M_i > d) = \epsilon$ as well as $a \gg d$ such that $Pr(z < z_{M_i}) = O(\epsilon)$. Later we show that $(x(0), 0)$ is not always optimal under the ordering to upper bound policy.

4.5.2.2 The Expedite Model

Under the expedite setting, the manufacturer can fulfill any shortage by purchasing from an open market or expediting the production at a price of p_e per unit. For $p + c_a = p_s + c_s = p_e$, the term $(p + c_a - p_s - c_s)P(\hat{M}_i > q)$ in (4.61) no longer

4.5 Binomial Model with Informed and Uninformed Customers

exists and (4.60) becomes a standard newsvendor problem. As a result,

$$q^* = \hat{\mu}_i + \hat{\mu}_u + \Phi^{-1}\left(1 - \frac{c_m}{p_e}\right)\hat{\sigma}, \quad (4.66)$$

$$L(q) = (p_s + c_s)\phi(z_a)\hat{\sigma}, \quad (4.67)$$

$$\hat{\pi}(x, p, \hat{M}_i, \hat{M}_u) = m_i x + (p - c_m)\hat{\mu}_i + (p_s - c_m)\hat{\mu}_u - L(q). \quad (4.68)$$

where $z_a = \Phi\left(1 - \frac{c_m}{p_e}\right)$. The manufacturer's optimal profit can be expressed as follows,

$$\Pi_u = \max_{x,p} \mu_i x + (p - c_m)\mu_i \bar{F}(p) + (p_s - c_m)\mu_u \bar{F}(p_s) - p_e \phi(z_a) E_{M_1}[\hat{\sigma}], \quad (4.69)$$

where $E_{M_1}[\hat{\sigma}] \leq [\mu_i \bar{F}(p)F(p) + \mu_u \bar{F}(p_s)F(p_s) + \sigma_u^2(1 - \rho^2)\bar{F}^2(p_s)]^{1/2}$ and (x, p) satisfies condition (4.53), $x, p \geq 0$.

4.5.2.3 Comparing Option Selling to The Newsvendor

Under the newsvendor scheme, the manufacturer's expected profit is as follows,

$$\Pi_N = (p_s - c_m)(\mu_i + \mu_u)\bar{F}(p_s) - (p_s + c_s)\phi(z)\tilde{\sigma},$$

where $\tilde{\sigma} \equiv [(\mu_i + \mu_u)\bar{F}(p_s)F(p_s) + (\sigma_i^2 + \sigma_u^2 + 2\sigma_u\sigma_i\rho)\bar{F}^2(p_s)]^{1/2}$ is the standard deviation of demand in period 2.

The comparison between the expected profits of adopting advance selling and the newsvendor is presented as follows, $\mu_i(x(0) - c_m) + (p_s - c_m)\mu_u \bar{F}(p_s) - (p_s + c_s)\phi(z)E[\hat{\sigma}_u] \geq (p_s - c_m)(\mu_i + \mu_u)\bar{F}(p_s) - (p_s + c_s)\phi(z)\tilde{\sigma}$. After some cancellations

4.5 Binomial Model with Informed and Uninformed Customers

and transpositions we then have $x(0) \geq p_s \bar{F}(p_s) + c_m F(p_s) - (p_s + c_s) \phi(z) (\tilde{\sigma} - E[\hat{\sigma}_u]) / \mu_i$

Proposition 4.5.10. *Advance selling dominates the newsvendor scheme when $\nu - \nu_w \geq \Delta\nu$, $\Delta\nu = \bar{F}(p_s) + c_m F(p_s) - (p_s + c_s) \phi(z) (\tilde{\sigma} - E[\hat{\sigma}_u]) / \mu_i$.*

For the expedite setting, clearly $(x(p_s), p_s)$ is a feasible solution to (4.69). The corresponding $E_{M_1}[\hat{\sigma}] \leq [(\mu_i + \mu_u) \bar{F}(p_s) F(p_s) + \sigma_u^2 (1 - \rho^2) \bar{F}^2(p_s)]^{1/2} \leq [(\mu_i + \mu_u) \bar{F}(p_s) F(p_s) + (\sigma_i^2 + \sigma_u^2 + 2\sigma_u \sigma_i \rho) \bar{F}^2(p_s)]^{1/2} = \tilde{\sigma}$. Therefore,

$$\Pi_u \geq \mu_i x(p_s) + (p_s - c_m)(\mu_i + \mu_u) \bar{F}(p_s) - (p_s + c_s) \phi(z) E_{M_1}[\hat{\sigma}] \quad (4.70)$$

$$> (p_s - c_m)(\mu_i + \mu_u) \bar{F}(p_s) - (p_s + c_s) \phi(z) \tilde{\sigma} \quad (4.71)$$

$$= \Pi_N \quad (4.72)$$

Proposition 4.5.11. *The option strategy under the expedite setting always dominates the newsvendor scheme.*

