SUPPLY CHAIN DESIGN:
CAPACITY, FLEXIBILITY AND WHOLESALE PRICE STRATEGIES

by

Brian T. Tomlin

B.E. Chemical Engineering

University College Dublin, Ireland, 1992

Submitted to the

Alfred P. Sloan School of Management

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Management

at the

Massachusetts Institute of Technology

February 2000

Signature of Author____

Alfred P. Sloan School of Management
November 15, 1999

Certified by____

Stephen C. Graves
Abraham J. Siegel Professor of Management
Co-director, Leaders for Manufacturing Program

Accepted by____

Birger Wernerfelt
Professor of Management Science
Chair, Doctoral Program

This page intentionally left blank

SUPPLY CHAIN DESIGN:
CAPACITY, FLEXIBILITY AND WHOLESALE PRICE STRATEGIES

by

Brian T. Tomlin

Submitted to the Alfred P. Sloan School of Management
on November 15, 1999 in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Management

Abstract

Increasing recognition is being placed, both in industry and in academia, on effective supply chain management. The term supply chain management presupposes that there exists a supply chain to be managed. With a focus on supply chains in which demand uncertainty is the key challenge, this dissertation develops strategies and models to aid in the design of certain supply chain features, namely capacity, flexibility and wholesale price schedules.

Firstly, this dissertation studies capacity investments in single-product supply chains in which the participants make investments to maximize their individual expected profits. Using a stylized game theoretic model of a supply chain comprising a supplier and a manufacturer, simple non-linear wholesale price schedules, whether they be quantity premium or quantity discount schedules, are shown to outperform simple linear schedules in terms of the total supply chain profit achieved. While the model is stylized, it provides insight into how actual wholesale price schedules can be structured to induce near optimal supply chain capacity investments.

Next, this dissertation then extends the work of Jordan and Graves (1995) so as to develop process flexibility strategies for multiple-product multiple-stage supply chains. The ability of multiple-stage supply chains to fill product demands is shown to be affected by two inefficiencies, termed stage-spanning bottlenecks and floating bottlenecks, that do not affect single-stage supply chains. Flexibility configurations differ in the protection they provide against these inefficiencies. The chaining strategy of Jordan and Graves (1995), with augmentation if either the number of stages or number of products is large, is shown to provide a high degree of protection and therefore to enable multiple-stage supply chains to better meet demand.

Finally, this dissertation studies the capacity decision in multiple-product multiple-stage supply chains. Solution approaches to the capacity investment problem in which there is either an expected shortfall bound or a service level bound are developed. The service level problem, while widely studied in inventory theory, has not been studied in the multiple-product multiple-stage supply chain capacity literature to date. In addition to developing solution approaches, insights into the optimal capacity decisions in multiple-product multiple-stage supply chains are provided.

Thesis Supervisor: Stephen C. Graves
Title: Professor of Management Science

This page intentionally left blank

Acknowledgements

The past four years at MIT have been an immensely rewarding educational and intellectual experience. Much of the credit for this must go to my adviser Professor Stephen C. Graves. Interacting with Steve has developed my analytical skills more than I could have imagined. Observing Steve, both in the class room and in meetings, has taught me that clarity of expression is rooted in a deep and thorough understanding of the subject matter at hand. This is an insight I will always remember.

I am extremely grateful for Steve's contributions to my research. This thesis is immeasurably better as a result of his comments. Much credit also goes to Professors Georgia Perakis and Yashan Wang, the other members of my committee. Their feedback has been invaluable. I would also like to thank Professor Andreas Schulz for his helpful comments.

My time at MIT has been greatly enriched by all my fellow Operations Management and System Dynamics doctoral students. Hari Abhyankar, Sharon Novak and Sean Willems deserve special mention. The four of us spent considerable time and energy pursuing our doctorates. I'm glad to say that we have all succeeded, in no small part because of the friendship that we developed over the years. I look forward to hearing of the graduations of Amit, Ariel, Juan Carlos, Laura, Paulo G., Paulo O. and René.

I would especially like to thank my great friends Jeffrey Furman and Martin Haugh for the afternoon breaks in Au Bon Pain and the enjoyable evenings in Grafton Street. I can only hope that in the years to come we will continue our conversations, even if distance makes them less frequent.

I would also like to thank both the Montgomery family and the O'Shea family for their frequent hospitality.

I owe a tremendous debt of gratitude to my parents Breffni and Nollaig and to my sisters Aoife and Dervla. The frequent weekend conversations were always eagerly anticipated, especially the conversations with my nieces Maeve, Ciara and Fiona. I look forward to talking with my nephew Kevin when he's a little older. The encouragement I received, not only from my immediate family, but also from my brothers-in-law Bryan and Daniel, was very valuable.

Finally I want to thank Sarah Montgomery for all her support over the past few years. Sarah has put up with the long hours that I have spent in my office with incredible patience and good will. I am eternally grateful to her. I can only hope to repay her in the coming years.

This page intentionally left blank

To my parents, Breffni and Nollaig Tomlin

This page intentionally left blank

Table of Contents

1 Introduction

1.1 Problem Context

The past decade has seen an increasing recognition of the importance of supply chain management. In industry, the rapid growth of supply chain software companies testifies to the significance that businesses place on the efficient management of their supply chains. Research in this area has become a key focus of the operations management academic community in recent years; a comprehensive review of this literature can be found in Tayur, Ganeshan and Magazine (1998).

Numerous definitions of a supply chain exist (see for example, Lee and Billington, 1993), and while they may differ in terminology, they are reasonably consistent in meaning. A supply chain can be thought of as a network of entities interacting to transform raw material into finished product for customers. Each entity provides some activity necessary for this transformation. Interactions can take the form of material, information or monetary flow. Defining the boundaries of any particular supply chain is somewhat arbitrary, as what constitutes finished product for one supply chain may be raw material for another. While supply chains typically refer to manufacturing systems, as is reflected by the terminology used, supply chains also arise in service systems.

The term "supply chain management" presupposes that there exists a supply chain to be managed. How does this supply chain come into being? The answer is most likely a combination of legacy, happenstance and design; where design implies that a conscious decision concerning some supply chain feature is made with regard to the overall supply chain performance. This dissertation aims to develop strategies and models to aid in the design of certain supply chain features.

Supply chain design encompasses a very large number of decisions. Product development, in which product functions and features are determined, dictates certain supply chain features. The set of processing technologies chosen to deliver product functionality specifies some of the necessary supply chain activities. The entities that carry out the activities need to be selected; and may be internal or external to the firm. Willems (1999) studies the tradeoff between cost and lead time in the entity selection problem. Rules and contracts governing entity interaction need to be specified. Willems (1999) provides a brief review of the research on the material quality aspects of supplier-manufacturer interactions. Tsay, Nahmias and Agrawal (1998) reviews the recent literature on supply chain contracts. Decisions need to be made on whether the supply chain will be a make-to-stock or make-to-order system, or some hybrid of the two. An entity providing a supply chain activity needs to determine how this activity will be delivered. Will it use a single resource (i.e. plant or machine) or multiple resources? What resource capacity is needed? This list of supply chain decisions is by no means exhaustive but it serves to highlight the complex nature of supply chain design. Supply chain design is often further complicated by decentralized decision-making. Extra complexity arises if the supply chain processes multiple products – aspects such as resource flexibility must be considered.

Given the large number of interdependent supply chain decisions and the scarcity of existing design strategies and tools, it is not altogether surprising that supply chains can evolve on a somewhat ad hoc basis. It is unlikely that an all-encompassing supply chain design tool is feasible. Indeed it is very doubtful whether such a tool would even be useful; different supply chains face different challenges, and as such, features critical to one supply chain may be unimportant to another.

I would argue in favor of a hierarchical approach to supply chain design. Firstly the critical challenges facing supply chains should be categorized. For each category, the key features that determine the ability of the supply chain to meet the challenge should be identified. By identifying the critical challenge and key features, the supply chain design problem becomes more manageable, allowing design strategies and tools to be developed.

One possible categorization of supply chain challenges is cost, time and uncertainty. In some supply chains, the unit product cost may be the overriding competitive challenge. In others, time may be critical; whether it be final product lead time, new product introduction time, or production ramp time. In yet other supply chains, uncertainty may be the most important challenge. Lee and Billington (1993) identify three sources of uncertainty, “demand (volume and mix), process (yield, machine downtimes, transportation reliabilities), and supply (part quality, delivery reliability)”. Of course, cost, time and uncertainty are important in different degrees to all supply chains. However, if one aspect dominates, this helps to simplify the complex challenge of developing appropriate strategies or tools for supply chain design.

This dissertation focuses on the design of supply chains in which demand uncertainty is the key challenge.

1.2 Strategies for Dealing with Demand Uncertainty

Inventory, lead-time and capacity are common strategies used in supply chains facing uncertain or variable demand. Each provides a buffering mechanism to absorb the uncertainty.

In some supply chains, products must be produced before demand is realized. This typically arises in situations where production takes a long time and the selling period is short, fashion goods for instance. As demand is not known when production occurs, firms must determine how much inventory to build. A large inventory enables the firm to fulfill demand with a high probability. In other supply chains, production and sales occur over multiple periods. A supply chain may have a nominal capacity that enables it to produce a certain quantity each period; that is the supply chain has an installed capacity that can not be altered in the time frame needed to meet demand. If no inventory exists in the system, lost sales or backorders occur in periods in which demand exceeds capacity. By carrying inventory, a supply chain can meet some of the excess demand. Inventory can be used to smooth the demand process. Businesses incur significant costs for holding inventory, and as such, it is important that the quantity and placement of inventory be judiciously chosen. A strategic inventory placement tool is detailed in Graves and Willems (1998). The selection of entities to perform the supply chain activities can affect the inventory requirement through the lead time of the various activities. Willems (1999) develops a tool for this selection problem.

Lead time, defined as the actual time between order placement and delivery, can also be used to cope with demand uncertainty. If the lead time between a supplier and manufacturer is very small, then extra orders can be placed with the supplier in periods of high demand. Customer lead time can also play a role. Finished goods service time, defined as the time allowed between customer order placement and actual delivery, has a large impact on the supply chain's ability to deal with demand uncertainty. If the service time is small, and the demand exceeds the supply chain capacity, then a shortfall will occur if there is insufficient inventory. As the service time increases, the supply chain will be better able to absorb demand uncertainty. Customer orders in periods of high demand may take longer to fill, due to capacity constraints, but as long as this increase does not cause the customer lead time to exceed the finished goods service time, demand can still be fulfilled. Of course, increasing finished goods service time may come at the cost of a decrease in customer satisfaction.

The nominal capacity chosen is another mechanism that firms can use to cope with demand uncertainty. Demand can be fulfilled as long as the capacity is not exceeded. The larger the capacity, the less likely it is that some demand can not be met. Supply chain capacity depends on the capacity of the entities. Each entity performs a distinct activity and this activity may be provided by one or more resources, where a resource may be a plant, a machine or even a person. The capacity of each entity depends on the capacity of its resources. Capacity is costly and therefore one should be mindful of the tradeoff between cost and customer service when the capacities of the various supply chain entities are chosen.

In multiple-product supply chains, that is supply chains that process more than one finished good, it is not only the capacity of the entities, but the nature of the capacity, that plays a role in coping with demand uncertainty. Resources may be dedicated to one product or may be capable of performing the activity for multiple products – that is a resource may be flexible. Flexibility enables supply chains to cope with demand uncertainty because the allocation of resource capacity to products can occur after demand is realized.

Inventory, lead time and the level and flexibility of capacity are all features of a supply chain that can be used to deal with demand uncertainty. This dissertation develops strategies and models for the supply chain capacity decision. The dissertation contains three main chapters, with each chapter focusing on a different aspect of the supply chain capacity problem.

Chapter 2 studies the capacity level problem in single-product supply chains with multiple agents. That is, the individual supply chain stages make their own capacity choices taking into account the demand uncertainty, purchase price and sales price. Independent agent problems are of interest because agents acting in their own interest typically make supply chain sub-optimal decisions. Cachon and Lariviere (1997, 1999) study this problem for a supply chain comprising a manufacturer and supplier. Only the supplier need invest in capacity. Chapter 2 analyzes supply chains in which both the supplier and the manufacturer must invest in capacity. Moreover, it introduces a new wholesale price schedule contract not studied by Cachon and Lariviere (1997, 1999). This contract is a quantity premium contract whereby the manufacturer pays a higher average unit cost as the order size increases. Such a contract may seem counter intuitive as buyers often expect to be rewarded for placing larger orders. Quantity discounts are prevalent both in practice and in the operations management literature. However, it is shown that a quantity premium contract induces the supplier to invest in more capacity which in turn benefits the manufacturer. A correctly priced quantity premium contract results in optimal supply chain capacity decisions and at the same time enables the manufacturer to capture the total supply chain profit.

Chapter 3 develops flexibility strategies for multiple-product supply chains. The flexibility decision is one of determining what products a resource should be able to process. Jordan and Graves (1995) study this problem for a single stage supply chain and determine that a "chaining" strategy is effective. This single-stage work is extended to multiple-stage supply chains in Chapter 3. Supply chains with multiple stages are shown to suffer from inefficiencies that do not affect single stage supply chains. Flexibility configurations differ in the protection they provide against these inefficiencies. A chaining strategy similar to that of Jordan and Graves (1995) is shown to provide effective protection against these inefficiencies. The following flexibility strategy is developed. For multiple-stage systems, the single-stage guidelines of Jordan and Graves (1995) should be followed to create a chain structure for each of the supply chain stages. In supply chains with a large number of products or stages, additional flexibility is advisable, especially for stages in which the capacity is not much greater than the expected demand. This extra layer of flexibility should again be added in accordance with the guidelines of Jordan and Graves (1995) to create another chain structure overlaying the initial chain structure.

Chapter 4 focuses on the capacity level decision in multiple-product multiple-stage supply chains in which there is a central decision-maker. The existing literature has focused primarily on single-product or single-stage supply chains. In multiple-product supply chains, floating bottlenecks can occur and thus it is necessary to consider the multiple stages of the supply chain. For an expected profit criteria, Eberly and Van Mieghem (1997) develop the optimal policy structure, or capacity levels, for multiple-product multiple-stage supply chains in which each stage is totally flexible. However the authors do not develop any algorithms for determining the actual capacity levels of the stages. Studying a special case of the model developed in Eberly and Van Mieghem (1997), Harrison and Van Mieghem (1999) prove that the optimal capacity investment involves some hedging that would never be optimal if the demands were known. In other words, there may be no demand scenario for which all stages operate at capacity. The authors provide an algorithm to determine the optimal capacity levels. However this algorithm is not appropriate unless the number of stages is very low. Indeed, the purpose of the paper is really to provide insight into the notion that multiple-stage capacity planning under uncertainty involves hedging rather than to provide an efficient solution method. Chapter 4 studies the capacity decision in multiple-product multiple-stage supply chains and develops problem formulations that can be solved using the Microsoft Excel Solver. In addition to developing a solution approach to the problem in which there is an expected shortfall bound, I develop an approach to the problem in which there is a service level bound. The service level problem, while widely studied in inventory theory, has not been studied in the multiple-product multiple-stage supply chain capacity literature to date. In addition to providing solution approaches, Chapter 4 provides some insights into the optimal capacity decisions in multiple-product multiple-stage supply chains.

2 Capacity Decisions in Supply Chains with Independent Agents

2.1 Introduction

Often firms must invest in production capacity before a product is bought to market. At the time of investment, product demand is uncertain, and the capacity decision might be made using a forecast of the product sales, where the forecast may take the form of a probability distribution. This uncertainty in the product demand complicates the capacity investment decision and any reduction in the uncertainty would be desirable. In some circumstances, the company may have the ability to make capacity investments over a number of periods. Sales in prior periods may contain valuable information that can then be used to update and refine the forecast for future period sales. Firms do not always have the option of building capacity over the life of the product. Short life cycle products may require the total capacity to be built before any sales are made. The focus of this dissertation (Chapter 2) is on such one-time capacity investment decisions.

More often than not, production of a product requires more than one operation. Multiple-component products are a prevalent example. In such multiple-component products, the capacity level for the components and the final assembly must be determined. If all components are specific to the particular end product, then an equal capacity should be chosen for all components and assembly, where capacities are expressed in units of the end product. While I use multiple component products as the motivating example, this research covers the more general class of products whose production requires multiple operations.

Multiple component products can be produced by a completely integrated firm or by a supply chain of independent firms where the production of certain components is outsourced to suppliers. The manufacturer, the firm selling the end product to the market, may produce some of the components or may simply assemble the supplier components into the end product. In the case of an integrated firm, the capacity investment decision is made by a single agent; while in the case of a supply chain, the manufacturer and component suppliers make their capacity investment decisions as independent agents. The supply chain capacity is determined by the operation with the least capacity.

In the integrated firm, the manufacturer has control over all the component capacity decisions. When the manufacturer outsources some of the component production, this is no longer the case. A supplier may choose to invest in less capacity than the manufacturer would like. Likewise the manufacturer might invest in less capacity than a supplier would like. This arises due to externalities in the supply chain. Positive (negative) externalities occur when the action of one agent benefits (harms) another. As agents act in their own interest, they are likely to engage in too little (too much) of actions that cause positive (negative) externalities. In this supply chain case, the capacity investment by one firm affects the profit of another firm as the total end product sales, and thus component sales, are limited by the overall supply chain capacity.

Supply chains with multiple agents are characterized by decentralized decision making while an integrated firm is typically characterized by a central or coordinated decision-maker. It should be noted that while decentralized decision-making is more likely to occur in multiple firm supply chains and coordinated decision-making more likely in an integrated firm, the key underlying structures are that of a coordinated decision-maker or decentralized decision makers acting in their own interest. Both of these decision-making structures may arise in either integrated firms or in supply chains.

Supply chains with multiple agents, or decision-makers, have begun to receive a lot of attention in the operations management literature. This interest arises primarily because of the fact that independent agents acting in their own self-interest often make decisions that lead to supply chain sub-optimal performance. By sub-optimality, I mean that the total supply chain profit, Π_d, is strictly less than that achievable by a single central decision-maker with complete information, Π_c. Therefore decentralized control is said to be inefficient if $\Pi_d<\Pi_c$. As noted by Cachon (1998), supply chain inefficiency arises due to supply chain externalities. Much of the existing research focuses on how this gap can be decreased by implementing different classes of contracts. Tsay, Nahmias, and Agarwal (1998) refer to this as the system-wide performance improvement objective. I use the more commonly used term of channel coordination.

Supply chains studied in this literature are typically simple in structure. While some papers have studied a serial supply chain inventory problem with multiple agents, e.g. Cachon and Zipkin (1997), or a single supplier with multiple retailers, e.g. Cachon (1997), the vast majority of the literature focuses on a single un-capacitated upstream party supplying a single downstream party.

An overview of the supply chain model common in the literature is given in the review paper of Tsay, Nahmias, and Agarwal (1998).