4.5.2.4 Is Informing All of The Customers Optimal?

Think of a situation in which all customers of a market are informed of option selling. As the whole market size can be observed in period 1, demand uncertainty in period 2 has been reduced to the most extent. As a result, the probability of lost sale has been minimized. The unit lost sale cost has increased from c_s to c_a for the customers coming from the initial uninformed segment. In addition, making effort to inform all customers incurs more cost, like advertise expense. The tradeoff involving uncertainty reduction and increase in the unit lost sale and advertising costs affects the decision whether to make uninformed customers

4.5 Binomial Model with Informed and Uninformed Customers

aware of option selling.

We have shown the consequence if $Pr(z \geq z_{M_i}) = O(\epsilon)$ in (4.65) in Section (4.5.2.1). If $Pr(z < z_{M_i}) = O(\epsilon)$, the expected profit $\Pi_u \geq \mu_i x + (p - c_m)\mu_i \bar{F}(p) + (p_s - c_m)\mu_u \bar{F}(P_s) - (p_s + c_s)\phi(z) [(E[\hat{\sigma}_i]^2 + E[\hat{\sigma}_u]^2)]^{1/2}$. For option $(x(p_s), p_s)$ clearly $(E[\hat{\sigma}_i]^2 + E[\hat{\sigma}_u]^2)^{1/2} < \tilde{\sigma}$, where $\tilde{\sigma}$ is the standard deviation of demand in period 2 under the newsvendor scheme. Thus in such situation option $(x(p_s), p_s)$ achieves better profit than the newsvendor scheme. We have shown in Section (4.5.2.1) that manipulating the size of the two segments could construct a situation in which $Pr(z < z_{M_i}) = O(\epsilon)$ and consequently option selling outperforms the newsvendor. Most importantly, this process does not involve the cost structure.

However, there could exist a scenario of c_m, p_s, c_s, c_a such that when both of the two segments are informed option selling obtains less expected profit than the newsvendor, i.e. condition (4.51) not satisfied.

Proposition 4.5.12. *There exists a market scenario with $Pr(z < z_{M_i}) = O(\epsilon)$ in which option selling performs better than the newsvendor. And there exists a specific cost structure such that informing the initial uninformed customers would worsen the manufacturer's expected profit.*

It implies that the manufacturer's expected profit would not be improved if she manages to inform all customers of option selling. That is to say, the manufacturer should keep part of her customers uninformed. The ratio of the market size of the informed segment to that of the uninformed segment depends on the cost structure.

Furthermore, "a little bit" option selling is always beneficial. We can show as long as the market size of the informed segment is small enough, the condi-

tion $Pr(z < z_{M_i}) = O(\epsilon)$ can be met so that the expected profit improves over the newsvendor. The reasons are twofold, firstly, the manufacturer gains more profit from the informed segment by passing the risk to the uninformed segment; secondly the realization of the size of the informed segment can bring in more accurate demand information about the size of the uninformed segment via the correlation coefficient ρ .

With a market scenario in which option selling outperforms the newsvendor, we can also create a combination of p_s, c_s, c_m such that the condition in Proposition 4.5.10 unfulfilled. As a result the optimal option under the ordering to upper bound policy is not necessarily to be $(x(0), 0)$.

4.6 Conclusions

We study real option selling in the context where a newsvendor manufacturer faces a joint decision making of option pricing and production planning. Consumers are uncertain about their valuations on the product and the uncertainty will be resolved in the spot market. Taking advantage of the uncertainty, the newsvendor manufacturer offers real options to induce customers to buy before the spot market. The demand uncertainty about the spot market then is reduced or eliminated after the manufacturer observes the market sizes at the end of the option-selling season.

We show that option selling as an instrument of revenue management and information updating can achieve better profit for the manufacturer both in the single-option and single-market case and two-option and two-market case under

the fluid model. We provide the optimal option prices and order quantities in both cases. We also compare the different policies the manufacturer can employ and identify the conditions for the optimal policy when there are two market segments. Moreover, the optimal option price shows that advance selling is inferior to option selling.

We also investigate the demand information updating problem by modelling a market with correlated informed and uninformed segments with binomial demand in the spot market. We incorporate advance selling into this model and introduce an expedite mode. We show option selling under the expedite mode always dominates the newsvendor. And option selling dominates advance selling in any case. But interestingly advance selling could be optimal in a special case where the manufacturer adopts an ordering up to upper bound policy or forced compliance. Finally a surprising result is that it is not always beneficial to let all customers know option selling due to the cost structure.

5

Final Remarks and Future Research

This thesis has studied the strategies of process flexibility and real options to hedge against demand uncertainty. We have shown both strategies can improve the performances of manufacturing systems, though in different terms.

In Chapter 2 we have studied the VHC method for balanced and nonidentical systems and then extended it to unbalanced systems. In both setting our numerical studies show VHC can achieve 90% benefits of full flexibility without using more links than the long chain. This makes future improvement in performance very difficult for such systems. A natural extension of VHC is to study the unbalanced and asymmetrical systems (numbers of products and plants are not equal). We have provided a version of VHC for such case in the remark of this chapter. However, the general performance of VHC in this setting has not been tested. The difficulty of asymmetric systems lies in that one plant could be shared by multiple products. It is not trivial to recognize how much capacity of a plant is left to participate in the process flexibility. Consequently, it is difficult to identify the

bottleneck in a chain. Besides, the long chain efficiency bounds in this chapter are numerically verified but it seems very difficult to prove them analytically, given the fact that so far only the efficiency of identical systems increasing in system size has been proved in the literature.

In Chapter 3 we have studied process flexibility in multistage setting. We show the existence of a sparse structure that achieves $1 - \epsilon$ benefits of full flexibility. We also introduce a CPCMM model to cope with the evaluation on process flexibility structures with unknown distribution. Better design of process flexibility structures even for small number of stages, say 3 stages, is a challenging problem. To simplify the problem, our model ignores the interstage logistic routes. If taking into account the logistic routes, it is a daunting task even for a system with two stages and a small system size. Moreover, the multistage process flexibility involves supply chain coordination. If only an isolated stage adopts process flexibility the final output still will be limited by other stages. In other words, the supply chain strategy for each stage should be aligned with each other. So it is interesting to consider the decentralized supply chain with process flexibility. A natural extension of the results in this chapter would be to use just the mean and variance information as in [Natarajan et al. \(2009\)](#) who show that the worst case expected maximum flow is solvable as a second order cone program (SOCP). Since the dual of the maximum output problem in process flexibility setting is a 0-1 problem, we can apply this result to formulating the process flexibility design problem as a mixed integer SOCP. In addition, since both CPCMM by [Natarajan et al. \(2011\)](#) and SOCP by [Natarajan et al. \(2009\)](#) can be applied to formulating the process flexibility design problem, a comparison between them is of interest.

In Chapter 4 we have investigated the joint option pricing and production planning problem. We incorporate advance selling into our option selling framework. And we show our option selling scheme dominates the newsvendor under the fluid model and the binomial model with the expedite setting. One direct extension of this work could be the n market segments case. It is interesting to study whether there exist a bundle of options which can discriminate the multiple segments. Another possible future research is production cost could be a function of the order quantity. It is interesting to know how the optimal option price will change with the production cost. Moreover, in our essay we assume customers are risk neutral. Naturally risk averse customers can be studied in a similar way. Notice we have not studied the market with correlated informed and uninformed segments under the fluid model. Because this case can eventually be transformed to a problem comprised of the option selling and the newsvendor and bridged by the correlation coefficient, we did not extend our research to this setting. In contrast, adding multiple markets with different valuation distributions to the binomial model with correlated informed and uninformed customers would be challenging. We also notice that [Tang et al. \(2004\)](#), [McCardle et al. \(2004\)](#) and [Weng and Parlar \(1999\)](#) do not use uncertain valuation to model the advance-selling problem. Instead they introduce the likelihood to purchase or a fraction function to determine the number of customers to buy in advance. [Weng and Parlar \(1999\)](#) argue that using likelihood to purchase is “*consistent with the choice behavior of a large majority of customers*”. We are especially interested in the circumstance where customers have heterogeneous likelihoods to buy. We conjecture that modelling option selling in this way will lead to results similar to

those in this essay. We are going to explore this topic following this essay.