> Consider a supply chain structure in which an upstream party (which we refer to as a manufacturer) provides a single product to a downstream party (which we refer to as a

> retailer), who in turn serves market demand. This scenario could describe the link between any two consecutive nodes in a supply chain, and indeed on occasion the tandem may be referred to as supplier and manufacturer, manufacturer and distributor, or, most generally, supplier and buyer. Researchers commonly make the following simplifying assumptions to render the analysis more tractable. The manufacturer produces (or acquires) the product at a common unit cost of c and charges a retailer a wholesale/transfer price of $W(Q)$ for Q units. $W(Q)$ may be exogenous, or a decision variable under the control of one or more parties. The retailer in turn sells the product at a price p per unit. Market demand, denoted as $D(p)$, in reality is both price-sensitive and uncertain. While some models include both these features, it is more common to either take the retail price as fixed and represent market demand as a random variable (as in the operations research literature), or assume a deterministic, downward-sloping demand curve (as in the economics and marketing literatures). In the latter case the retailer's decision is primarily p, whereas in the former it is Q. A simpler underlying structure allows traditional inventory models to treat more complex problem settings including multiple periods, continuous review, and finite and infinite horizons. However, most contract papers assume only a one-period problem (i.e., a newsvendor setting), since the resulting models are often too complex to be amenable to multi-period analysis.

So the basic paradigm is that of a downstream party, facing a newsvendor problem, who buys goods from an un-capacitated upstream party. Pasternack (1985) showed that the double marginalization effect, first noted by Spengler (1950) for deterministic downward-sloping demand, also occurs when retail price is fixed but demand uncertain. The downstream party will purchase less than the supply chain optimal quantity unless the upstream party prices its goods at marginal cost. Pasternack showed that a buy-back policy whereby the upstream party promises to buy back unsold goods for a fixed price (lower than the wholesale price) induces the downstream party to purchase more. A properly priced buy-back policy achieves complete channel coordination. If the upstream party has control over the wholesale and buy-back prices, it can essentially capture the total expected supply chain profit.

This supply chain problem has received a lot of attention. Lariviere (1998) analyzes this problem in more detail, investigating the upstream party's wholesale optimal price when no buy-backs are allowed (i.e. a price only contract). The author develops conditions on the demand distribution for the upstream party's optimization problem to be well behaved. The framework used when allowing the upstream party to set the wholesale price is that of a Stackleberg game in which the upstream party is the leader. In other words, the upstream party sets the wholesale price and then the downstream party orders its optimal quantity given this price. Because all the information is assumed to be known to both agents, the upstream party knows with certainty what quantity the downstream party will order. Note that this optimal quantity depends on the demand distribution.

The downstream party need not necessarily buy the goods before the demand is realized. However if the downstream party must commit to an order quantity before demand is known, he faces the same problem as buying before demand. With this in mind, alternative contract structures from buy-back policies have also been investigated. Minimum quantity contracts force the downstream party to commit to a minimum quantity that can be added to when demand is realized. Quantity flexible contracts allow the downstream party to adjust its order upward or downward but at a cost. Such contracts have been shown to increase channel coordination. These two type of contracts have also been studied in multiple period problems where the retailer commits to a total purchase quantity over the whole horizon and then places orders in each period.

While the common framework tends to be that of inventory, it is closely related to capacity decisions. The minimum quantity and quantity flexible contracts can be directly mapped into capacity problems where the downstream party must reserve capacity at an upstream party.

The papers by Cachon and Lariviere (1997, 1999) deserve special mention, as they are the most closely related papers to the capacity model studied here. In the 1997 paper, a manufacturer faces a single period stochastic demand that is either high (with probability β) or low (with probability 1-β). Retail price, *r*, is fixed. The only capacity constraint that the manufacturer faces is that of a single component supplier. The supplier can build capacity at a cost of *c* per unit. Unused capacity can be salvaged by the supplier at *v* per unit. All other costs are normalized to zero. The manufacturer acts as the Stackleberg leader and sets the wholesale price, *w*, paid to the supplier per unit of component ordered. Components are bought after demand has been realized. Given the price, the supplier determines its capacity investment, which is either high or low. β is assumed to be restricted to a range in which a central decision-maker would invest in high capacity ($\beta_I \leq \beta \leq 1$). The authors investigate three contracts, a price only contract, a termination fee contract and a specific type of minimum order contract. A price only contract is one in which the only contract parameter is the constant wholesale price per unit to be paid to the supplier. A termination fee contract requires the manufacturer to specify a wholesale price, an initial order quantity and a per unit cancellation fee for any of these orders that are not actually wanted. The minimum purchase quantity is closely related and shown to be equivalent for this model. The authors raise an important distinction between capacity and inventory, namely that capacity investments are difficult to verify and thus contracts based on capacity can be prohibitively difficult to enforce, making contracts analogous to inventory buy-back contracts impossible. The authors identify two problem regimes: (i) contract compliance can be enforced and the manufacturer offers a contract that includes a wholesale price and minimum order quantity and (ii) compliance is voluntary and the manufacturer cannot force the supplier to invest in sufficient capacity but can only offer inducements based on the contract. In the voluntary compliance regime, the authors show that if both parties have complete information, the manufacturer never offers the termination fee or minimum purchase quantity contracts. They also prove that a price only contract can lead to supply chain sub-optimality for a certain range of β. Termination fee and minimum purchase quantity contracts are shown to be useful when the supplier has less information on the demand uncertainty – these contracts can act as signals of the manufacturer's demand type.

Cachon and Larivere (1999) updates the 1997 paper in two key aspects. The termination and minimum order quantity are mapped into a more general class of contracts termed advance contracts. Advance contracts specify an initial commitment quantity, m, from the manufacturer, that are purchased at a price w_m per unit, an optional quantity, o, that costs w_o per unit. The manufacturers actual order after demand is realized must lie between m and $m+o$. Optional units carry a per unit exercise price of w_e, but do not have to be purchased. Assuming that the supplier's capacity choice is a concave function of the wholesale price in a price only contract, the authors allow the demand distribution to be continuous. The authors develop a sufficient, but not necessary condition, for concavity, namely that the distribution be an increasing failure rate (IFR) distribution.

A constant wholesale price schedule, whereby the retailer pays a constant price per unit to the manufacturer is common in the existing multiple decision-maker supply chain literature. However, more complex pricing schedules exist. A quantity discount schedule is one in which the average price per unit decreases as the retailer's order increases. The two most common discount schedules are incremental quantity discounts and all unit quantity discounts. An upstream party may offer a quantity discount schedule to induce the downstream party to order in larger quantities, which can in turn reduce the frequency with which the supplier incurs the fixed cost of production. Such quantity discounts, when demand is deterministic and operating costs include inventory holding and set up costs, have been studied by Monahan (1984), Lal and Staelin (1984), Lee and Rosenblatt (1984) and Banerjee (1984). The focus of these works is on the increased profit the upstream party can gain from offering a quantity discount.

Jeuland and Shugan (1983) demonstrate the channel coordination attributes of a specific quantity discount schedule for a certain supply chain model. In their model, an upstream party sells a product to a downstream party who in turn serves the final market. Demand is deterministic but price sensitive. Both parties incur a fixed cost. The decision to be made by both parties is the price to charge, the upstream party decides the wholesale price to charge the downstream party, and the downstream party decides the retail price to charge the market. The only constant wholesale price that would induce the retailer to price at the supply chain optimal retail price is the upstream party's marginal cost. However the upstream party does not price at marginal cost, as he must recoup the fixed cost. If both parties agree to share the total channel profits, then they choose the supply chain optimal prices. The authors show that there exists a specific quantity discount schedule that is equivalent to profit sharing and thus leads to channel coordination.

Weng (1995) integrates both of the above streams of quantity discount research to both analyze and determine the supplier's optimal all unit quantity discount schedule. He shows that quantity discounts alone are not sufficient to achieve complete channel coordination.

In a comment on the model of Jeuland and Shugan (1983), Moorthy (1987) notes that a quantity discount schedule is not unique in achieving channel coordination and develops the necessary and sufficient condition on the pricing schedule for complete channel coordination to be achieved in this model. A wide range of pricing schedule classes have instances that meet these conditions, including quantity discounts, quantity surcharges and two-part tariffs (which the author argues to be the best pricing class).

In Section 2.2, I introduce the specific supply chain capacity model studied in Chapter 2. The coordinated supply chain in which a central decision-maker optimizes the total profits is covered in Section 2.3.

Two specific games are analyzed in Section 2.4, one in which the wholesale price is exogenous, and the other in which it is under the control of the manufacturer. In both these games, the wholesale price per unit is constant for all order quantities. In the first game I show that the channel does not achieve complete coordination unless the exogenous wholesale price happens to be a certain unique price. In the second game, I show that the channel never achieves complete coordination.

A specific variable wholesale price schedule is introduced in Section 2.5. This variable schedule is a quantity premium (surcharge) schedule. It is shown that if the manufacturer chooses a specific incremental quantity premium, the total supply chain profits strictly increase; that is, there is a reduction in the supply chain inefficiency. However, the channel is not be completely coordinated. More interestingly, from the manufacturer's perspective, the manufacturer's expected profit strictly increases with such a quantity premium schedule. A continuous quantity premium schedule is shown to completely coordinate the channel and enables the manufacturer to capture the entire profit.

Section 2.6 studies the supply chain games when the supplier has control over the wholesale price schedule. The results are equivalent to the games in which the manufacturer controls the wholesale price schedule but instead of offering a quantity premium, the supplier offers a quantity discount.

Section 2.7 extends the constant wholesale price per unit games to supply chains in which the manufacturer has multiple suppliers. As in Section 2.4, there exists a unique set of wholesale prices for which the channel is completely coordinated. Again, when the manufacturer controls the wholesale prices, the channel is not completely coordinated.

Section 2.8 studies two-party supply chains in which only the supplier must invest in capacity. For discrete demand distributions with two (three) demand states, it is shown that a quantity premium wholesale price schedule with one (two) breakpoint(s) completely coordinates the channel.

Concluding remarks are presented in Section 2.9. Proofs of all lemmas can be found in Chapter 5 (Appendix 1).

2.2 The Supply Chain Model

The single end product supply chain consists of a manufacturer and a supplier facing a single period uncertain demand. The end product is sold at a fixed retail price of r per unit. Production of one unit of end product requires the use of both one unit of manufacturer capacity and one unit of supplier capacity. For exposition purposes, I assume that the product requires two components; the supplier produces one component and the manufacturer produces the other component and assembles the two components into the end product. However, any supply chain in which the end product requires the use of both supplier and manufacturer capacity fits into the model.

Both the manufacturer and the supplier must make their capacity investment before demand is known. The manufacturer (supplier) has a per unit capacity cost of c_M (c_S) and a salvage value on unused capacity of v_M (v_S). Salvage values are strictly less than capacity costs. The manufacturer (supplier) incurs a per unit marginal production cost of p_M (p_S). After demand becomes known, the manufacturer orders and pays for a quantity of the supplier's components, assembles these with its own component and then sells the end product. An order of Q units costs $W(Q)$.

All wholesale price schedules $W(Q)$ are assumed to meet the following two conditions. Firstly, $W(Q)$ is non-decreasing in Q. If this did not hold for all Q, i.e. $W(Q_2)<W(Q_1)$ for some $Q_2>Q_1$, then the manufacturer would prefer to order and pay for Q_2 units even if it only wanted Q_1. Such a price schedule is not reasonable. Secondly, $(r-p_M)Q-W(Q)$ is non-decreasing in Q. That is the manufacturer's profit is non-decreasing in the number of units sold. If this did not hold for all Q, i.e. if $(r-p_M)Q_2-W(Q_2)<(r-p_M)Q_1-W(Q_1)$ for some $Q_2 > Q_1$, then for a demand of Q_2 units the manufacturer would prefer to order and sell only Q_1 units as this would yield a larger profit. The supplier only receives $W(Q_1)$. By setting $W_{new}(Q_2)=(r-p_M)(Q_2-Q_1)+W(Q_1)>W(Q_1)$, then $(r-p_M)Q_2-W_{new}(Q_2)=(r-p_M)Q_1-W(Q_1)$ and the manufacturer is willing to buy and sell Q_2 units. The supplier's profit is larger than if the manufacturer only bought Q_1 units. Thus neither party has an incentive to implement a wholesale price schedule for which the second condition does not hold. These two conditions assure that if the demand is Q then the manufacturer orders and sells $\min\{y,Q\}$ units, where y is the supply chain capacity. In regions where the wholesale price schedule is continuous and differentiable, the conditions can be stated in terms of the marginal wholesale price. The marginal wholesale price must be non-negative and also less than or equal to the retail price less the manufacturer's marginal processing cost.

The retail price is assumed to be large enough so that a central decision-maker would invest in positive capacity i.e. $r>c_M+c_S+p_M+p_S$. It is also assumed that the only wholesale price schedules considered are such that both the supplier and manufacturer acting in their own interest would invest in positive capacity; for a constant wholesale price schedule, i.e. $W(Q)=wQ$, this implies $c_S+p_S<w<r-c_M-p_M$.

For constant wholesale price schedules, let m_M be the manufacturer's margin on each unit sold ($m_M=r-p_M-w$) and m_S be the supplier's margin on each unit sold ($m_S=w-p_S$).

The demand is assumed to have a continuous and differentiable cumulative distribution function $F_X(x)$ over the range $[a,b]$ where a is the minimum possible demand and b is the maximum possible demand.

Both supplier and manufacturer are assumed to be risk neutral, a strong but common assumption in the literature. This implies that both parties are willing to make a positive capacity investment if the expected profit is non-negative.

I also assume that all information is known to both parties when capacity investment decisions are made. That is both the supplier and manufacturer know each other's costs and the demand distribution. Think of the manufacturer sharing its demand forecast with the supplier. The information symmetry assumption is a strong one. However it is frequently made in the literature and can serve as an important first step in modeling a supply chain. Recently, the literature has begun to address the notion of information asymmetry, whereby one or more parties may have incomplete or imperfect information about either some of the other party's costs or the demand distribution. In the games presented in Chapter 2, all information is known to each party.

Contracts specifying the capacity choices of either party are assumed to be prohibitively expensive or unenforceable, that is the voluntary compliance regime of Cachon and Lariviere (1997, 1999) is assumed.

While there are a number of differences between this model and that of Cachon and Lariviere (1997, 1999), it is worth highlighting the three key differences. Firstly, both the manufacturer and supplier must make capacity investment in this model whereas the manufacturer is uncapacitated in the Cachon and Lariviere model. This introduces a new complexity and realism to the model. Secondly, the class of wholesale price contracts is different. I allow for a non-linear wholesale price schedule in which the manufacturer is not required to make any up front quantity commitment. In Cachon and Lariviere (1999), the price schedule is non-linear only if an up front quantity commitment is made. While this distinction may seem trivial, it has important implications. The authors show that an advance (non-linear) contract would never be offered if both manufacturer and supplier had the same information on the demand distribution; a constant wholesale price schedule would be offered by the manufacturer. In this dissertation, I show that when a non-linear price schedule with no up front commitment is allowed, the manufacturer offers a non-linear contract instead of the constant wholesale price contract. Moreover, not only does the manufacturer increase its profits with this contract but also the total supply chain profit increases. The third difference lies in the class of demand distributions allowed. In their 1997 paper, Cachon and Lariviere, assume a Bernoulli demand. In their 1999 paper, the authors allow IFR distributions as they show that IFR is a sufficient but not necessary condition for the supplier's capacity choice to be a concave function of the constant wholesale price *w*. They also show that the distribution $F_X(x)=1-x^{-k}$, while not IFR, would induce concavity. Simultaneously and independently of Cachon and Lariviere (1999), I have developed the sufficient and necessary condition for concavity. It should be noted that while the model presented here assumes a continuous distribution, a Bernoulli distribution could be readily handled.

It is also important to realize that the purpose of Cachon and Lariviere (1997, 1999) is different from the purpose of this dissertation. While Cachon and Lariviere cover the coordination loss issue, they also introduce the important notion of voluntary compliance and proceed to show the value of advanced contracts when the manufacturer has private information on the demand distribution that is not shared with the suppler. The purpose of this dissertation (Chapter 2) is to concentrate on the coordination issue and show that a quantity premium increases both the manufacturer's and supply chain's expected profit.

There is an analogous inventory model to this capacity model. The inventory model is one in which there are two critical components. One produced by the manufacturer and one produced by the supplier. Both components must be produced before demand is realized. After demand is realized, the manufacturer places an order, Q, with the supplier, who then processes Q units and delivers them to the manufacturer who then combines them with the other components to produce the finished units which are then sold.

Table 1 presents the data used for the numerical examples in Chapter 2. For both the manufacturer and supplier, the capacity and marginal costs can be either high or low. The sixteen examples cover all combinations of high and low costs (the high and low values are 8 and 3). Salvage values are taken to be 20% of the capacity costs. In each example the retail price is 35. The demand is normally distributed with a mean of 200 and a standard deviation of 80. Section 2.3 introduces expressions for the expected profit of the supply chain, the manufacturer and the supplier. These expressions are dependent on the demand distribution. For a normal distribution, a closed form version of the expressions can be derived in terms of the mean and standard deviation. A truncated Normal distribution would be more appropriate as demand should not be negative. However for analytical tractability, a non-truncated Normal distribution was assumed. The probability of a negative demand is less than 1% for the demand distribution used.

Table 1

	c_M	v_M	p_M	c_S	v_S	p_S	r	Mean Demand	St. Dev. Demand
Example 1	3	0.6	3	3	0.6	3	35	200	80
Example 2	3	0.6	3	3	0.6	8	35	200	80
Example 3	3	0.6	3	8	1.6	3	35	200	80
Example 4	3	0.6	3	8	1.6	8	35	200	80
Example 5	3	0.6	8	3	0.6	3	35	200	80
Example 6	3	0.6	8	3	0.6	8	35	200	80
Example 7	3	0.6	8	8	1.6	3	35	200	80
Example 8	3	0.6	8	8	1.6	8	35	200	80
Example 9	8	1.6	3	3	0.6	3	35	200	80
Example 10	8	1.6	3	3	0.6	8	35	200	80
Example 11	8	1.6	3	8	1.6	3	35	200	80
Example 12	8	1.6	3	8	1.6	8	35	200	80
Example 13	8	1.6	8	3	0.6	3	35	200	80
Example 14	8	1.6	8	3	0.6	8	35	200	80
Example 15	8	1.6	8	8	1.6	4	35	200	80
Example 16	8	1.6	8	8	1.6	8	35	200	80

2.2.1 Nomenclature

For ease of reference, the nomenclature used in this chapter is presented.

a: The minimum possible product demand

b: The maximum possible product demand

c_M: The per unit capacity cost of the manufacturer

c_S: The per unit capacity cost of the supplier

c_I: The per unit capacity cost of a central decision-maker = c_M+c_S

$F_X(x)$: Cumulative distribution function for the end product demand, x

$f_X(x)$: Probability density function for the end product demand, x

m_M: The per unit margin for the manufacturer = r-p_M-w

m_S: The per unit margin for the supplier = w-p_S

m_I: The per unit margin of a central decision-maker = r-p_M-p_S

N: The number of supplier is in the multiple supplier model (n=1,…,N)

$PD_M(Q)$: The price difference schedule for the manufacturer = rQ-$W(Q)$

$PD_S(Q)$: The price difference schedule for the supplier = $W(Q)$

Q: The size of an order the manufacturer places with the supplier.