References

- O. Z. Aksin and F. Karaesmen. Characterizing the performance of process flexibility structures. *Oper. Res. Lett.*, **35**(4):477–484, 2007. [18](#)
- J. M. Amend. GM manufacturing operations more flexible after bankruptcy. Retrieved March 12, 2012 from www.wardsauto.com, 2010. [10](#)
- Anonymous. Alibaba: The world’s greatest bazaarr. Retrieved June 24, 2013 from www.economist.com, 2013. [2](#)
- D. Barnes-Schuster, Y. Bassok, and R. Anupindi. Coordination and flexibility in supply contracts with options. *Manufacturing Service Oper. Management*, **4**(3):171–207, 2002. [100](#), [101](#)
- A. Bassamboo, R. S. Randhawa, and J. A. Van Mieghem. Optimal flexibility configurations in newsvendor networks: Going beyond chaining and pairing. *Management Sci.*, **56**(8):1285–1303, 2010. [18](#)
- R. Berry, M. Honig, T. Nguyen, V. Subramanian, H. Zhou, and R. Vohra. Newsvendor model of capacity sharing. Working paper, 2012. [62](#)

REFERENCES

- D. Bertsimas, K. Natarajan, and C. P. Teo. Probabilistic combinatorial optimization: Moments, semidefinite programming and asymptotic bounds. *SIAM J. Optim.*, **15**(1):185–209, 2004. [91](#)
- D. Bertsimas, K. Natarajan, and C. P. Teo. Persistence in discrete optimization under data uncertainty. *Math. Programming, Ser. B, Issue Optim. Under Uncertainty*, **108**(2–3):251–274, 2006. [91](#)
- E. K. Bish, A. Muriel, and S. Biller. Managing flexible capacity in a make-to-order environment. *Management Sci.*, **51**(2):167–180, 2005. [18](#)
- D. L. Bryan and M. E. O’Kelly. Hub-and-spoke networks in air transportation: An analytical review. *Journal of Regional Science*, **39**(2):275–295, 1999. [16](#)
- T. A. Burgin. The gamma distribution and inventory control. *Opl Res. Q.*, **26**(3(i)):507–525, 1975. [62](#)
- A. Burnetas and P. Ritchken. Option pricing with downward-sloping demand curves: The case of supply chain options. *Management Sci.*, **51**(4):566–580, 2005. [102](#)
- G. P. Cachon and M. A. Lariviere. Contracting to assure supply: How to share demand forecasts in a supply chain. *Management.Sci.*, **47**(5):629–646, 2001. [100](#), [141](#)
- G. Calafiore and M. C. Campi. Uncertainty convex programs: Randomized solutions and confidence levels. *Math.Programming*, **102**(1):25–46, 2005. [80](#)

REFERENCES

- F. Chen and M. Parlar. Value of a put option to the risk-averse newsvendor. *IIE Transactions*, **39(5)**:481–500, 2007. [102](#)
- X. Chen and Z.J. Shen. An analysis of a supply chain with options contracts and service requirements. *IIE Transactions*, **44(10)**:805–819, 2012. [102](#)
- F. Cheng, M. Ettli, G. Y. Lin, M. Schwarz, and D. D. Yao. Flexible supply contracts via options. *Working paper*, 2003. IBM T. J. Watson Research Center, Yorktown Heights, NY. [100](#)
- S. H. Cho and C. S. Tang. Advance selling in a supply chain under uncertain supply and demand. *Manufacturing Service Oper. Management*, **15(2)**:305–319, 2013. [104](#)
- M. C. Chou, C. P. Teo, and H. Zheng. Process flexibility: Design, evaluation and applications. *Flexible Services and Manufacturing Journal*, **20(1-2)**:59–94, 2008. [21](#)
- M. C. Chou, G. A. Chua, C. P. Teo, and H. Zheng. Design for process flexibility: Efficiency of the long chain and sparse structure. *Oper. Res.*, **58(1)**:43–58, 2010. [10](#), [11](#), [12](#), [18](#), [20](#), [21](#), [24](#), [28](#), [47](#), [68](#), [78](#), [81](#), [92](#), [93](#)
- M. C. Chou, G. A. Chua, C. P. Teo, and H. Zheng. Process flexibility revisited: the graph expander and its applications. *Oper. Res.*, **59(5)**:1090–1105, 2011. [12](#), [18](#), [20](#), [21](#), [56](#)
- G. N. Cook and J. Goodwin. Airline networks: A comparison of hub-and-spoke