Q_P The breakpoint in a single breakpoint incremental quantity premium schedule

Q_P^i The i^{th} breakpoint in a two breakpoint incremental quantity premium schedule, i=1,2

p_M: The per unit production cost of the manufacturer

p_S: The per unit production cost of the supplier

p_I: The per unit production cost of a central decision-maker = p_M+p_S

r: The per unit retail price at which the manufacturer sells the end product

v_M: The per unit salvage value of unused capacity for the manufacturer

v_S: The per unit salvage value of unused capacity for the supplier

v_I: The per unit salvage value of unused capacity for a central decision-maker = v_M+v_S

$W(Q)$: The wholesale price schedule (the manufacturer pays the supplier $W(Q)$ for Q units)

w: The wholesale price in a constant wholesale price schedule or the initial wholesale price in an incremental quantity premium price schedule

w_{crit}: The unique wholesale price in a constant wholesale price schedule for which the supply chain is completely coordinated

x: The demand for the end product

y_I: The coordinated supply chain capacity

y_I*: The optimal coordinated supply chain capacity

$y_M(w)$: The manufacturer's optimal capacity if the supplier has infinite capacity (for a constant wholesale price schedule w)

$y_S(w)$: The supplier's optimal capacity if the manufacturer has infinite capacity (for a constant wholesale price schedule w)

$y_M(w,\Delta)$: The manufacturer's optimal capacity if the supplier has infinite capacity (for a single breakpoint quantity premium wholesale price schedule (w,Δ))

$y_S(w,\Delta)$: The supplier's optimal capacity if the manufacturer has infinite capacity (for a single breakpoint quantity premium wholesale price schedule (w,Δ))

Δ: The quantity premium in a single breakpoint incremental quantity premium schedule

Δ_i: The i^{th} quantity premium in a two breakpoint incremental quantity premium schedule, i=1,2.

Δ_{crit}: The unique quantity premium in a single breakpoint quantity premium wholesale price schedule (w,Δ) for which the supply chain is completely coordinated

Π_C: The expected supply chain profit obtained by a central decision-maker

Π_D: The expected supply chain profit obtained by decentralized decision-making

$\Pi_M(W(Q))$: The expected supply chain profit of the manufacturer for a given wholesale price schedule $W(Q)$

$\Pi_S(W(Q))$: The expected supply chain profit of the supplier for a given wholesale price schedule $W(Q)$

2.3 The Coordinated Supply Chain

In a coordinated (or integrated) supply chain, a central decision-maker determines the capacity of both the supplier and manufacturer. The optimum manufacturer and supplier capacities are equal and therefore the central decision-maker problem can be cast as a single capacity investment decision with the following revenues and costs. The retail price is r, the marginal production cost is $p_I=p_M+p_S$, the unit margin is $m_I=m_M+m_S=r-p_M-p_S$, the capacity cost is $c_I=c_M+c_S$, the salvage value is $v_I=v_M+v_S$. The expected supply chain profit as a function of the capacity level, y_I, is given by,

$$\Pi_I(y_I) = -c_I y_I + m_I \int_a^{y_I} x f_X(x)dx + m_I y_I[1-F_X(y_I)] + v_I \int_a^{y_I} [y_I - x] f_X(x)dx \tag{1}$$

The first term is the capacity cost, the second and third terms combined are the expected total margin, and the fourth term is the expected salvage revenue. This is a newsvendor type of problem and the optimal capacity is given by critical fractile solution.

Lemma 1

The optimal capacity choice, y_I, for a central decision-maker is given by,*

$$y_I^* = F_X^{-1}\left(\frac{m_I - c_I}{m_I - v_I}\right) = F_X^{-1}\left(\frac{r - p_M - p_S - c_M - c_S}{r - p_M - p_S - v_M - v_S}\right)$$

It should be noted that the total supply chain profit depends only on the supply chain capacity and not at all on the wholesale price w, which simply transfers profits from one party to the other. Therefore if the supply chain capacity is different from the optimal capacity y_I*, then the total expected supply chain profit is strictly less than that obtained by a central decision-maker.

If either the supplier or the manufacturer is faced with making a capacity investment decision when the other party has infinite capacity, the capacity investment decision is again a single capacity choice and can be expressed similarly to the central decision-maker's problem, equation (1), but with the capacity, salvage and unit margins correctly adjusted. For a constant wholesale price schedule $W(Q)=wQ$, let $y_M(w)$ be the manufacturer's optimal capacity if the supplier has infinite capacity and let $y_S(w)$ be equivalently defined for the supplier. These capacity choices prove to be important in the games developed in later sections.

Lemma 2

For a given wholesale price, w, the optimal capacity choice, $y_M(w)$, for the manufacturer, assuming the supplier has infinite capacity, is given by,

$$y_M(w) = F_X^{-1}\left(\frac{m_M - c_M}{m_M - v_M}\right) = F_X^{-1}\left(\frac{r - w - p_M - c_M}{r - w - p_M - v_M}\right)$$

Lemma 3

For a given wholesale price, w, the optimal capacity choice, $y_S(w)$, *for the supplier, assuming the manufacturer has infinite capacity, is given by,*

$$y_S(w) = F_X^{-1}\left(\frac{m_S - c_S}{m_S - v_S}\right) = F_X^{-1}\left(\frac{w - p_S - c_S}{w - p_S - v_S}\right)$$

Lemma 4

(i) $y_M(w)$ *is strictly decreasing in w and* (ii) $y_S(w)$ *is strictly increasing in w*

(Note that $F_X(x)$ *is assumed to be continuous and differentiable)*

Lemma 5

(i) There is a unique wholesale price, w, such that $y_M(w)=y_S(w)$

(ii) This unique wholesale price, w_{crit}, *is given by,*

$$w_{crit} = \frac{(r - p_M)(c_S - v_S) + p_S(c_M - v_M) + c_M v_S - c_S v_M}{c_S - v_S + c_M - v_M}$$

(iii) At this wholesale price, w_{crit}, $y_M(w_{crit})=y_S(w_{crit})=y_I*$

Lemma 6

The necessary and sufficient condition on the demand distribution, $F_X(x)$, *for* $y_S(w)$ *to be a concave function of w is given by,*

$$\frac{f_X'(x)[1 - F_X(x)]}{[f_X(x)]^2} \geq -2 \quad \forall\, x \in [a,b] \qquad (2)$$

For strict concavity, the inequality needs to be strict.

The condition given by equation (2) is referred to as the (strict) *concavity condition*. The concavity of $y_S(w)$ is used to show that the expected profits of the manufacturer and the supplier are concave functions of the wholesale price. This property is helpful in determining the optimal wholesale price in games where one of the parties controls the wholesale price.

What types of distributions meet this concavity condition? In Appendix 1, I show that IFR distributions satisfy the strict concavity condition (Lemma 7). Therefore a wide class of distributions, including the exponential, normal, uniform, weibull (as long as the shape parameter exceeds 1) and gamma (similar restriction to the weibull) distributions, satisfy the strict concavity condition.

2.4 Two Games with a Constant Wholesale Price Per Unit

In this section, two different games are analyzed, one in which the wholesale price schedule is exogenous and not under the control of either party, and another in which the wholesale price schedule is set by the manufacturer. In both games the wholesale price per unit, w, is constant regardless of the number of units ordered by the manufacturer, that is $W(Q)=wQ$. I assume that the manufacturer acts as the Stackleberg leader in both of the games, that is the manufacturer makes its decisions first and then the supplier makes its decision knowing the manufacturer's decision. As the manufacturer knows how the supplier respond to its decisions, the manufacturer benefits from being the leader.

2.4.1 The Exogenous Wholesale Price Game

In this game (*Game A*), the wholesale price, w, is exogenous and cannot be controlled by either manufacturer or supplier. The only decision to be made by the manufacturer is its capacity. After deciding its capacity, the manufacturer announces this to the supplier who then determines its capacity. Let y_M* be the optimal capacity chosen by the manufacturer and y_S* be the optimal capacity chosen by the supplier in this Stackleberg game. The following lemma characterizes the optimal solution.

Lemma 8

In Game A, where the wholesale price is exogenous and the manufacturer is the Stackleberg leader, the manufacturer and supplier choose their capacities such that $y_M^*=y_S^*=min\{y_S(w),y_M(w)\}$, *where* $y_M(w)$ *and* $y_S(w)$ *are given by Lemma 2 and Lemma 3 respectively.*

Let y_D* denote the optimal supply chain capacity for Game A, as given in Lemma 8. From Lemma 5, there is a unique wholesale price such that $y_S(w)=y_M(w)$ and at this wholesale price w_{crit}, $y_S(w_{crit})=y_M(w_{crit})=y_I^*$, the coordinated supply chain optimal capacity. For all other w, the supply chain capacity under decentralized decision making is less than the supply chain capacity under coordination, i.e. $y_D^*=min\{y_S(w),y_M(w)\}<y_I^*$. As a consequence, the total profit for Game A is strictly less than the coordinated channel profit, unless the wholesale price w is equal to w_{crit}. Note that this critical wholesale price

is independent of the demand distribution. It depends only on the parties' costs and the retail price.

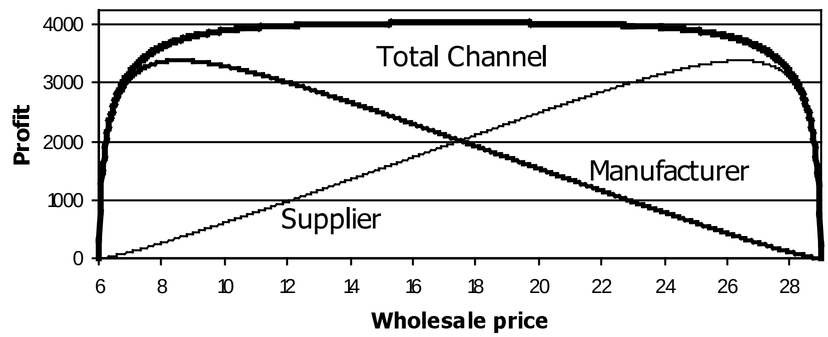
For Example 1 from Table 1, Figure 1 shows the Stackleberg equilibrium supply chain capacity as the exogenous wholesale price varies from c_S+p_S to $r-c_M-p_M$, i.e. over the allowable range.

Figure 1
Supply Chain Capacity (Example 1)

For this example, w_{crit} equals 17.5. For wholesale prices below w_{crit}, the supply chain capacity is equal to $y_S(w)$, in other words it is being dictated by the supplier's reluctance to invest in more capacity. Above w_{crit}, , the supply chain capacity is equal to $y_M(w)$ and it is the manufacturer's reluctance to invest in more capacity that dictates the supply chain capacity.

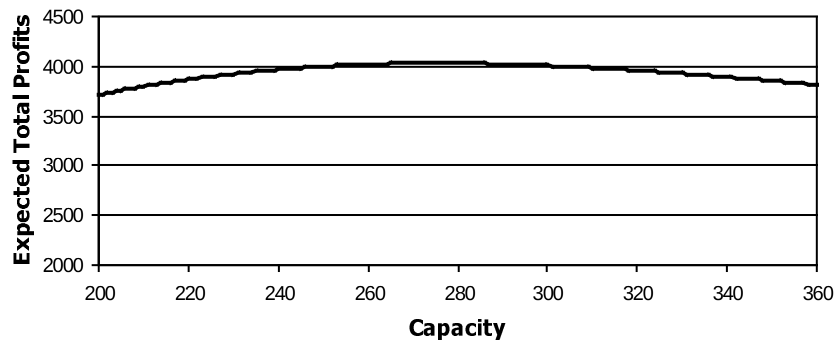
Figure 2 shows the expected profit of the supplier, the manufacturer, and the total channel as the wholesale price varies.

Figure 2
Supply Chain Profits as Exogenous Wholesale Price Varies (Example 1)



As can be seen in this example, the total expected supply chain profit is relatively insensitive to the exogenous wholesale price. Unless the wholesale price is much larger or much smaller than w_{crit}, the total expected supply chain profits are not too far below the optimal coordinated expected supply chain profit. This is caused by two factors. Firstly, the supply chain capacity is less sensitive to the wholesale price near w_{crit}, as can be seen from Figure 1. Secondly, the expected total profits are quite insensitive to the supply chain capacity for a reasonable range of capacities about the supply chain optimum (see Figure 3).

Figure 3
Total Supply Chain Profit as Capacity Increases (Example 1)



The manufacturer and supplier expected profits are very sensitive to the wholesale price with the manufacturer (supplier) doing better for lower (higher) wholesale prices unless the price is too low (high). As the exogenous wholesale price approaches its lower (upper) bound, the supplier (manufacturer) is unwilling to invest in much capacity.

Figures 4, 5, and 6 show the expected profits for Examples 2, 3, and 4 from Table 1.

Figure 4
Supply Chain Profits as Exogenous Wholesale Price Varies (Example 2)

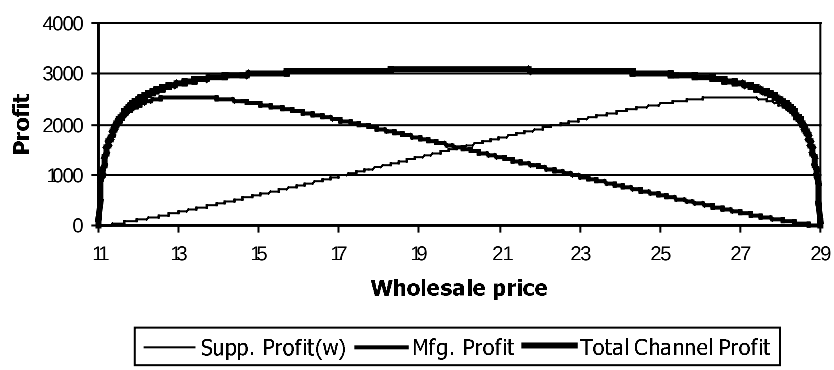


Figure 5
Supply Chain Profits as Exogenous Wholesale Price Varies (Example 3)

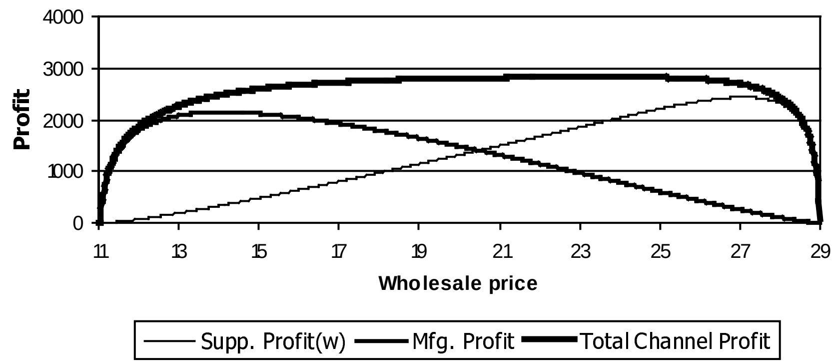
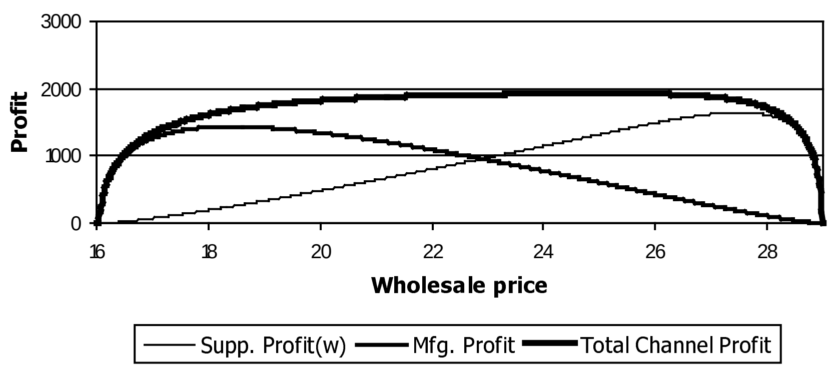
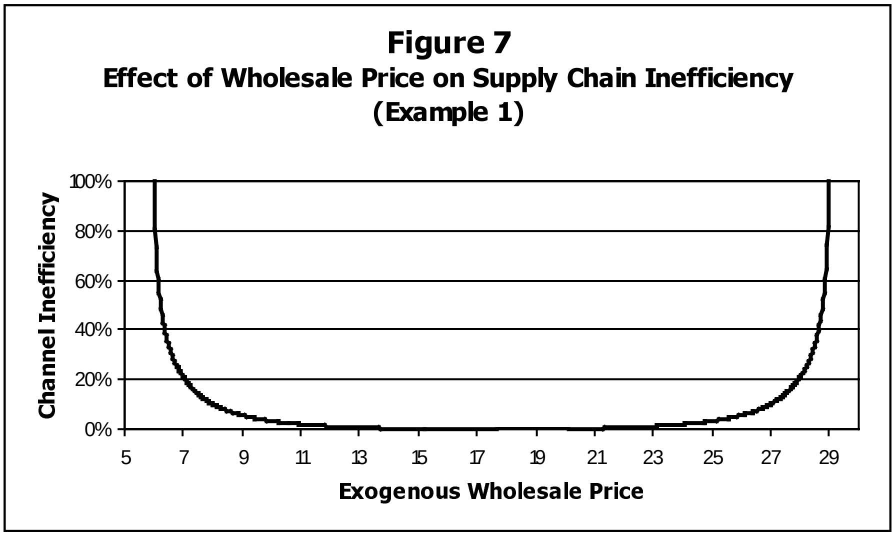


Figure 6
Supply Chain Profits as Exogenous Wholesale Price Varies (Example 4)



These figures display similar relationships between the various expected profits and the wholesale price. For example, the insensitivity of the expected total supply chain profit to the wholesale price occurs in all four examples. It should also be noted that the two manufacturer's and supplier's expected profits are equal at w_{crit} if the costs are all symmetric, as Figure 2 shows for Example 1. However this is not true in general as can be seen most clearly from Figures 5 and 6.

Let the channel inefficiency be defined as $100*(\Pi_C-\Pi_D)/\Pi_C$, i.e. the loss in expected supply chain profit resulting from decentralized decision making, expressed as a percentage of the coordinated system expected supply chain profit. Figure 7 shows the effect of the exogenous wholesale price on channel inefficiency for Example 1.



As mentioned above, the expected total supply chain profit, and hence channel inefficiency, is relatively insensitive to the wholesale price over a reasonable range about w_{crit} (w_{crit} equals 17.5 in the above example). As the wholesale price diverges from w_{crit}, the channel inefficiency increases from 0% all the way up to 100%. So in this exogenous wholesale price game, a constant wholesale price per unit achieves channel coordination only when the wholesale price equals to w_{crit}.

2.4.2 The Wholesale Price Controlled by the Manufacturer

In this game (*Game B*), the wholesale price w is under the control of the manufacturer. Therefore the manufacturer has two decisions to make, its capacity and the wholesale price. Both decisions are made simultaneously before demand is known. For a given wholesale price, the manufacturer's and supplier's optimal capacity decisions are given in Lemma 8 above, i.e. $y_M^*=y_S^*=min\{y_S(w),y_M(w)\}$. So the manufacturer's problem can be expressed as choosing the wholesale price to maximize its expected profit subject to the resulting capacity choice being that given above.

Lemma 9

In Game B, where the manufacturer chooses the wholesale price, the manufacturer never chooses a wholesale price, w, such that $w>w_{crit}$.

The basic idea behind this lemma can be found by looking at Figure 1. For any supply chain capacity (apart from the maximum), there are two wholesale prices that induce this capacity, one smaller than w_{crit} and one larger. The maximum capacity is obtained by w_{crit} alone. The manufacturer never chooses a wholesale price greater than w_{crit} as there exists a lower wholesale price which induces the same supply chain capacity but gives a larger unit margin to the manufacturer (as seen in Figure 1).

Therefore, the manufacturer's wholesale price choice can be restricted to $w\leq w_{crit}$. For the supplier to invest in capacity, the wholesale price, w, must be larger than the sum of unit capacity cost and unit marginal cost ($w>p_S+c_S$). So the manufacturer's optimal wholesale price falls within $p_S+c_S<w\leq w_{crit}$. In this range, the equilibrium (i.e. the game's optimum for both manufacturer and supplier) supply chain capacity is given by $y_M^*=y_S^*=y_S(w)$.

Let $\Pi_M(w)$ be the manufacturer's expected profit as a function of the wholesale price w. In the region, $p_S+c_S<w\leq w_{crit}$, $\Pi_M(w)$ is given by,

$$\Pi_M(w) = -c_M y_S(w) + (r - p_M - w)\int_a^{y_S(w)} x f_X(x)dx + (r - p_M - w)y_S(w)[1 - F_X(y_S(w))]$$
$$+ v_M \int_a^{y_S(w)} [y_S(w) - x] f_X(x)dx$$

using $y_S(w)$ as the capacity induced by a wholesale price of $w\leq w_{crit}$.

Lemma 10

If $F_X(x)$ satisfies the concavity condition, then the manufacturer's profit, $\Pi_M(w)$, restricted to $p_S+c_S\leq w\leq w_{crit}$, is a strictly concave function.

The optimum wholesale price is then given by either the first order conditions or one of the boundaries, $[p_S+c_S, w_{crit}]$. The following lemma shows that the optimal wholesale price lies in the interior. When I use the term optimum, it should be understood to mean the manufacturer's optimum choice.

Lemma 11

If $F_X(x)$ satisfies the concavity condition of equation (2), then the optimal w for Game B is strictly greater than p_S+c_S and strictly less than w_{crit}.

The fact that the manufacturer's optimal wholesale price is strictly less than w_{crit} has important implications for the expected total channel profit.

Lemma 12

If $F_X(x)$ *satisfies the concavity condition, then in Game B, the total channel profit,* Π_D*, is strictly less than the total channel profits obtained by a central decision-maker,* Π_C.

So this lemma tells us that a constant wholesale price per unit fails to achieve channel coordination when the wholesale price can be set by the manufacturer. In contrast, when the wholesale price is exogenous, it was shown earlier that the channel can achieve complete coordination if the wholesale price happens to coincide with w_{crit}. How large or small is the effect of decentralized decision making on the expected total supply chain profit in this game? Table 2 gives the expected total supply chain profit under coordination for the sixteen examples in Table 1. It also gives the expected profits for the manufacturer, supplier and total supply chain for Game B, i.e. when the manufacturer chooses the wholesale price to maximize its expected profit. The manufacturer's optimization problem was solved by a simple search algorithm. The channel inefficiency, the optimum wholesale price and the critical wholesale price, w_{crit} are also given in Table 2. As defined in Section 2.4.1, the channel inefficiency is the loss in expected supply chain profit resulting from decentralized decision making, expressed as a percentage of the coordinated system expected supply chain profit.

Table 2

	Coordination	Game B					
	Total Channel Profit	**Manuf'er Profit**	**Supplier Profit**	**Total Channel Profit**	**Inefficiency**	**Optimal *w***	w_{crit}
Example 1	4031.58	3387.96	348.74	3736.70	7.31%	8.53	17.50
Example 2	3073.54	2549.34	286.78	2836.12	7.72%	13.16	20.00
Example 3	2825.10	2148.08	332.97	2481.04	12.18%	14.01	24.09
Example 4	1924.75	1433.52	254.12	1687.64	12.32%	18.46	25.45
Example 5	3073.54	2549.34	286.78	2836.12	7.72%	8.16	15.00
Example 6	2129.68	1736.18	220.91	1957.09	8.10%	12.76	17.50
Example 7	1924.75	1433.52	254.12	1687.64	12.32%	13.46	20.45
Example 8	1064.77	771.02	151.71	922.72	13.34%	17.69	21.82
Example 9	2825.10	2437.80	246.58	2684.38	4.98%	7.91	10.91
Example 10	1924.75	1644.57	179.08	1823.66	5.25%	12.49	14.55
Example 11	1776.61	1381.46	228.89	1610.35	9.36%	13.28	17.50
Example 12	964.12	735.49	126.14	861.64	10.63%	17.48	20.00
Example 13	1924.75	1644.57	179.08	1823.66	5.25%	7.49	9.55
Example 14	1064.77	898.33	109.90	1008.23	5.31%	12.03	13.18
Example 15	964.12	735.49	126.14	861.64	10.63%	12.48	15.00
Example 16	257.13	190.49	37.66	228.15	11.27%	16.66	17.50

The inefficiency ranges from 5.0% to 13.3%, with an average of 9.0%. This inefficiency might be eliminated by an alternative price schedule. But the manufacturer is only interested in such a schedule if its expected profit is at least as large as for a constant wholesale price schedule.

2.5 Quantity Premium Price Schedules Controlled by the Manufacturer

In the exogenous constant wholesale price game (*Game A*), the supply chain capacity is limited by the supplier's capacity choice if the wholesale price is less than w_{crit}. The manufacturer would like the supplier to invest in more capacity but the supplier is unwilling to do so. Similarly in *Game B*, where the manufacturer controls the wholesale price, the supply chain capacity is always limited by the supplier's capacity choice and the manufacturer suffers from the supplier's reluctance to invest in more capacity. Is there a mechanism that the manufacturer can use to induce the supplier to invest in more capacity without reducing its own expected profit?

In this section, I show that a quantity premium wholesale price schedule, correctly priced, induces the supplier to invest in more capacity without sacrificing the manufacturer's own profit. Suppose the manufacturer sets a price schedule by a wholesale price, w, a quantity, Q_P, and a premium, $\Delta>0$. For each unit up to Q_P the unit price is w; for each unit above Q_P, the unit price is $w+\Delta$. The total price for $W(Q)$ units is illustrated in Figure 8 below. This is an example of an incremental quantity price premium schedule with one breakpoint.

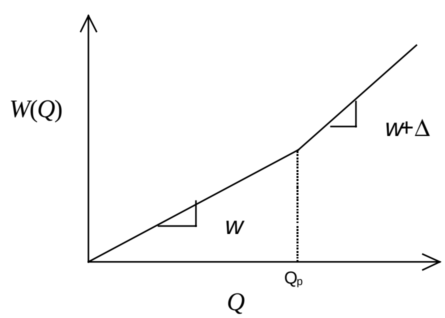


Figure 8

The intuition for quantity premiums is that they provide a mechanism for both the manufacturer and the supplier to share the benefits of high demand scenarios. If the demand is below Q_P, then the manufacturer does not have to pay the quantity premium on any of the units purchased. If the demand is above Q_P, then the manufacturer pays the quantity premium only on the demand above Q_P. The supplier profits in these high demand scenarios as its unit margin increases. The manufacturer gives up some of its unit margin in these high demand scenarios but still increases its expected profit because the supplier is now willing to invest in more capacity, thus allowing the manufacturer to capture more sales in these high demand scenarios. In essence, the manufacturer is trading some unit margin (and capacity investment) for increased expected sales. Of course, this intuition assumes that the quantity premium price schedule is constructed in a reasonable manner. If Q_P is too low or Δ too large, then the manufacturer gives up too much margin and suffers from the policy.

Quantity premium price schedules can take on many forms. They can be incremental price schedules, as in Figure 8, or continuous. A continuous schedule is one in which the marginal price continuously increases.

In this dissertation (Chapter 2) I restrict incremental price schedules to having either one or two breakpoints. If there is only one breakpoint, then for a wholesale price w this breakpoint, Q_P, is set equal to $y_S(w)$, the capacity the supplier is willing to invest in for this constant wholesale price w. The manufacturer is only interested in offering a quantity premium when it wants the supplier to invest in more capacity. The manufacturer has no interest in setting Q_P to be lower than $y_S(w)$, as the supplier is already willing to build capacity up to $y_S(w)$. Setting a breakpoint above $y_S(w)$ may be beneficial but causes the supplier's profit to be a bimodal function of its capacity. This bimodality makes the optimization problems slightly more difficult but does not change the insight that quantity premiums are a valuable tool. If there are two breakpoints, then the first one, Q_P^1, is set equal to $y_S(w)$, and the second, Q_P^2, is set equal to $y_S(w,\Delta_1)$, where Δ_1 is the price premium paid on units between Q_P^1 and Q_P^2 and $y_S(w,\Delta_1)$ is the supplier capacity induced by a single breakpoint schedule (w,Δ_1). An incremental wholesale price schedule is be denoted by (w,Δ) if there is only one breakpoint and (w,Δ_1,Δ_2) if there are two breakpoints. The actual breakpoints are suppressed as they are specified by w or (w,Δ_1) as above.

These are only two of many possible price schedules. Indeed I conjecture that schedules with more breakpoints would be even better for the manufacturer. In Section 2.5.3, I develop an optimal price schedule for the manufacturer and show that a schedule with one or two breakpoints cannot be optimal. This optimal schedule is a continuous quantity premium schedule. A continuous price schedule is likely to be cumbersome to put into practice, whereas an incremental quantity premium schedule with one or two breakpoints would be easier to implement. In Section 2.5.2.1 and 2.5.2.2, I show that while such incremental price schedules may not be optimal, they perform extremely well.

I will show that the expected profits of the manufacturer, the supplier and the total supply chain all increase when Game A is modified to enable the manufacturer to offer its optimal quantity premium given an exogenous wholesale price w less than w_{crit}.

I will also show that the expected profits of the manufacturer and the total supply chain both increase when Game B is modified to enable the manufacturer to choose a quantity premium in addition to the wholesale price. However the supplier's expected profit can decrease in Game B.

Before studying the effects of such a quantity premium policy on Games A and B, it is worthwhile analyzing the supplier and manufacturer capacity choices. For a given wholesale price $w \leq w_{crit}$, and quantity premium, Δ, let $y_M(w,\Delta)$ be the manufacturer's optimal capacity if the supplier had infinite capacity and let $y_S(w,\Delta)$ be equivalently defined for the supplier.

Lemma 13

For any allowable price schedule (w,Δ), $p_S+c_S \le w+\Delta \le r-c_M-p_M$ *and* $\Delta \ge 0$, *then* (*i*) *the optimal capacity choice,* $y_S(w,\Delta)$, *for the supplier, assuming the manufacturer has infinite capacity, is given by,*

$$y_S(w,\Delta) = F_X^{-1}\left(\frac{m_S+\Delta-c_S}{m_S+\Delta-v_S}\right) = F_X^{-1}\left(\frac{w+\Delta-p_S-c_S}{w+\Delta-p_S-v_S}\right) \ge y_S(w)$$

and (*ii*) *the optimal capacity choice,* $y_M(w,\Delta)$, *for the manufacturer, assuming the supplier has infinite capacity, is given by,*

$$y_M(w,\Delta) = F_X^{-1}\left(\frac{m_M-\Delta-c_M}{m_M-\Delta-v_M}\right) = F_X^{-1}\left(\frac{r-w-\Delta-p_M-c_M}{r-w-\Delta-p_M-v_M}\right) \le y_M(w)$$

Lemma 14

(*i*) $y_M(w,\Delta)$ *[*(*ii*) $y_S(w,\Delta)$*] is strictly decreasing [increasing] in both w and* Δ *(Note that* $F_X(x)$ *is assumed to be continuous and differentiable)*

Lemma 15

(i) For a given wholesale price $w \le w_{crit}$ *there is a unique quantity premium* Δ_{crit}, *given by*

$$w+\Delta_{crit} = \frac{(r-p_M)(c_S-v_S)+p_S(c_M-v_M)+c_M v_S - c_S v_M}{c_S-v_S+c_M-v_M} = w_{crit}\ ,$$

such that $y_M(w,\Delta_{crit}) = y_S(w,\Delta_{crit}) = y_I{}^*$.

Lemma 16

$y_S(w,\Delta)$ *is a concave function of both w and* Δ *iff the concavity condition given in equation (2) holds for* $F_X(x)$.

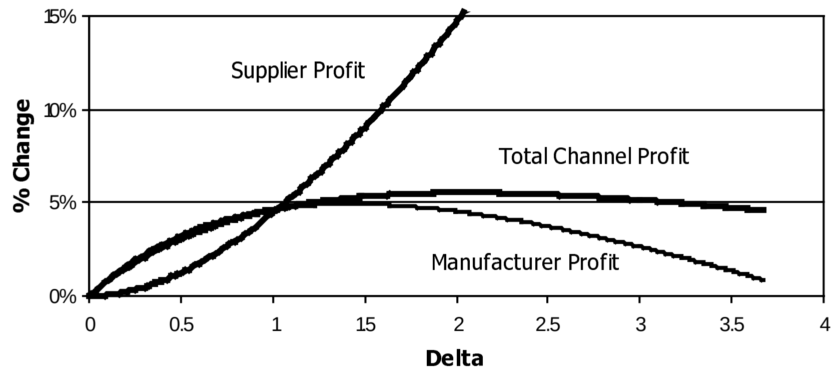
2.5.1 The Exogenous Wholesale Price Game with a Single Breakpoint Quantity Premium

In this game, *Game C,* the wholesale price w is exogenous and is less than or equal to w_{crit}. However the manufacturer chooses a quantity premium, Δ, so as to maximize its expected profit. The sequence of events is that the manufacturer simultaneously chooses the quantity premium and its capacity, announces these to the supplier, who then chooses its capacity. For a given (w,Δ) the optimal manufacturer and supplier capacities are given by $y_M^*=y_S^*=min\{y_S(w,\Delta),y_M(w,\Delta)\}$. As the quantity premium chosen dictates the manufacturer and supplier optimal capacity choices, the manufacturer's problem can be expressed as determining the optimal quantity premium with the understanding that the resulting capacities chosen are optimal for that quantity premium.

Consider Example 10 with the wholesale w price set equal to 12.49, the optimal wholesale price for the manufacturer when no quantity premium is allowed. Then, for this example, the quantity breakpoint Q_P is 175.96. In Figure 9, I show the percentage change (over the no quantity premium case) in the various expected profits as Δ increases.

Figure 9

% Change in Profits as Delta Increases (Example 10)



As can be seen, the expected profits all increase initially. Further increases in Δ still lead to increases in the supplier's expected profit but cause the manufacturer's and supply chain's expected profits to decrease.

The following lemma shows that when the manufacturer chooses the quantity premium, the possible range for Δ can be restricted to $[0,\Delta_{crit}]$.

Lemma 17

For Game C the manufacturer never chooses a quantity premium, Δ, such that $\Delta>\Delta_{crit}$.

Lemma 18

For Game C, if $F_X(x)$ satisfies the concavity condition given by equation (2), then the manufacturer's profit, $\Pi_M(w,\Delta)$, restricted to $0\leq\Delta\leq\Delta_{crit}$, is a strictly concave function of Δ.

Lemma 19

For Game C,

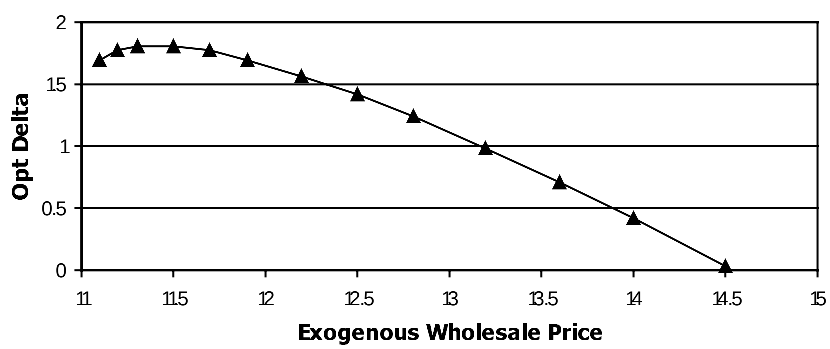
(i) the manufacturer chooses a positive quantity premium, i.e. Δ*(*w*)>0

(ii) the supplier's profit strictly increases with increasing Δ

(iii) the expected total supply chain profit for the price schedule (*w*,Δ*(*w*)), *is strictly greater than the expected total supply chain profit when no quantity premium is offered but the supply chain is not completely coordinated.*

Figure 10 shows the optimum quantity premium as a function of the wholesale price.

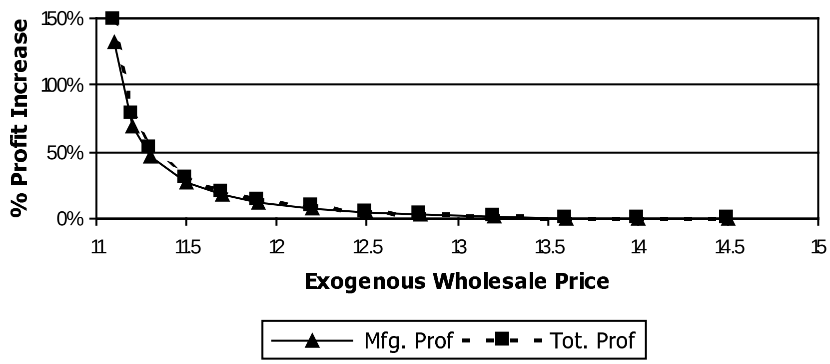
Figure 10
Optimum Premium as Wholesale Price Changes (Example 10)



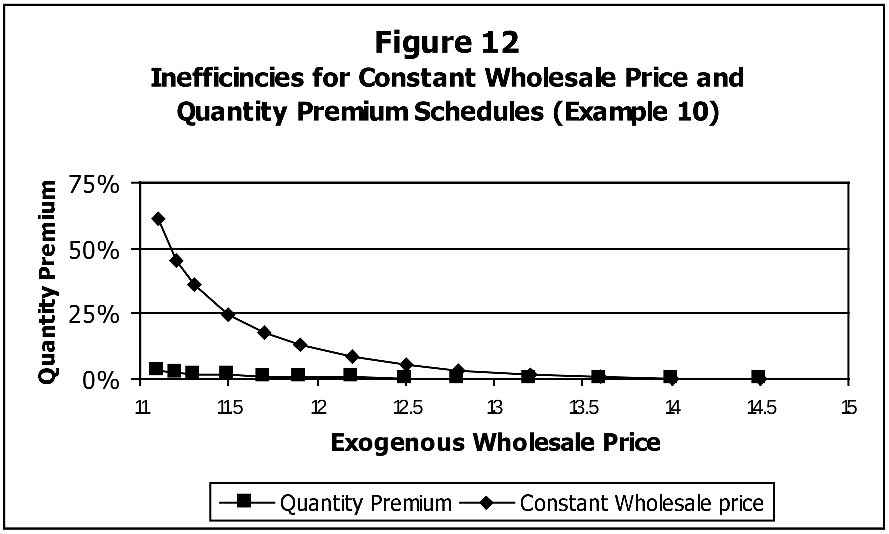
As the wholesale price increases, the increase in expected sales resulting from a quantity premium of Δ decreases. One might then expect the optimal Δ to decrease. While it does decrease after a certain wholesale price, it initially increases. A quantity premium schedule of (*w*,Δ) benefits the manufacturer in that its expected sales increase as the supplier invests in more capacity. However, this increase in expected sales comes at the cost of an increased average wholesale price. The increase in cost depends on both Δ and the range over which this quantity premium is in force. The manufacturer pays the premium for units over $y_S(w)$. For low wholesale prices, $y_S(w)$ is very low and therefore the quantity premium cost is incurred for a large quantity. As w increases, $y_S(w)$ increases rapidly at first (see Figure 1). The range over which the manufacturer pays the quantity premium decreases and therefore it is willing to pay a higher premium.