REFERENCES

- and point-to-point systems. *The Journal of Aviation/ Aerospace Education & Research*, **17(2)**:Art. 1, 2008. [16](#)
- A. Şen and A. X. Zhang. Newsboy problem with multiple demand classes. *IIE Transactions*, **31(5)**:431–444, 1999. [147](#)
- T. Deng and Z. J. Shen. Process flexibility design in unbalanced networks. *Manufacturing Service Oper. Management*, **15(1)**:24–32, 2013. [20](#)
- C. H. Fine and R. M. Freund. Optimal investment in product-flexible manufacturing capacity. *Management Sci.*, **36(4)**:449–466, 1990. [4](#), [5](#)
- Corporate Ford. Transformed Michigan assembly plant sets new standard for flexible manufacturing at ford. Retrieved March 12, 2012 from www.corporate.ford.com, 2010. [10](#)
- Media Ford. New era of flexible manufacturing begins at Fords Kentucky Truck Plant in Louisville. Retrieved March 12, 2012 from www.media.ford.com, 2009. [10](#)
- Q. Fu, C. Y. Lee, and C. P. Teo. Procurement management using option contracts: random spot price and the portfolio effect. *IIE Transactions*, **42(11)**:793–811, 2010. [102](#)
- G. Gallego. Production management. Lecture Notes, Department of Industrial Engineering and Operations Research, Columbia University, 1995. [62](#)
- G. Gallego and Ö. Şahin. Revenue management with partially refundable fares. *Oper. Res.*, **58(4)**:817–833, 2010. [6](#), [98](#), [102](#)

REFERENCES

- P. Golikeri. Hospital chains take hub-and-spoke route to hinterland. Retrieved March 20, 2014 from www.dnaindia.com, 2011. 16
- G. Golovachkina and J. R. Bradley. Supplier-manufacturer relationships under forced compliance contracts. *Working paper*, 2002. Cornell University, Ithaca, NY. 102
- M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html. 92
- M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>, March 2014. 92
- S. C. Graves and B. T. Tomlin. Process flexibility in supply chains. *Management Sci.*, **49(7)**:907–919, 2003. 18, 21, 68, 69, 72, 82
- S. Gurumurthi and S. Benjaafar. Modeling and analysis of flexible queueing systems. *Naval Res. Logist.*, **51(5)**:755–782, 2004. 18
- W. J. Hopp, E. Tekin, and M. P. Van Oyen. Benefits of skill chaining in serial production lines with cross-trained workers. *Management Sci.*, **50(1)**:83–98, 2004. 10, 18
- W. J. Hopp, S. M. R. Iravani, and W. L. Xu. Vertical flexibility in supply chains. *Management Sci.*, **56(3)**:495–502, 2010. 18

REFERENCES

- S. M. R. Iravani, M. P. Van Oyen, and K.T. Sims. Structural flexibility: A new perspective on the design of manufacturing and service operations. *Management Sci.*, **41(2)**:151–166, 2005. [12](#), [20](#), [21](#)
- S. M. R. Iravani, B. Kolfal, and M. P. Van Oyen. Call-center labor cross-training: It’s a small world after all. *Management Sci.*, **53(7)**:1102–1112, 2007. [18](#), [20](#), [21](#)
- W. C. Jordan and S. C. Graves. An analytic approach for demonstrating the benefits of limited flexibility. *Working paper*, 1991. Sloan School of Management, Massachusetts Institute of Technology. [68](#), [76](#)
- W. C. Jordan and S. C. Graves. Principles on the benefits of manufacturing process flexibility. *Management Sci.*, **41(4)**:577–594, 1995. [4](#), [5](#), [10](#), [12](#), [18](#), [21](#), [24](#), [69](#)
- P. R. Kleindorfer and D. J. Wu. Integrating long-term and short-term contracting via business-to-business exchanges for capitalintensive industries. *Management Sci.*, **49(11)**:1597–1615, 2003. [102](#)
- J. G. Klinecicz. Hub location in backbone/tributary network design: a review. *Location Science*, **6(1–4)**:307–335, 1998. [16](#)
- V. Martínez-de-Albéniz and D. Simchi-Levi. Competition in the supply option market. *Oper. Res.*, **57(5)**:1082–1097, 2009. [102](#)
- K. McCardle, K. Rajaram, and C. S. Tang. Advance booking discount programs under retail competition. *Management Sci.*, **50(5)**:701–708, 2004. [103](#), [160](#)