Figure 11 shows the percentage increase in the manufacturer and total supply chain expected profit resulting from this quantity premium.

Figure 11
% Increases in Profits as Wholesale Price Changes
(Example 10)



The channel inefficiency, with and without a quantity premium, is shown in Figure 12. With a quantity premium chosen by the manufacturer, the inefficiency ranges from 3.22% to 0.01%, which is much lower than the inefficiency of the constant wholesale price case. The difference in inefficiencies is greatest for wholesale prices furthest from w_{crit} (14.55).



2.5.2 The Wholesale Price and Quantity Premium Controlled by the Manufacturer

2.5.2.1 A Single Breakpoint Schedule

In this game, *Game D*, the sequence of events is the same as in *Game B*, but the manufacturer now chooses a wholesale price and a quantity premium. The necessary and sufficient conditions for a wholesale price and quantity premium to be optimal are given in the following lemma.

Lemma 20

(i) If $F_X(x)$ *satisfies the concavity condition, then the manufacturer's profit* $\Pi_M(w,\Delta)$*, restricted to* $p_S+c_S<w+\Delta<w_{crit}$*,* $w>0$ *and* $\Delta>0$*, is a strictly concave function of* w *and* Δ*.*

(ii) The first order conditions for w and Δ are necessary and sufficient for (w,Δ*) to be optimal.*

Clearly the manufacturer is better off when it is allowed to choose a wholesale price and quantity premium than when it is only allowed to choose a wholesale price. However, it is less clear whether the total supply chain profits are greater when a quantity premium can be chosen. The following lemma shows that the total supply chain profits are strictly greater for Game D than for Game B. However it also shows that complete channel coordination is not achieved.

Lemma 21

(i) If $F_X(x)$ *satisfies the concavity condition given by equation (2), then the total expected channel profit when the manufacturer chooses both a wholesale price and a quantity premium is strictly greater than the total channel profit when the manufacturer only chooses a wholesale price but it is strictly less than the total supply chain profit when the supply chain is completely coordinated.*

When the manufacturer can offer a quantity premium, it has a wider range of potential price schedules to choose from. Because the manufacturer is the Stackleberg leader in these games, this wider range of choices gives it more "power" in its relationship with the supplier. This power is derived from the manufacturer's ability to structure a more favorable price schedule from the wider range of schedules available in Game D. In what follows, this increase in power is reflected by the increase in the manufacturer's profit. This increase comes partly at the expense of the supplier's profit.

For the sixteen test examples given in Table 1, Table 3 gives the expected profits, the optimal wholesale price, w^*, and the optimal quantity premium, Δ^*, for Game D. It also gives the sum $w^*+\Delta^*$ for Game D, the optimal wholesale price w for Game B and w_{crit}.

Table 3

	Game D One Breakpoint						**Game B** Constant *w*	
	Manuf'er Profit	**Supplier Profit**	**Total Channel Profit**	**Optimum *w***	**Optimum Δ**	$w^*+\Delta^*$	**Optimum *w***	w_{crit}
1	3734.28	217.31	3951.58	6.92	4.02	10.94	8.53	17.50
2	2827.66	176.68	3004.34	11.81	3.19	15.00	13.16	20.00
3	2472.03	236.58	2708.61	12.31	4.24	16.55	14.01	24.09
4	1664.06	184.91	1848.96	17.13	3.33	20.46	18.46	25.45
5	2827.66	176.68	3004.34	6.81	3.19	10.00	8.16	15.00
6	1941.05	138.33	2079.38	11.72	2.43	14.14	12.76	17.50
7	1664.06	184.91	1848.96	12.13	3.33	15.46	13.46	20.45
8	905.35	110.24	1015.59	16.76	2.13	18.88	17.69	21.82
9	2656.59	137.85	2794.44	6.69	2.45	9.14	7.91	10.91
10	1800.65	100.95	1901.60	11.53	1.81	13.34	12.49	14.55
11	1577.72	147.86	1725.58	11.98	2.71	14.68	13.28	17.50
12	846.89	87.99	934.88	16.68	1.73	18.41	17.48	20.00
13	1800.65	100.95	1901.60	6.53	1.81	8.34	7.49	9.55
14	989.33	62.00	1051.33	11.39	1.13	12.52	12.03	13.18
15	846.89	87.99	934.88	11.68	1.73	13.41	12.48	15.00
16	222.03	26.79	248.82	16.35	0.65	16.99	16.66	17.50

As proven in Lemma 13, the supplier's capacity choice in Game D depends on w and Δ only through the sum $w+\Delta$. The closer this sum is to w_{crit}, the closer the supply chain capacity is to the coordinated supply chain capacity. In Game B, the supplier's capacity choice depends on w. Again the closer this is to w_{crit}, the closer the supply chain capacity is to the coordinated supply chain capacity. For each example, Table 3 shows that the sum $w^*+\Delta^*$ for Game D is larger than the optimal wholesale price w for Game B.

Therefore the supply chain capacity in Game D should be closer to the coordinated capacity than the is capacity in Game B. As proven in Lemma 19, the channel inefficiency should thus be lower in Game D than in Game B.

Table 4 gives the inefficiencies for both Game B and Game D. It also give the percentage change in expected profits of Game D over those in Game B, and the percentage of the lost total profit in Game B recovered in Game D. The loss in total profit in Game B is the coordinated channel profit less the Game B channel profit.

Table 4

	Game B Constant *w*	**Game D** One Breakpoint				
	Inefficiency	**Inefficiency**	**% Change in Manuf'er Profit from Game B**	**% Change in Supplier Profit from Game B**	**% Change in Total Channel Profit from Game B**	**% of Game B Lost Total Profit Recovered**
1	7.31%	1.98%	10.22%	-37.69%	5.75%	72.87%
2	7.72%	2.25%	10.92%	-38.39%	5.93%	70.85%
3	12.18%	4.12%	15.08%	-28.95%	9.17%	66.14%
4	12.32%	3.94%	16.08%	-27.24%	9.56%	68.04%
5	7.72%	2.25%	10.92%	-38.39%	5.93%	70.85%
6	8.10%	2.36%	11.80%	-37.38%	6.25%	70.86%
7	12.32%	3.94%	16.08%	-27.24%	9.56%	68.04%
8	13.34%	4.62%	17.42%	-27.33%	10.06%	65.38%
9	4.98%	1.09%	8.97%	-44.10%	4.10%	78.21%
10	5.25%	1.20%	9.49%	-43.63%	4.27%	77.10%
11	9.36%	2.87%	14.21%	-35.40%	7.16%	69.31%
12	10.63%	3.03%	15.15%	-30.24%	8.50%	71.47%
13	5.25%	1.20%	9.49%	-43.63%	4.27%	77.10%
14	5.31%	1.26%	10.13%	-43.59%	4.27%	76.23%
15	10.63%	3.03%	15.15%	-30.24%	8.50%	71.47%
16	11.27%	3.23%	16.56%	-28.88%	9.06%	71.31%

The inefficiencies for Game D range from 1.1% to 4.6% with an average of 2.7%, compared to an average of 9.0% for Game B. In every example, the inefficiency strictly decreases, as expected from Lemma 21.

The manufacturer's expected profit increase by an average of 13.0% and the total expected channel profit increases by an average of 7.0%. However, the supplier's profits decrease in each example, by an average of 35.1%. In the exogenous wholesale price game, the supplier was strictly better off when the manufacturer chose the quantity premium. In these examples, when the manufacturer can also choose the wholesale price, it chooses a lower wholesale price then when no quantity premium is offered, see Table 3. This reduces the supplier's expected profit. This is a manifestation of the extra power gained by the manufacturer when it uses a quantity premium policy.

The manufacturer's expected profit increases because the total supply chain capacity increases but also because the manufacturer captures some of the supplier's profit relative to Game B, about one third in the sixteen test examples.

The total expected channel profit increases in Game D. The single breakpoint price schedule recovers 71.6% of Game B's lost profit on average, i.e. Game D closes 71.6% of the profit gap between Game B and a completely coordinated supply chain.

2.5.2.2 A Schedule with Two Breakpoints

As the manufacturer benefits from a single breakpoint schedule, a schedule with two breakpoints is analyzed to determine the effect on both the manufacturer's profit and the supply chain inefficiency. This game is referred to as Game E.

For the sixteen test examples given in Table 1, Table 5 gives the expected profits, the optimal wholesale price w^* and the optimal quantity premiums Δ_1^* and Δ_2^* for Game E. It also gives the sum $w^*+\Delta_1^*+\Delta_2^*$ for Game E and $w^*+\Delta^*$ for Game D.

Table 5

	Game E Two Breakpoints							Game D One Breakpoint
	Manuf'er Profit	Supplier Profit	Total Channel Profit	Optimum w	Optimum Δ_1	Optimum Δ_2	$w+\Delta_1+\Delta_2$	$w+\Delta$
1	3845.22	153.27	3998.50	6.58	1.75	4.22	12.54	10.94
2	2918.18	126.24	3044.42	11.45	1.45	3.34	16.24	15.00
3	2590.85	182.41	2773.26	11.79	2.21	4.34	18.34	16.55
4	1750.24	134.10	1884.34	16.66	1.67	3.13	21.47	20.46
5	2918.18	126.24	3044.42	6.45	1.45	3.34	11.24	10.00
6	2009.37	99.12	2108.49	11.39	1.16	2.47	15.03	14.14
7	1750.24	134.10	1884.34	11.66	1.67	3.13	16.47	15.46
8	957.07	85.01	1042.08	16.52	1.16	1.94	19.63	18.88
9	2721.61	90.91	2812.52	6.39	1.08	2.20	9.67	9.14
10	1847.74	69.72	1917.46	11.32	0.84	1.65	13.81	13.34
11	1646.68	107.07	1753.75	11.59	1.36	2.51	15.45	14.68
12	886.78	64.78	951.56	16.44	0.89	1.55	18.88	18.41
13	1847.74	69.72	1917.46	6.32	0.84	1.65	8.81	8.34
14	1017.40	42.18	1059.59	11.22	0.55	0.98	12.74	12.52
15	886.78	64.78	951.56	11.44	0.89	1.55	13.88	13.41
16	233.81	19.58	253.39	16.24	0.37	0.54	17.14	16.99

Table 6

	Game D One Breakpoint	Game E Two Breakpoints				
	Inefficiency	Inefficiency	% Change in Manuf'er Profit from Game B	% Change in Supplier Profit from Game B	% Change in Total Channel Profit from Game B	% of Game B Lost Total Profit Recovered
1	1.98%	0.82%	13.50%	-56.05%	7.01%	88.78%
2	2.25%	0.95%	14.47%	-55.98%	7.34%	87.73%
3	4.12%	1.83%	20.61%	-45.22%	11.78%	84.93%
4	3.94%	2.10%	22.09%	-47.23%	11.66%	82.96%
5	2.25%	0.95%	14.47%	-55.98%	7.34%	87.73%
6	2.36%	0.99%	15.74%	-55.13%	7.74%	87.72%
7	3.94%	2.10%	22.09%	-47.23%	11.66%	82.96%
8	4.62%	2.13%	24.13%	-43.97%	12.94%	84.03%
9	1.09%	0.45%	11.64%	-63.13%	4.77%	91.06%
10	1.20%	0.38%	12.35%	-61.07%	5.14%	92.79%
11	2.87%	1.29%	19.20%	-53.22%	8.90%	86.25%
12	3.03%	1.30%	20.57%	-48.64%	10.44%	87.75%
13	1.20%	0.38%	12.35%	-61.07%	5.14%	92.79%
14	1.26%	0.49%	13.25%	-61.62%	5.09%	90.84%
15	3.03%	1.30%	20.57%	-48.64%	10.44%	87.75%
16	3.23%	1.46%	22.74%	-48.02%	11.06%	87.07%

Table 6 gives the inefficiencies for both Game D and Game E. It also give the percentage change in expected profits of Game E over those in Game B, and the percentage of the lost total profit in Game B recovered in Game E. The loss in total profit in Game B is the coordinated channel profit less the Game B channel profit.

The average inefficiency in Game E is 1.2%. As in Game D, the manufacturer's expected profit is larger than in Game B, by 17.5% on average. This is partly due to the increase in the total supply chain capacity. The increase also occurs because the manufacturer captures some of the supplier's profit relative to Game B, about one half in the sixteen test examples. On average 87.7% of the lost profit in Game B is recovered in Game E.

Table 7 summarizes the results for the different price schedules. A price schedule with two breakpoints both increases the manufacturer's expected profit and nearly achieves complete channel coordination.

Table 7

Minimum, Maximum and Average Taken Over The 16 Test Examples

		Inefficiency	**% Change in Manuf'er Profit from Game B**	**% Change in Supplier Profit from Game B**	**% Change in Total Channel Profit from Game B**	**% of Game B Lost Total Profit Recovered**
Constant Wholesale Price (Game B)	min	4.98%	-	-	-	-
	max	13.34%	-	-	-	-
	ave	**8.98%**	-	-	-	-
One Breakpoint (Game D)	min	1.09%	8.97%	-44.10%	4.10%	65.38%
	max	4.62%	17.42%	-27.24%	10.06%	78.21%
	ave	**2.65%**	**12.98%**	**-35.14%**	**7.02%**	**71.58%**
Two Breakpoints (Game E)	min	0.38%	11.64%	-63.13%	4.77%	82.96%
	max	2.13%	24.13%	-43.97%	12.94%	92.79%
	ave	**1.18%**	**17.49%**	**-53.26%**	**8.65%**	**87.70%**

2.5.3 A Continuous Quantity Premium Schedule

In this section, I introduce a continuous and twice differentiable quantity premium price schedule that completely coordinates the supply chain. Moreover, it is an optimal price schedule from the manufacturer's perspective, allowing the manufacturer to capture the entire expected supply chain profit.

$W(Q)$ is the total price paid by the manufacturer to the supplier for Q units. If the marginal price is increasing ($d^2W(Q)/dQ^2>0$), then $W(Q)$ is said to be a continuous quantity premium price schedule. I assume that $W(0)=0$, that is the manufacturer does not pay any fixed price for ordering. A price schedule is then specified by the marginal wholesale price of the Q^{th} unit.

If there is positive demand with probability one, then the existence of a fixed price F in $W(Q)$ does not change the coordination properties of the price schedule introduced in this section. It does however serve to transfer a profit of F from the manufacturer to the supplier.

A continuous price schedule allows the manufacturer to choose the schedule from an infinite number of possible schedules. As discussed in Section 2.5.2.1, this enlarged range of choices should give the manufacturer even more power. The following lemma shows that the manufacturer can capture the entire expected supply chain profit, leaving the supplier with nothing.

Lemma 22:

The following continuous quantity premium price schedule[1] *is an optimal wholesale price schedule for the manufacturer,*

$$\frac{dW(Q)}{dQ} = \frac{c_S - v_S F_X(Q)}{1 - F_X(Q)} + p_S \tag{3}$$

Furthermore, it completely coordinates the supply chain but leaves the supplier with an expected profit of zero.

The price schedule in (3) is a quantity premium schedule as

$$\frac{d^2W(Q)}{dQ^2} = \frac{(c_S - v_S) f_X(Q)}{\left(1 - F_X(Q)\right)^2} > 0$$

Recall that $c_S > v_S$, i.e. salvage value is strictly less than the capacity cost.

The intuition behind this price schedule is that it leaves the supplier with an expected profit of zero regardless of its capacity choice. The supplier's marginal cost of providing the Q^{th} unit of an order comprises two elements, the marginal processing cost p_S and the marginal capacity cost. The manufacturer reimburses the supplier for the processing cost by the second term in (3). The supplier's marginal capacity cost is c_S. However, the supplier recoups some of this if the demand is less than Q. Effectively the supplier's capacity cost is $c_S - v_S F_X[Q]$. The probability that the manufacturer orders at least Q

[1] The motivation for this price schedule arose from personal correspondence with Prof. Martin Lariviere. For a supply chain model in which only the supplier needs to invest in capacity and in which salvage values and marginal costs are normalized to zero, he conjectured that $dW(Q)/dQ = c/[1-F_X(Q)]$ would coordinate the channel.

units is 1-$F_X[Q]$. For the supplier's expected marginal capacity cost to be zero, the manufacturer must reimburse the supplier $(c_S - v_S F_X[Q])/(1-F_X[Q])$ for the Q^{th} unit. This is the fist term in (3). The supplier is therefore indifferent to any capacity choice as its expected profit is zero [(2)]. The manufacturer captures all of the expected supply chain profit. This profit is maximized by choosing a capacity equal to the coordinated supply chain capacity.

A continuous price schedule would be very difficult to implement in practice. In Sections 2.5.1 and 2.5.2, incremental quantity premium schedules with one or two breakpoints were analyzed and shown to perform quite well with regard to channel coordination.

Table 8 shows the inefficiencies for the sixteen test examples for the four different price schedules analyzed, a constant wholesale price schedule, an incremental quantity premium schedule (one and two breakpoints) and the continuous quantity premium schedule (3). It also shows the minimum, maximum and average over the sixteen examples. For each schedule, the results are shown for the manufacturer's optimal choice. The two breakpoint schedule has an average inefficiency of 1.18% whereas the continuous schedule has an inefficiency of zero from Lemma 22. This means that the total expected supply chain profit increases on average by 1.18% if the continuous price schedule is implemented instead of the two breakpoint schedule. The benefit to the

[(2)] I make the assumption that the supplier still chooses the same capacity as that announced by the manufacturer even though it is indifferent to all capacity choices. One can add an arbitrarily small quadratic term $\varepsilon(y_I - Q)^2$ to the wholesale price schedule $W(Q)$ to ensure that the supplier's optimal capacity choice is the coordinated supply chain capacity. The results of Lemma 22 still hold in this case.

supply chain of implementing a continuous schedule is therefore quite small. The performance of the more easily implemented two breakpoint schedule is almost identical.

As the manufacturer is interested in its expected profit rather than the supply chain profit, the percentage increase in the manufacturer's expected profit from implementing the continuous quantity premium schedule is of interest. This is shown in Table 8 for the constant wholesale price schedule and the incremental quantity premium schedules (one and two breakpoints).

Table 8

	Inefficiency				**% Increase in Manufacturer Profit from implementing the Continuous Quantity Premium Price Schedule**		
	Constant Wholesale Price	**One Breakpoint**	**Two Breakpoints**	**Continuous Quantity Premium**	**over Constant Wholesale Price**	**over One Breakpoint**	**over Two Breakpoints**
1	7.31%	1.98%	0.82%	0.00%	19.00%	7.96%	4.85%
2	7.72%	2.25%	0.95%	0.00%	20.56%	8.70%	5.32%
3	12.18%	4.12%	1.83%	0.00%	31.52%	14.28%	9.04%
4	12.32%	3.94%	2.10%	0.00%	34.27%	15.67%	9.97%
5	7.72%	2.25%	0.95%	0.00%	20.56%	8.70%	5.32%
6	8.10%	2.36%	0.99%	0.00%	22.66%	9.72%	5.99%
7	12.32%	3.94%	2.10%	0.00%	34.27%	15.67%	9.97%
8	13.34%	4.62%	2.13%	0.00%	38.10%	17.61%	11.25%
9	4.98%	1.09%	0.45%	0.00%	15.89%	6.34%	3.80%
10	5.25%	1.20%	0.38%	0.00%	17.04%	6.89%	4.17%
11	9.36%	2.87%	1.29%	0.00%	28.60%	12.61%	7.89%
12	10.63%	3.03%	1.30%	0.00%	31.08%	13.84%	8.72%
13	5.25%	1.20%	0.38%	0.00%	17.04%	6.89%	4.17%
14	5.31%	1.26%	0.49%	0.00%	18.53%	7.63%	4.66%
15	10.63%	3.03%	1.30%	0.00%	31.08%	13.84%	8.72%
16	11.27%	3.23%	1.46%	0.00%	34.98%	15.81%	9.98%
min	4.98%	1.09%	0.38%	0.00%	15.89%	6.34%	3.80%
max	13.34%	4.62%	2.13%	0.00%	38.10%	17.61%	11.25%
ave	8.98%	2.65%	1.18%	0.00%	25.95%	11.38%	7.11%

On average over the sixteen examples, the continuous schedule increases the expected manufacturer profit by 25.95% over the constant wholesale price schedule, by 11.38% over the one breakpoint schedule and by 7.11% over the two breakpoint schedule. By implementing a two breakpoint schedule, the manufacturer is giving up 7.11% potential expected profit. While this is a lot less than the 25.95% given up by a constant wholesale price schedule, this lost profit may still be significant.

A two breakpoint schedule appears to be a reasonable compromise. The manufacturer's expected profit is not too far from its maximum possible expected profit. From the supply chain perspective, the total channel profit is very close to the coordinated channel profit, with an inefficiency of only 1.18% over the sixteen examples.

2.6 Games in which the Wholesale Price Schedule is Controlled by the Supplier

Up to this point, in all the games where there was some element of the wholesale price schedule under the control of the supply chain participants, I have assumed that the manufacturer had control over the wholesale pricing decision. In some supply chains, the supplier may have control over the wholesale pricing decisions. This section analyzes games in which the supplier controls the wholesale price schedule.

It turns out that all the earlier results have directly equivalent results for the case of a supplier controlled price schedule. These results can be proved independently of the earlier sections. However the results can be shown to follow directly from the previous sections by considering the earlier games in terms of a leader and a follower rather than a manufacturer and a supplier. The leader is the participant who has control over the wholesale price schedule and is the first to announce capacity.

Unfortunately one cannot immediately infer that the games analyzed in the previous section follow directly. The difficulty lies in the payments and receipts of the participants. If the manufacturer is the leader, then for an order of Q units, it pays a total price of $W(Q)$. It then sells these Q units for a total revenue of rQ. The supplier (the follower) pays nothing for the Q units but receives a total revenue of $W(Q)$ from the supplier. If the supplier is the leader then the payments and receipts of the leader and follower are reversed. As the leader chooses a wholesale price schedule to maximize its expected profit, it would then seem necessary to specify whether the leader is the manufacturer or supplier.

In fact one can avoid this difficulty by specifying the problem in terms of a *price difference* schedule rather than a wholesale price schedule. Define the total price difference of Q units $PD(Q)$ as the total revenue received from selling the Q units less the total price paid for the Q units. For the manufacturer, $PD_M(Q)=rQ-W(Q)$. For the supplier $PD_S(Q)=W(Q)$. For a given retail price, the wholesale price schedule is completely specified by the price difference schedule of either the manufacturer or supplier. Note also that $PD_M(Q)= rQ-PD_S(Q)$ and $PD_S(Q)= rQ-PD_M(Q)$, so the price difference schedule of one participant is completely specified by the price difference schedule of the other participant. Moreover if the follower's price difference schedule is specified by $PD_F(Q)$, then the leader's price schedule is given $rQ-PD_F(Q)$ irrespective of which participant is the leader. Let p_L and p_F be the marginal processing cost of the leader and follower respectively. These are assumed to be constant for all Q throughout this dissertation (Chapter 2). The leader's total profit from fulfilling an order of Q units is then given by $rQ\text{-}PD_F(Q)\text{-}p_LQ$ and the follower's by $PD_F(Q)\text{-}p_FQ$. These profit expressions are the same regardless of which participant is the leader. Thereby by working with the follower's price difference schedule rather than the wholesale price schedule, one can avoid the necessity of specifying whether the leader is the manufacturer or the supplier.

In previous sections, the manufacturer was the leader and the games were analyzed in terms of the wholesale price schedule. The wholesale price schedule is the same as the supplier's price difference schedule. The entire analysis is identical if one thinks in terms of a leader and a follower's price difference schedule rather than a manufacturer and a supplier's wholesale price schedule. Therefore all the results proven for the case in which the manufacturer is the leader carry through directly to the case where the supplier is the leader.

There is however one caveat. All references to a wholesale price schedule or specific wholesale price should instead be interpreted as the follower's price difference schedule or follower's specific price difference. The quantity premium price schedule results still hold but the quantity premium refers to the follower's price difference schedule rather than the wholesale price schedule. If the supplier is the leader then a quantity premium price difference schedule for the manufacturer is a quantity discount wholesale price schedule. To see this remember that $PD_M(Q)=rQ\text{-}W(Q)$. Consider a continuous and twice differentiable wholesale price schedule. By definition a quantity premium price difference schedule is convex, i.e. the average price difference is increasing in the order size Q. The corresponding wholesale price schedule must be concave as

$$\frac{d^2 PD(Q)}{dQ^2} = -\frac{d^2 W(Q)}{dQ^2}.$$

Therefore the average wholesale price is decreasing in the order size which corresponds to a quantity discount schedule.

If the supplier is the leader, then quantity discount wholesale price schedules provide it with a mechanism to increase both its profit and the total supply chain profit.

In Game A, the exogenous wholesale price game, *both* parties can increase their expected profits by agreeing on a single breakpoint price schedule as long as w does not equal w_{crit}. The manufacturer proposes a quantity premium if w is less than w_{crit}. The supplier proposes a quantity discount if w is greater than w_{crit}. If w equals w_{crit}, then neither a premium or discount schedule benefits *both* parties, and thus they are not both willing to accept a single breakpoint schedule.

In games where the *manufacturer controls* the wholesale price schedule, the manufacturer can capture one third of the supplier's profit by implementing a single breakpoint *quantity premium* schedule rather than a constant wholesale price schedule. A two breakpoint schedule enables the manufacturer to capture one half of the supplier's profit. These fractions are based on the sixteen test examples. A specific continuous premium schedule enables the manufacturer to capture all of the supplier's profit. This is true for any example. The manufacturer benefits while the supplier suffers from the quantity premium schedule. See Section 2.5.2.1 for more details.

In games where the *supplier controls* the wholesale price schedule, the supplier can capture one third of the manufacturer's profit by implementing a single breakpoint *quantity discount* schedule rather than a constant wholesale price schedule. A two breakpoint schedule enables the supplier to capture one half of the manufacturer's profit. These fractions are based on the sixteen test examples[(3)]. A specific continuous discount schedule enables the supplier to capture all of the manufacturer's profit. This is true for any example. The supplier benefits while the manufacturer suffers from the quantity discount schedule.

2.7 The Multiple Supplier Model

Manufacturers often require multiple components when producing a product, with the various components being sourced from different suppliers. In such situations, the manufacturer requires more than one supplier to invest in capacity. In this section, the supply chain is assumed to comprise one manufacturer and *N* suppliers, as opposed to a

[(3)] There is a direct equivalence between Game 1 in which the manufacturer is the leader and Game 2 in which the supplier is the leader if $c^2_S=c^1_M$, $v^2_S=v^1_M$, $p^2_S=p^1_M$, $c^2_M=c^1_S$, $v^2_M=v^1_S$, $p^2_M=p^1_S$. The superscript specifies the game to which the parameter refers. The expected profits of the supplier (manufacturer) in Game 2 is the same as the manufacturer (supplier) in Game 1 when the two games have this mirror structure. Because the sixteen test examples are created in a high/low manner (see Table 1), each example has a specific mirror example contained in the sixteen examples. The results developed for the games in which manufacturer is the leader can be used to determine the results for the game in which the supplier is the leader.

single supplier as in previous sections. The suppliers do not produce the same component but rather each produces a distinct component. The manufacturer assembles these components into the end product.

As in previous sections, demand for the end product is uncertain. The manufacturer and suppliers must each invest in capacity before demand is realized. After demand is realized, the manufacturer purchases components from each of the suppliers, assembles them into the end product, and then sells them. The model parameters are the same as in the single supplier case but supplier parameters are labeled n=1,…,N.

I restrict attention to constant wholesale price schedules, $W(Q)=wQ$. I analyze two different games in this section, the first in which the wholesale prices are exogenous, and the second in which they are under the control of the manufacturer. In each game all parties are assumed to know all the other parties costs. All parties know the end product demand distribution also.

A central decision-maker invests in the same capacity for all suppliers and the manufacturer. The central decision-maker's problem can be mapped into a single supplier problem with the single supplier parameters being replaced by the summation over the N suppliers. From Lemma 1, the optimal capacity is then given by,

$$y_I^* = F_X^{-1}\left(\frac{r - p_M - p_s^{Tot} - c_M - c_s^{Tot}}{r - p_M - p_s^{Tot} - v_M - v_s^{Tot}}\right) \text{where } p_s^{Tot} = \sum_{n=1}^{N} p_{S_n} ,\ c_s^{Tot} = \sum_{n=1}^{N} c_{S_n} \text{, and } v_s^{Tot} = \sum_{n=1}^{N} v_{S_n} .$$

2.7.1 The Exogenous Wholesale Prices Game

In this game (as in *Game A* for a single supplier), the wholesale prices, $w_1,\ldots,w_N$, are exogenous and cannot be controlled by either manufacturer or supplier. The only decision to be made by the manufacturer is its capacity. After deciding its capacity, the manufacturer announces this to the suppliers whom then determine their own capacities. Suppliers are assumed to make their capacity decisions in the following order, supplier 1 decides its capacity and announces this capacity, then supplier 2 follows, then supplier 3 etc. Let y_M* be the optimal capacity chosen by the manufacturer and y_{Sn}* be the optimal capacity chosen by supplier n in this game.

For a given set of wholesale prices $w_1,\ldots,w_N$, let $w_{Tot}=\sum^{N}_{n=1}w_{\mathrm{n}}$. The manufacturer pays a total wholesale price of w_{Tot} for each unit of end product produced. Let $y_M(w_{Tot})$ be the manufacturers optimal capacity if the suppliers all have infinite capacity and let $y_{Sn}(w_n)$ be supplier n's optimal capacity choice if the manufacturer and all other suppliers have infinite capacity. As in Lemmas 2 and 3 earlier,

$$y_M(w_{Tot}) = F_X^{-1}\left(\frac{r - w_{Tot} - p_M - c_M}{r - w_{Tot} - p_M - v_M}\right) \quad , \quad y_{S_n}(w_n) = F_X^{-1}\left(\frac{w_n - p_{S_n} - c_{S_n}}{w_n - p_{S_n} - v_{S_n}}\right)$$

Let $y_S^{min}(w_1,\ldots,w_N)=min_n\{y_{Sn}(w_n)\}$.

Lemma 23

In the N supplier exogenous wholesale price game, the manufacturer and suppliers choose their capacities to be $y_M^*=y_{S1}^*=..=y_{SN}^*=min\{y_S^{min}(w_1,\ldots,w_N),y_M(w_{Tot})\}$.

In the single-supplier exogenous wholesale price game, there existed a unique wholesale price, w_{crit}, such that for $w=w_{crit}$, $y_M(w)=y_S(w)=y_I*$. At this critical wholesale price, the supply chain capacity chosen was the same as a central decision-maker because $y_M*=y_S*=min\{y_M(w),y_S(w)\}$. The expected total supply chain profit depends only on the capacity choice, so the channel is completely coordinated if the exogenous wholesale price is equal to w_{crit}. For all other $w \neq w_{crit}$, $min\{y_M(w),y_S(w)\}<y_I*$ and the channel fails to be coordinated. The following lemma generalizes the critical wholesale price result to the N supplier case.

Lemma 24:

(i) $y_M(w_{Tot})=y_I*$*, the coordinated channel capacity, iff* $w_{Tot} = w_{Tot}^{crit}$*, where*

$$w_{Tot}^{crit} = \frac{(r - p_M)(c_s^{Tot} - v_s^{Tot}) + (c_M - v_M)p_s^{Tot} + c_M v_s^{Tot} - v_M c_s^{Tot}}{c_M - v_M + c_s^{Tot} - v_s^{Tot}}$$

and $y_M(w_{Tot}) \geq y_I*$ *iff* $w_{Tot} \leq w_{Tot}^{crit}$

(ii) $y_{Sn}(w_n)=y_I*$ *iff* $w_n = w_n^{crit}$*,where*

$$w_n^{crit} = \frac{(r - p_M - p_s^{Tot})(c_{S_n} - v_{S_n}) + p_{S_n}(c_M - v_M + c_s^{Tot} - v_s^{Tot}) - c_S(v_M + v_s^{Tot}) + v_S(c_M + c_s^{Tot})}{c_M - v_M + c_s^{Tot} - v_s^{Tot}}$$

and $y_{Sn}(w_n) \geq y_I*$ *iff* $w_n \geq w_n^{crit}$.

(iii) $\sum_{n=1}^{N} w_n^{crit} = w_{Tot}^{crit}$

(iv) Complete channel coordination occurs iff $w_n=w_n^{crit}$ $n=1,\ldots,N$.

So in the N supplier case, there are N critical wholesale prices, one for each supplier. If the exogenous wholesale prices happen to equal these critical prices, then the channel is completely coordinated. Otherwise the channel fails to be coordinated.

2.7.2 The Wholesale Prices Controlled by the Manufacturer

In this game (as in *Game B* for the single supplier case, Section 2.4.2), the wholesale prices, $w_1,\ldots,w_N$, are under the control of the manufacturer. Therefore the manufacturer has two decisions to make, its capacity and the wholesale prices. Both decisions are made simultaneously before demand is known. For a given set of wholesale prices, the manufacturer's and supplier's optimal capacity decisions are given in Lemma 23 above, i.e. $y_M{*}=y_{S1}{*}=\ldots=y_{SN}{*}=min\{y_S^{min}(w_1,\ldots,w_N),y_M(w)\}$. So the manufacturer's problem can be expressed as choosing the wholesale prices to maximize its expected profits subject to the resulting capacity choice being that given above.

This multiple-supplier game can be transformed into a single-supplier game that is almost identical to Game B. This can be seen as follows.

Firstly, it is never optimal for the manufacturer to set a wholesale price w_k for supplier k such that $y_{Sk}(w_k)>y_S^{min}(w_1,\ldots,w_N)=min_n\{y_{Sn}(w_n)\}$ as w_k can be reduced by an arbitrarily small amount without any effect on the supply chain capacity. The manufacturer's expected sales remain the same but its margin increases. Therefore its expected profit can be strictly increased. For an optimal set of wholesale prices, $w_1{*},\ldots,w_N{*}$, we have $y_{S1}(w_1{*})=y_{Sk}(w_k{*})=\ldots=y_{SN}(w_N{*})$. It is therefore only necessary to consider a single supplier, say supplier 1, with the understanding that the other wholesale prices need to be set such that $y_{S1}(w_1{*})=y_{Sk}(w_k{*})=\ldots=y_{SN}(w_N{*})$. The manufacturer's problem is now to set the wholesale price for supplier 1. The supply chain capacity choice is given by $min\{y_{S1}(w_1),y_M(w_{Tot})\}$.

Secondly, as in Lemma 9 for the single-supplier case, the manufacturer's optimal wholesale price for supplier 1 is always be such that $y_{S1}(w_1)\leq y_M(w_{Tot}^{crit})$, i.e. $w_1\leq w_1^{crit}$. In other words, it is never optimal for the manufacturer to offer a wholesale price such that the supply chain capacity choice is given by $y_M(w_{Tot})$. The manufacturer could lower the wholesale price and still induce the same supply chain capacity. So in any optimal solution, the supply chain capacity is given by $y_{S1}(w_1)$. Supplier 1 only invests in capacity if $w_1>p_{S1}+c_{S1}$, therefore the manufacturer's optimal supplier 1 wholesale price can be restricted to $p_{S1}+c_{S1}<w_1\leq w_1^{crit}$.

This game is almost identical to Game B, the single supplier game of Section 2.4.2. However there is one difference. The manufacturer does not pay w_1 per unit but actually pays $w_{Tot}=\sum_{n=1}^{N} w_n$ as there are N suppliers to pay. As developed above, $y_{S1}(w_1{*})=y_{Sk}(w_k{*})=\ldots=y_{SN}(w_N{*})$ must hold in an optimal solution. It is therefore possible to express w_{Tot} as a function of w_1 only.

This can be done as follows. As $F_X(x)$ is continuous and increasing, then $y_{S1}(w_1)=y_{Sk}(w_k)$ if and only if,

$$\left(\frac{w_{S_1}-p_{S_1}-c_{S_1}}{w_{S_1}-p_{S_1}-v_{S_1}}\right)=\left(\frac{w_{S_k}-p_{S_k}-c_{S_k}}{w_{S_k}-p_{S_k}-v_{S_k}}\right)$$

This is equivalent to $w_k = H_k^1 w_1 + L_k^1$ where

$$H_k^1 = \frac{c_{S_k}-v_{S_k}}{c_{S_1}-v_{S_1}} \quad \text{and} \quad L_k^n = \frac{(c_{S_1}+p_{S_1})(p_{S_k}+v_{S_k})-(c_{S_k}+p_{S_k})(p_{S_1}+v_{S_1})}{c_{S_1}-v_{S_1}}$$

If the manufacturer pays supplier 1 a wholesale price of w_1, then it pays a total $G(w_1)$ to the N suppliers, where $G(w_1)=A_1w_1+B_1$, with $A_1 = 1+\sum_{n=2}^{N} H_1^n > 0$ and $B_1 = \sum_{n=2}^{N} L_1^n$.

If the N suppliers are all identical in terms of costs, then $G(w_1)=Nw_1$.

So, the manufacturer's problem is to choose $p_{S1}+c_{S1}<w_1\leq w_1^{crit}$ such that

$$\Pi_M^N(w_1) = -c_M y_{S_1}(w_1) + (r - p_M - G_1(w_1)) \int_a^{y_{S_1}(w_1)} x f_X(x)dx$$

$$+ (r - p_M - G_1(w_1)) y_S(w_1)[1 - F_X(y_{S_1}(w_1))] + v_M \int_a^{y_{S_1}(w_1)} [y_{S_1}(w_1) - x] f_X(x)dx$$

is maximized.

In the single supplier case, the channel is not completely coordinated when the manufacturer chooses the wholesale price. The following lemma generalizes this result for the N supplier case.

Lemma 25

In the N supplier wholesale price game, if $F_X(x)$ satisfies the concavity condition, then

(i) *The manufacturer's expected profit as a function of the suppler 1 wholesale price w_1, $\Pi^N_M(w_1)$, restricted to $p_{S1}+c_{S1}<w_1\leq w_1^{crit}$, is a strictly concave function.*

(ii) *The optimal wholesale prices w_n^* are strictly less than w_n^{crit}, $n=1,\ldots,N$.*

(iii) *The channel fails to be completely coordinated.*

So as in the case of a single supplier, the channel fails to be coordinated when the manufacturer controls the wholesale prices.