REFERENCES

- A. Muriel, A. Somasundaram, and Y. Zhang. Impact of partial manufacturing flexibility on production variability. *Manufacturing Service Oper. Management*, **8(2)**:192–205, 2006. [18](#)
- K. Natarajan, M. Song, and C. P. Teo. Persistency model and its applications in choice modeling. *Management Sci.*, **55(3)**:453–469, 2009. [159](#)
- K. Natarajan, C. P. Teo, and Z. Zheng. Mixed 0-1 linear programs under objective uncertainty: a completely positive representation. *Oper. Res.*, **59(3)**:713–728, 2011. [69](#), [85](#), [86](#), [87](#), [88](#), [89](#), [90](#), [91](#), [159](#)
- A. Prasad, K. E. Steckel, and X. Zhao. Advance selling by a newsvendor retailer. *Production and Oper. Management*, **20(1)**:129–142, 2011. [6](#), [7](#), [103](#), [112](#), [136](#), [137](#), [144](#), [146](#)
- E. Rasmusen. *Games and Information: An Introduction to Game Theory*. Blackwell, Auflage, third edition, 2001. [127](#)
- P. H. Ritchken and C. S. Tapiero. Contingent claims contracting for purchasing decisions in inventory management. *Oper. Res.*, **34(6)**:864–970, 1986. [100](#)
- A. K. Sethi and S. P. Sethi. Flexibility in manufacturing: A survey. *Internat. J. Flexible Manufacturing Systems*, **2**:289–328, 1990. [4](#)
- D. Simchi-Levi and Y. Wei. Understanding the performance of the long chain and sparse designs in process flexibility. *Oper. Res.*, **60(5)**:1125–1141, 2012a. [11](#), [19](#)

REFERENCES

- D. Simchi-Levi and Y. Wei. Worst-case analysis of process flexibility designs. *Working Paper*, 2012b. MIT, Massachusetts, US. [20](#), [21](#)
- S. B. Tan and C. Leong. The evolution of public transport policies in singapore. Case study, Lee Kuan Yew School of Public Policy, National University of Singapore, 2013. [16](#)
- C. S. Tang, K. Rajaram, A. Alptekinoglu, and J. Ou. The benefits of advance booking discount programs: Model and analysis. *Management Sci.*, **50(4)**: 465–478, 2004. [103](#), [160](#)
- G. Volpato and A. Stocchetti. Managing product life cycle in the auto industry:evaluating carmakers effectiveness. *Int. J. Automotive Technology and Management*, **8(1)**:22–41, 2008. [2](#)
- R. B. Wallace and W. Whitt. A staffing algorithm for call centers with skill-based routing. *Manufacturing Service Oper. Management*, **7(4)**:276–294, 2005. [18](#)
- X. Wang and J. Zhang. Process flexibility: A distribution-free bound on the performance of k-chain. *Working paper*, 2013. New York University, New York, US. [19](#), [21](#)
- Z. K. Weng and M. Parlar. Integrating early sales with production decisions: Analysis and insights. *IIE Transactions*, **31(11)**:1051–1060, 1999. [103](#), [160](#)
- Chris Woodyard. Ford focuses on flexibility at its factories. Retrived March 16, 2012 from www.usatoday.com, 2011. [10](#)

REFERENCES

- D. J. Wu and P. R. Kleindorfer. Competitive options, supply contracting, and electronic markets. *Management Sci.*, **207**:668–675, 2005. [102](#)
- D. J. Wu, P. R. Kleindorfer, and J. E. Zhang. Optimal bidding and contracting strategies for capital-intensive goods. *Eur. J. Oper. Res.*, **137(3)**:657–676, 2002. [101](#)
- J. Xie and S. Shugan. Electronic tickets, smart cards, and online prepayments: When and how to advance sell. *Marketing Sci.*, **20(3)**:219–243, 2001. [103](#)
- M. Yu, R. Kapuscinski, and H. S. Ahn. Advance selling with interdependent customer valuations. *Working paper*, 2011. Hong Kong University of Science and Technology, Hong Kong. [103](#)
- Y. Zhao, S. Wang, T. C. E. Cheng, X. Yang, and Z. Huang. Coordination of supply chains by option contracts: A cooperative game theory approach. *Eur. J. Oper. Res.*, **137(3)**:657–676, 2010. [100](#)
- H. Zheng. *Sparse flexibility structures: Design and application*. PhD thesis, National University of Singapore, Singapore, 2007. [69](#)