2.8 Quantity Premiums when the Manufacturer has Unlimited Capacity

In this section, as in Cachon and Lariviere (1997, 1999), the only capacity constraint is that of the supplier. Effectively the manufacturer has infinite capacity. The game proceeds in the following sequence: the manufacturer offers the supplier a price schedule $W(Q)$, the supplier determines its capacity, y_S, demand, x, is realized and the manufacturer acquires $\min\{x, y_S\}$ units from the supplier and sells them at a retail price of r per unit. As in Cachon and Lariviere, the only cost is the supplier's capacity cost c. All other costs are normalized to zero. The salvage value v is less than the capacity cost c.

The demand distribution is assumed to be discrete with three possible states, small (S), medium (M), or large (L). The probability of each demand state is β_S, β_M, and β_L respectively. The only possible optimal capacity choices for the supplier (or a central decision-maker) are S, M, or L. If the supplier's expected profit is the same for two different capacity choices, it is then indifferent between the two. I assume that it chooses the larger of the two capacities. As in Cachon and Lariviere (1997), I assume that the probabilities are such that a central decision-maker would not invest in small capacity. If the central decision maker would invest in small capacity (S), the manufacturer could simply offer a price schedule of $W(Q)=cQ$, the supplier would be willing to invest in a capacity of S as its expected profit is still zero and the manufacturer would capture all the supply chain expected profit. The channel would be completely coordinated.

Let the quantity premium price schedule whereby the manufacturer pays c per unit for the first S units and $c+\Delta_M$ per unit for units above S be denoted by (c, Δ_M). Similarly let (c, Δ_M, Δ_L) denote the quantity premium price schedule whereby the manufacturer pays c per unit for the first S units, $c+\Delta_M$ per unit for the next (M-S) units and $c+\Delta_M+\Delta_L$ per unit for units above M.

Lemma 26

(i) If $\beta_S<(r-c)/(r-v)$, then a central decision maker would invest in either medium or large capacity, investing in medium capacity if $\beta_L\leq(c-v)/(r-v)$ and in large capacity otherwise.

(ii) If $\beta_S<(r-c)/(r-v)$ and $\beta_L\leq(c-v)/(r-v)$, then a quantity premium price schedule of (c,Δ_M^*) induces the supplier to invest in medium capacity, where

$\Delta_M^* = \left(\frac{\beta_S}{1-\beta_S}\right)(c-v)$. *Furthermore, the channel is completely coordinated and the manufacturer captures all the expected supply chain profit.*

(iii) If $\beta_S<(r-c)/(r-v)$ and $\beta_L>(c-v)/(r-v)$, then a quantity premium price schedule of $(c,\Delta_M^*,\Delta_L^*)$ induces the supplier to invest in large capacity, where Δ_M^* is given above and $\Delta_L^* = \left(\frac{1-\beta_L}{\beta_L} - \frac{\beta_S}{1-\beta_S}\right)(c-v)$. *Furthermore, the channel is completely coordinated and the manufacturer captures all the expected supply chain profit.*

In the case of a discrete demand distribution with two states, Cachon and Lariviere (1997) show that when the manufacturer is restricted to choosing a constant wholesale price schedule, $W(Q)=wQ$, then there is a certain range of $\beta_S<(r-c)/(r-v)$ such that the channel fails to be completely coordinated. One can use Lemma 26 for the case of two demand states, small and medium, to see that the quantity premium schedule (c,Δ_M*) completely coordinates the schedule for all β_S<(r-c)/(r-v) and β_L=0.

Lemma 26 tells us that if β_S<(r-c)/(r-v), then the channel can be completely coordinated via a one or two breakpoint quantity premium schedule. If there are two demand states, then only one breakpoint is needed; if there are three demand states, then a two-breakpoint schedule is needed. In both cases the manufacturer captures all of the expected profit.

2.9 Conclusion

In Chapter 2, I show that decentralized single period capacity and pricing decisions can lead to channel inefficiency in multiple-operation supply chains. That is, the total expected supply chain profit is strictly less than that obtained by a central decision-maker. If the wholesale price is constant and exogenous, then the channel inefficiency depends on the wholesale price and is zero at a unique wholesale price w_{crit}. If the manufacturer or supplier controls a constant wholesale price schedule, then the channel inefficiency is always strictly positive. These results have been developed for a single end product and a single period uncertain demand.

In two party supply chains, one party may be reluctant to invest in as much capacity as the other party would like. If the supplier is the reluctant party, then the manufacturer should offer a quantity premium. If the manufacturer is the reluctant party, then the supplier should offer a quantity discount. These piecewise linear price schedules provide a mechanism by which the two parties can share both the benefit of high demand states and the risk of under utilized capacity in low demand states.

If the manufacturer has control over the wholesale price schedule, then quantity premiums are shown to improve channel coordination. A particular continuous quantity premium schedule completely coordinates the channel while an incremental premium schedule with one or two breakpoints approaches complete coordination. Moreover, it is shown that the manufacturer can increase its expected profit by offering a quantity premium price schedule (unless $w \geq w_{crit}$ in the exogenous wholesale price game). In other words, it is in the manufacturer's interest to offer a quantity premium. Whether the supplier benefits from a quantity premium depends on whether the wholesale price is exogenous or under the control of the manufacturer. If the supplier controls the wholesale price schedule then analogous results hold for quantity discounts.

Moorthy (1987) noted that a specific quantity premium schedule would coordinate a two party supply chain in which demand is deterministic but price sensitive and the parties must make pricing decisions. This dissertation shows that quantity premiums also result in increased supply chain profits when the supply chain is faced with making capacity decisions in the face of uncertain but price-insensitive demand.

The bargaining scheme used in this dissertation assumes that the manufacturer (or supplier) has the power. In reality the bargaining process is likely to be more complex. In addition the parties may not share all the cost and demand distribution information. However, the insight that quantity premiums (and discounts) provide a mechanism for the manufacturer and supplier to increase the total supply chain profit should still hold.

This page intentionally left blank

3 Process Flexibility in Supply Chains

3.1 Introduction

Many firms operate in environments characterized by uncertainty, variability, or rapid change, and as such must develop an ability to cope with these challenges. Flexibility is often represented as a key competency for these firms. However, flexibility is a term that, while widely used, has many varying interpretations. Even manufacturing flexibility has a wide range of manifestations (for a survey see De Toni and Tonchia, 1998). One category of manufacturing flexibility is the ability of a plant or machine to process more than one product at the same time. Such flexibility is referred to as process flexibility. In this dissertation, the term flexibility is understood to refer to process flexibility.

Manufacturing firms must often install plant capacity in anticipation of future product demand. However, at the time of capacity commitment, demand is often uncertain. A plant might be dedicated to one product, that is it can process one product and that product only. If a dedicated plant's capacity exceeds the actual product demand, this excess capacity cannot be used for another product whose demand exceeds its plant's capacity. A firm might decide to build sufficient capacity to ensure that the total demand can be met with a high probability; however this strategy is likely to prove expensive as some plants only use a portion of their capacity.

Flexibility provides firms with an alternative means of coping with demand uncertainty. By enabling plants to process multiple products and ensuring that more than one plant can process each product, demand uncertainty can be accommodated more effectively than if dedicated plants are built. When product demands become known, allocations of products to plants can be done more effectively if a plant that processes a low demand product also happens to be capable of processing a high demand one.

A number of authors (e.g. Fine and Freund, 1990, Gupta et al., 1992, Li and Tirupati, 1994, 1995, 1997 and van Mieghem, 1998) have focused on investments in dedicated plants versus totally flexible plants, where a totally flexible plant can process all products. Partial flexibility, whereby a plant can produce a subset of products has received less attention (Jordan and Graves, 1995 and Gavish, 1994).

Jordan and Graves (1995) investigated process flexibility in a single-stage manufacturing system with multiple products and plants. This dissertation (Chapter 3) has a similar focus but aims to understand the role of process flexibility in multiple-stage supply chains and to develop insights into general strategies for process flexibility deployment in supply chains.

If demands are variable or uncertain, supply chains are faced with an issue that does not arise in single-stage systems – the bottleneck stage can vary with demand, where the bottleneck stage is that stage that limits throughput. In an automotive supply chain, the ability to meet customer demand might be limited by engine capacity in one demand scenario but by assembly capacity in another.

In determining the capacity and flexibility configuration of its supply chain, a firm must recognize the possibility of "floating bottlenecks" (e.g. Hopp and Spearman, 1997). If capacity in one stage of the supply chain is significantly more expensive than all other stages, then it may suffice to designate this stage as being the "bottleneck" or limiting stage, with the understanding that other stage capacities are set significantly higher than this stage's. In some sense this approach attempts to ensure that the capital-intensive stage's throughput is never limited by some other stage's capacity. Using the terminology of the theory of constraints, all other stages are subordinate to this constrained stage.

What happens if there is no one stage that dominates the capacity cost or if the supply chain cuts across various firms? In these cases, one may not want or even be able to subjugate other stages' capacity and flexibility investments to one particular stage's decision. Such supply chains still need to make investment decisions that ensure a high expected throughput, but must focus their attention on capacity investments in more than one stage. This research aims to develop an understanding of flexibility investments in these multiple-stage systems.

A flexibility configuration for a particular stage denotes which products can be processed in which plants for that stage. It is most easily represented by the type of graph shown in Figure 1.

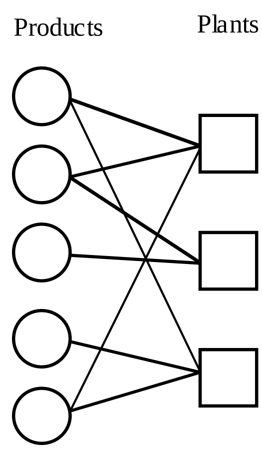


Figure 1

In this representation, circles denote products and squares denote plants. A link from a product to a plant indicates that the plant can process that product. Flexibility investments are then investments in these product-plant links.

The investment must be made before product demands are known and the investment can be costly. If flexibility were cheap, then a firm could enable every plant to process every product. Such total flexibility is an extreme case that provides the best hedge against demand uncertainty. However, it is likely to be prohibitively expensive. I seek flexibility configurations for the different stages of the supply chain that require significantly less investment (i.e. fewer link) but that still perform well.

Jordan and Graves (1995) analyzed a single-stage system and introduced the concept of "chaining" as an effective flexibility strategy. The following definition is taken from the authors. "A chain is a group of products and plants which are all connected, directly or indirectly, by product assignment decisions. In terms of graph theory, a chain is a connected graph. Within a chain, a path can be traced from any product or plant to any other product or plant via the product assignment links. No product in a chain is built by a plant from outside that chain; no plant in a chain builds a product from outside that chain." Figure 2 shows three different configurations for a six-product six-plant stage.

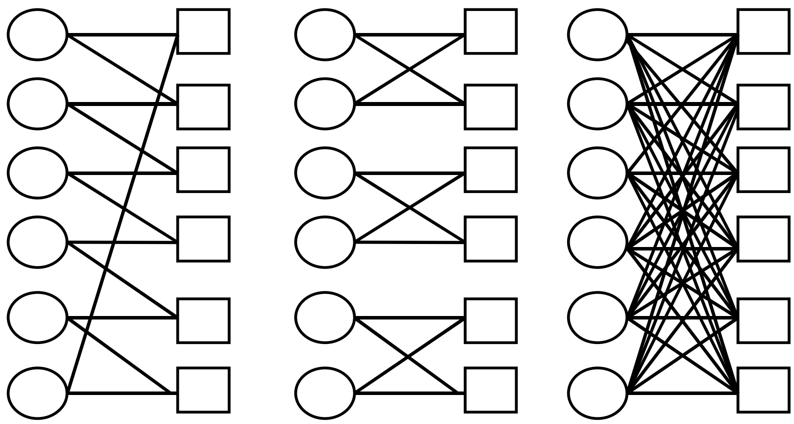


Figure 2

Both the first and second configuration contain chains in which each product can be processed in exactly two plants and each plant can process exactly two products. However, their performance is not equivalent. Jordan and Graves demonstrate that the complete chain configuration, in which all products and plants are contained in one chain, and the chain is "closed", significantly outperforms the configuration that has numerous distinct chains. In fact the complete chain configuration performs remarkably like the total flexibility in terms of expected throughput (or equivalently shortfall) even though it has much fewer product-plant links. The authors develop some flexibility guidelines, (i) try to equalize the capacity to which each product is directly connected, (ii) try to equalize the total expected demand to which each plant is directly connected and (iii) try to create a chain(s) that encompasses as many plants and products as possible.

The authors also develop a flexibility measure, $\Pi(M^*)$, "the maximal probability over all groupings or sets of products (M) that there will be unfilled demand for a set of products while simultaneously there is excess capacity at plants building other products". This is a surrogate measure for the probability that the shortfall of a configuraiton exceeds that of a totally flexible configuraiton. The smaller $\Pi(M^*)$ is, the better. For chains $\Pi(M^*)$ is very small.

Gavish (1994) builds upon the work of Jordan and Graves (1995) to investigate a chaining strategy in a particular seven-product two-stage supply chain, an automotive supply chain comprising the component and assembly stages. He shows that the chaining strategy that is effective for single-stage systems is effective for this particular two-stage supply chain. However, while this suggests that chaining may work well in general supply chains, it is not conclusive. Firstly it studies a particular system. Secondly, the resulting configuration can be shown to have the very special property that the assembly is always the bottleneck stage regardless of the demand realization. This may be a desireable property in that there is no issue of a floating bottleneck, but in general, systems will not exhibit this property and floating bottlenecks may significantly effect performance.

Before presenting the model and research, I will briefly introduce some configuration terminology. In the complete chain example above, each product is connected to two plants. If one numbers the products and plants i=1,…,6, then product i is connected to plants i and i+1 where i+1=1 if i=6. A chain can be represented by the product path starting and ending at product 1. The chain in Figure 2 could be represented by {1,2,3,4,5,6,1}. Numerous other chains are possible e.g. {1,5,3,2,4,6,1}. Instead of having every product being connected to two plants, one could have every product connected to three plants, i.e. product i connected to plants i,i+1 and i+2 for the complete chain shown in Figure 2. A chain in which each product is connected to h plants will be termed an h-type chain. Unless otherwise stated an h-type chain is assumed to be a complete chain. The second chain configuration in Figure 2 can be thought of as a pairwise allocation of products to plants. The products are paired together and then assigned to particular plants. Products that are not paired together cannot be produced in the same plant. Such a configuration will be called a “pairs” configuration. In general one could group together *n* products instead of just two. Such a configuration will be referred to as an *n*-tuple configuration.

Section 3.2 discusses the supply chain model used for this research. While I use the term supply chain, this research is applicable to any multiple-stage manufacturing system. In Section 3.3, I identify two supply chain inefficiencies, in which an inefficiency is a phenomenon that affects multiple-stage systems but not single-stage systems. Performance measurement is covered in Section 3.4. A general class of configurations is introduced in Section 3.5, in which members of the class are distinguished by the lower bound on capacity available to any subset of products. The performance of pairs and chain configurations is analyzed in Sections 3.6 and 3.7. Concluding remarks are presented in Section 3.8. Proofs of all lemmas can be found in Chapter 5(Appendix 2).

3.2 The Model

The supply chain model is a direct extension of the single-stage model of Jordan and Graves (1995) to multiple-stage systems.

The supply chain consists of K different stages, $k=1,\ldots,K$. The numbering of the stages does not imply any specific processing sequence and the supply chain need not be serial in nature. Any general multiple-stage production system in which there are K distinct operations, where an operation is distinct if it requires a different processing resource from all other operations, is allowed. An automotive supply chain might be modeled as a four-stage supply chain, comprising the component, engine, body, and assembly operations.

The supply chain produces I different products, $i=1,\ldots,I$, with each product requiring processing at each stage. This assumption is easily relaxed but is assumed to hold throughout Chapter 3. Each stage k has J_k different plants, $j=1,\ldots,J_k$, where the term plant refers to any processing resource with an associated capacity constraint, e.g. a machine, a department within a plant or a whole plant. Plant j of stage k has a capacity of c^k_j. Plant j of stage k is able to process the set of products $P^k(j)$. The set of plants at stage k that can process product i is given by $Q^k(i)$. Unless otherwise stated, all products $i \in P^k(j)$ are assumed to require the same amount of plant j's capacity per unit processed. Plant capacities are all expressed as the total product units that can be processed in the planning horizon.

The flexibility configuration decision determines which products can be processed in each of the plants and therefore specifies the sets $P^k(j)$ and $Q^k(i)$. A particular configuration can be denoted by the sets $P^k(j)$ $j=1,\ldots,J_k$, $k=1,\ldots,K$. Alternatively, the configuration at each stage k can be denoted by a set of ordered pairs $A_{k,}$ where $(i,j) \in A_k$ iff $i \in P^k(j)$. The supply chain configuration is then denoted by $\mathbf{A}=\{A_1,\ldots,A_K\}$.

The configuration decision must be made before demand is realized. After demand is realized, production planning occurs with the objective of minimizing the total shortfall. Shortfall of product *i* is defined as the maximum of zero and product *i*'s demand less its production. The production planning problem is covered in the next section, after a brief discussion of some notational convention.

Certain problem parameters are stochastic, while others are deterministic. Deterministic parameters, such as plant capacities, are written in lower case. Random variables are written in upper case, and actual realizations in lower case. Thus the product demand random vector is denoted by $\mathbf{D}=\{D_1,\ldots,D_I\}$ and a demand realization vector by $\mathbf{d}=\{d_1,\ldots,d_I\}$. Product shortfalls will also be random variables, with $\mathbf{S}=\{S_1,\ldots,S_I\}$ denoting the random variable vector for shortfalls. $\mathbf{S}$ will be a function of $\mathbf{D}$.

3.2.1 The Production Planning Problem

The flexibility configuration must be determined in advance of product demands being known. After product demands are realized production planning occurs, a common assumption in the flexibility and capacity planning literature (Eppen, Martin and Schrage, 1989, Jordan and Graves, 1995, Harrison and Van Mieghem, 1999). As noted by Jordan and Graves (1995), "this is an approximation since in practice one must allocate production capacity in real time as demand is realized." However the approximation is reasonable if the product demand is relatively constant over the length of the horizon or if demand can be backlogged until the end of the horizon.

For a given demand realization, $\mathbf{d}=\{d_1,\ldots,d_I\}$, and flexibility configuration, $\mathbf{A}=\{A_1,\ldots,A_K\}$, the production planning problem can be formulated as the following linear program, **P1**(**d**,**A**),

$$
\begin{aligned}
sf(\mathbf{d},\mathbf{A}) = \underset{\mathbf{x},\mathbf{s}}{\text{Min}} \{\sum_{i=1}^{I} s_i\} & \\
\text{subject to} & \\
\sum_{j\in P^k(i)} x_{ij}^k + s_i \geq d_i \qquad & i=1,\ldots,I \quad k=1,\ldots,K \\
\sum_{i\in Q^k(j)} x_{ij}^k \leq c_j^k \qquad & j=1,\ldots,J_k \quad k=1,\ldots,K \\
\mathbf{x},\mathbf{s} \geq \mathbf{0} &
\end{aligned}
$$

where *sf* is the minimum possible shortfall, x^k_{ij} is the amount of product i processed in plant j (at stage k) over the planning horizon, and the other parameters are defined above.

While all parameters are deterministic for the production planning problem, the demands are stochastic when the flexibility configuration is being decided. For a given configuration **A**, I will consider the expected total shortfall, E[*SF*(**D,A**)], where the expectation is over the demand vector **D**.

The configuration problem is then to determine a configuration, **A**, that results in a low expected total shortfall without requiring a prohibitively expensive investment in flexibility. For a specific problem instance, in which the product-plant link costs, plant capacities and demand scenarios are specified, one could use **P1**(**d,A**) to formulate the configuration problem as a stochastic linear integer program with recourse. The flexibility configuration **A** is determined subject to a budget constraint, so as to minimize E[*SF*(**D,A**)], where *sf*(**d,A**) is given by **P1**(**d,A**) for each demand realization. Birge and Louveaux (1997) provide a stochastic program formulation for the single-stage supply chain flexibility problem.

The purpose of this dissertation is neither to solve particular instances of the stochastic program nor to develop efficient solution techniques. Rather the aim is to develop an understanding of what configuration properties drive the shortfall (or throughput) in multiple-stage supply chains and from this understanding, develop general flexibility strategies that yield low expected shortfalls without requiring large capital investments.

In the following sections the dependence of problem **P1** on **d** and **A** may be suppressed in the notation.

It is worth commenting on the assumption of equal product capacity usage in the model (i.e. the same amount of a plant's capacity is required per unit of each product it can process). This is clearly a strong assumption as in practice the usage will likely vary. Product dependent capacity usages introduce an added complexity to the configuration decision, as now product to plant allocations should also consider the capacity usage of each product-plant pairing. In order to enable analytical tractability and simpler analyses, I assume that, as in Jordan and Graves (1995), the capacity usage is not product dependent unless otherwise stated.

3.2.1.1 A Lower Bound on the Minimum Shortfall

For certain supply chains, an alternative expression for the minimum total shortfall obtained in **P1(d,A)** is given in the following lemma. In general, this expression is a lower bound on the minimum total shortfall.

Lemma 1:

(i) A lower bound for the minimum shortfall in problem ***P1(d,A)*** *is given by problem* ***P2(d,A)****,*

$$\underset{M}{\text{Max}}\ \{\sum_{i\in M} d_i - \min_{L_1,..L_K}\{\sum_{k=1}^{K}\sum_{j\in P^k(L^k)} c_j^k\}\}$$

subject to

(i) $M \subseteq \{\varnothing,1,\ldots,I\}$

(ii) $L_k \cap L_{k'} = \varnothing \quad \forall k \neq k'$

(iii) $\bigcup_{k=1}^{K} L_k = M$

(ii) If either the number of stages K or the number of products I is less than three, then the minimum shortfall in problem ***P1(d,A)*** *is equal to the lower bound in (i).*

$P^k(L_k)$ is the set of plants at stage k that can process any product $i \in L_k$. **P2(d,A)** is a generalization, for multiple-stage systems, of the shortfall expression, $V(A)$, of Jordan and Graves (1995).

Consider any subset of products, M. This subset can be partitioned into K subsets $L_1,\ldots,L_K$. For each stage $k=1,\ldots,K$, the production of products in subset L_k is bounded by the total capacity at stage k available to that subset of products. The production of products in M is bounded by the sum of these K upper bounds. The internal minimization in **P2(d,A)** partitions the subset M so that the upper bound on the production of products in M is minimized. The shortfall for subset M is at least as large as the total demand for products in M less the upper bound on production of products in M. **P2(d,A)** maximizes this lower bound over all possible product subsets.

If either the number of stages or the number of products is less than three, then the shortfall equals this lower bound. However if both the number of stages and number of products are greater than two, then in general **P2(d,A)** only gives a lower bound on the shortfall. To see this, consider the 3-product 3-stage supply chain depicted in Figure 3. Let every plant have the same capacity *c*. The total capacity at each stage is 2*c*. However, the maximum production that this supply chain can achieve is 3*c*/2. The reason for this is that at each stage there is a pair of products that can be processed at one plant only. In addition this pair of products is different at each stage, being {1,2}, {2,3} and {1,3} for stages 1,2 and 3 respectively. If one produces *x* units of one product, then one can produce at most *c*-*x* of each of the other two products. The maximum production is achieved by producing *c*/2 of each product. Therefore in any demand scenario, every stage of this supply chain has at least one quarter of its total capacity, *c*/2, that is not being used. This is not a well designed supply chain.

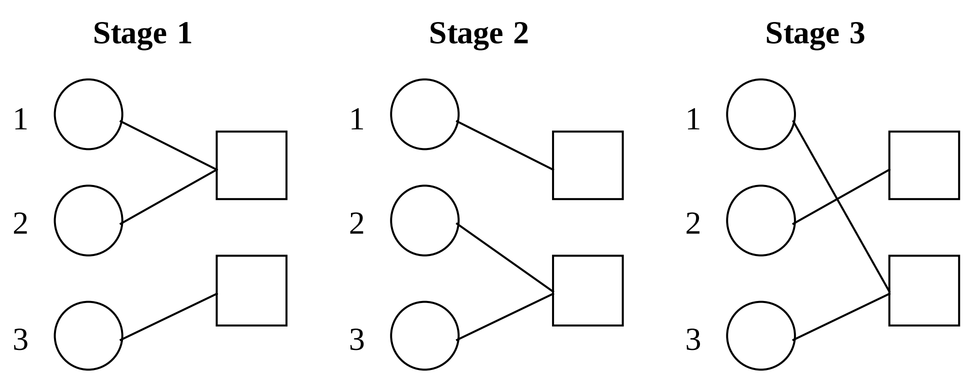


Figure 3

To show that **P2(d,A)** only gives a lower bound and not the actual shortfall for this supply chain, consider a scenario in which each product has a demand of 2*c*/3. In this case, the total demand equals 2*c*, the total capacity of each stage. As the total production is limited by $3c/2$, the shortfall must be at least $c/2$. In fact it is exactly $c/2$ as production of each product can be set equal to $c/2$. In **P2(d,A)**, if $|M|$=0,1, or 3, then the objective value is less than or equal to zero. If $|M|=2$, then the maximum objective function is $c/3$, e.g. $M=\{1,2\}$ $L_1=\{1,2\}$, $L_2=L_3=\{\varnothing\}$. So the optimal value to P2(d,A) is $c/3$ which is strictly less than the actual shortfall of $c/2$.

If either the number of stages or the number if products is less than three, then **P2(d,A)** gives the exact minimum total shortfall. For supply chains in which every stage is totally flexible or in which every stage is totally dedicated, **P2(d,A)** gives the exact minimum total shortfall as such supply chains can be mapped into one or more single product supply chains. In fact, as noted in the proof of Lemma 1, if the dual solution to **P1(d,A)** is integral, then **P2(d,A)** gives the exact minimum total shortfall rather than simply a lower bound.

Tests were run to see how often the dual solutions to **P1(d,A)** were integral in different supply chains. In each supply chain tested, the number of plants at each stage equaled the number of products. Stage capacities were either all equal to 100 or else were randomly generated from a normal distribution with mean 100 and standard deviation of 30 and then rounded to the nearest integer. Each supply chain had either a pairs configuration at each stage or a chain configuration at every stage. The particular pairs (chain) configuration was randomly generated for each stage, so each stage had a different pairs (chain) configuration. Each of the twenty eight supply chains was solved for either 2,500 or 10,000 different demand realizations. Product demands were randomly generated from a normal distribution with mean 100 and standard deviation of

30 and then rounded to the nearest integer. Details of the supply chains tested and the number of demand realizations used for each supply chain can be found in Table 1.

Table 1

Supply Chain	Number of Products	Number of Stages	Stage Configuration	Plant Capacity	Number of Instances
1	10	3	Pairs	100	10,000
2	10	4	Pairs	100	10,000
3	10	5	Pairs	100	10,000
4	10	6	Pairs	100	10,000
5	10	3	Chain	100	10,000
6	10	4	Chain	100	10,000
7	10	5	Chain	100	10,000
8	10	6	Chain	100	10,000
9	10	10	Chain	100	10,000
10	10	15	Chain	100	2,500
11	10	20	Chain	100	2,500
12	10	3	Chain	N(100,30)	10,000
13	10	4	Chain	N(100,30)	10,000
14	10	5	Chain	N(100,30)	10,000
15	10	6	Chain	N(100,30)	10,000
16	10	10	Chain	N(100,30)	10,000
17	10	15	Chain	N(100,30)	2,500
18	10	20	Chain	N(100,30)	2,500
19	20	3	Chain	100	10,000
20	20	4	Chain	100	10,000
21	20	5	Chain	100	10,000
22	20	6	Chain	100	10,000
23	20	10	Chain	100	10,000
24	20	3	Chain	N(100,30)	10,000
25	20	4	Chain	N(100,30)	10,000
26	20	5	Chain	N(100,30)	10,000
27	20	6	Chain	N(100,30)	10,000
28	20	10	Chain	N(100,30)	10,000

A total of 250,000 different instances of **P1(d,A)** were solved. In every single instance, the dual solution was integral. This implies that for these 250,000 instances **P2(d,A)** gives the exact total minimum shortfall and not simply a lower bound.

The supply chain in Figure 3 shows that it is possible to construct supply chains in which **P2(d,A)** gives a strict lower bound. However this supply chain is a rather poorly designed one as it contains capacity at every stage that can never be used in any demand scenario. The supply chains tested above were more reasonable in design. I conjecture that for most reasonably designed supply chains **P2(d,A)** gives the exact minimum total shortfall and not a strict lower bound(1). As noted in Lemma 1, **P2(d,A)** is exact for any supply chain in which either the number of stages or the number of products is less than three.

In this dissertation (Chapter 3), I assume that **P2(d,A)** gives the exact total minimum shortfall and not just a lower bound. As it always gives a lower bound for any supply chain, the analytical results that depend on **P2(d,A)** still have a closely related interpretation if one thinks of **P2(d,A)** as simply giving a lower bound.

Note that all the supply chain simulations in this dissertation (Chapter 3) were solved using **P1(d,A)** and so the shortfalls reported are exact.

3.2.2 Nomenclature

For ease of reference, the nomenclature used in this chapter is presented.

<u>Scalars</u>

c^k_j: The capacity of plant j of stage k

(1) In other words I conjecture that there exist sufficient conditions on the supply chain that guarantee that the dual solution to **P1(d,A)** is integral(2) which in turn implies that **P2(d,A)** gives the exact minimum shortfall. I also conjecture that most reasonable supply chains meet these conditions.

(2) Note that if the number of stages and products both exceed two then the dual problem to **P1(d,A)** does not have an integral polyhedron as the graph **G** in Lemma 1 (b) can be shown to be not perfect. However, the dual objective function, $\mathbf{f_d}$, is not completely arbitrary as the demand for a product is the same at all stages of the supply chain. The actual conjecture is that there exist sufficient conditions that guarantees that no allowable objective function $\mathbf{f_d}$ can be created such that for some dual feasible solution $\mathbf{x}$ the vector $\mathbf{f_d x}$ meets the dual polyhedron only at a fractional extreme point.

CI: The configuration inefficiency (defined in Section 3.5.3)

CFI: The configuration floating inefficiency (defined in Section 3.5.3)

CL: The configuration loss for the supply chain (defined in Section 3.5.1)

CSI: The configuration spanning inefficiency (defined in Section 3.5.3)

d_i: The demand realization for product i

D_i: The demand random variable for product i

g: A measure of the capacity available to each subset of products at a stage (defined in Section 3.6)

$g_{\min}$: The minimum g value over the stages in the supply chain

h: The number of products directly connected to each plant in a complete chain

I: The number of products (i=1,…,I)

J_k: The number of plants at stage k (j=1,…,J_k)

K: The number of stages in the supply chain (k=1,…,K)

s_i: The shortfall realization for product i

S_i: The shortfall random variable for product i

sf: The supply chain shortfall realization

SF: The supply chain shortfall random variable

sf_k: The shortfall realization for stage k as a stand alone stage

SF_k: The shortfall random variable for stage k as a stand alone stage

TC_k: Total capacity of stage k

$TC_{\min}$: The minimum total stage capacity

x^k_{ij}: The production of product i in plant j of stage k

μ_i: The expectation of demand for product i

σ_i: The standard deviation of demand for product i

$\Omega_S(M,L_1,\ldots,L_K)$: An upper bound on the probability that $(M,L_1,\ldots,L_K)$ is a stage-spanning bottleneck

$\Omega_S(I,g_{\min})$: An upper bound on the probability of occurrence of a stage-spanning bottleneck in an I-product $g_{\min}$ supply chain

Γ_K: The probability that the maximum stand-alone stage shortfall of the supply chain exceeds the shortfall of a totally flexible supply chain

$\Gamma_K(M_1^*,\ldots,M_K^*)$: A surrogate measure of the probability that the maximum stand-alone stage shortfall of the supply chain exceeds the shortfall of a totally flexible supply chain

Vectors:

A_k: The flexibility configuration of stage k

A: The flexibility configuration of the supply chain $=\{A_1,\ldots,A_K\}$

d: The demand realization vector $=\{d_1,\ldots,d_I\}$

D: The demand random vector $=\{D_1,\ldots,D_I\}$

TF_k: Denotes a total flexibility configuration for stage k

TF: Denotes a total flexibility configuration for the supply chain

Sets:

$P^k(j)$: The set of products that plant j of stage k can process

$Q^k(i)$: The set of plants at stage k that can process product i

M: A subset of the products $\{1,\ldots,I\}$

L_k: A subset of M, k=1,…,K, such that the L_k form a partition of M

$H(K)$: The set of stages l=1,…,K for which, $E[SF_l] = \underset{k=1,\ldots,K}{\text{Max}}\{E[SF_k]\}$.

3.3 Supply Chain Inefficiencies

The shortfall in a supply chain will always be greater than or equal to the minimum single-stage stand-alone shortfall, where the stand-alone shortfall of a stage is the shortfall that would result if it were a single-stage supply chain. In other words, a supply chain is only as effective at meeting demand as the “weakest link” in the chain.

Supply chains with multiple stages are susceptible to inefficiencies that do not affect single-stage supply chains. By inefficiencies, I mean phenomena that cause the supply chain shortfall to be strictly greater than the maximum stand-alone stage shortfall. So multiple-stage supply chains may actually be less effective at meeting demand than the weakest stage. This dissertation (Chapter 3) analyzes two potential supply chain inefficiencies.

The first, classified as stage-spanning bottlenecks, can occur even when demand is deterministic and constant through time.

The second inefficiency is an example of the floating bottleneck phenomenon noted previously in the literature (e.g. Nahmias, 1997, Hopp and Spearman, 1996) and in practice (Alcalde, 1997). It occurs only when demand stochastic.

When demand is stochastic, the expected or average shortfall is of interest. In such situations, an inefficiency occurs when the expected or average supply chain shortfall is strictly greater than the maximum single-stage stand-alone expected or average shortfall.

3.3.1 Stage-Spanning Bottlenecks

As noted above, for a given demand realization, the shortfall in a multiple-stage supply chain is at least as large as the maximum stand-alone stage shortfall. Using the set based formulation of the production planning problem, i.e. **P2**, to express this mathematically,

$$\underset{M}{\text{Max}}\ \{\sum_{i\in M} d_i - \min_{L_1,..L_K}\{\sum_{k=1}^{K}\sum_{j\in P^k(L^k)} c_j^k\}\} \quad \geq \quad \underset{k=1,\ldots K}{\text{Max}}\ \{\underset{\text{M}}{\text{Max}}\{\sum_{i\in \text{M}} d_i - \sum_{j\in P^k(\text{M})} c_j^k\}\}$$

subject to (left side):

(i) $M \subseteq \{\emptyset, 1, \ldots, I\}$

(ii) $L_k \cap L_{k'} = \emptyset \quad \forall k \neq k'$

(iii) $\bigcup_{k=1}^{K} L_k = M$

subject to (right side):

(i) $M \subseteq \{\emptyset, 1, \ldots, I\}$

(1)

It might be conjectured that this inequality would hold with equality, and indeed for certain instances it does. However, there are also instances in which the inequality is strict. Moreover, for some flexibility configurations, these instances can be quite numerous. For an example of a problem instance when the inequality is strict, consider a two-stage three-product supply chain that is configured as shown in Figure 4.

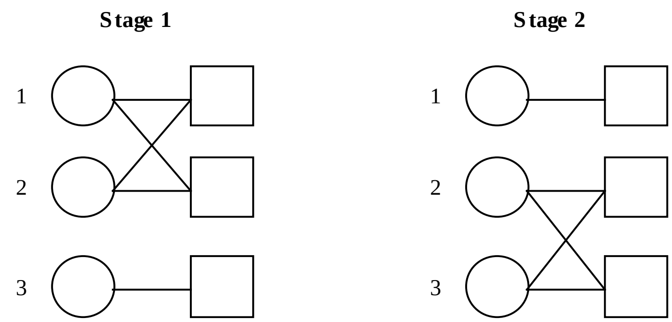


Figure 4

Let the demands for products 1, 2, and 3 be 150, 50, and 150 units respectively and the capacity be 100 units for all plants. The shortfall of either stage on a stand-alone basis is 50 units. For the two-stage supply chain, the optimal product sets for **P2** are $M^*=\{1,3\}$, $L_1^*=\{3\}$ and $L_2^*=\{1\}$ and the supply chain shortfall is 100 units. The bottleneck plants are those in bold in Figure 5 below. The bottleneck plants for stage k are those plants that can process the products in L_k^*, where L_k^*, $k=1,\ldots,K$ are the maximizing sets in **P2**.

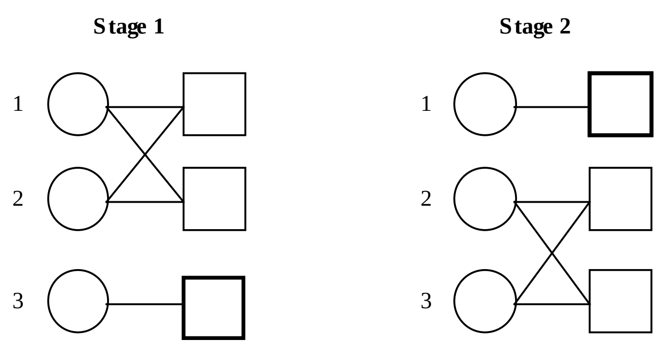


Figure 5

The inequality in (1) is strict only if more than one of the L_k* are non-empty. In other words, the supply chain and product demands are such that there exist two or more disjoint subsets of products, each subset having a positive shortfall. The shortfall of each subset is caused by the capacity constraints at a single stage, where this limiting stage is different for each subset. The products in the various subsets can be thought of as constrained products, that is their demand cannot be met; the plants constraining them are referred to as bottleneck plants. As there is more than one limiting stage, the bottleneck plants exist in more than one stage. Instances when the inequality in (1) is strict are referred to as stage-spanning bottlenecks.

The maximizing sets for **P2** may not be unique. I will focus on the optimal solution(s) with the minimum (out of all possible optimal solutions) number of non-empty L_k*. A stage-spanning bottleneck occurs only if,

$$\begin{array}{lcl}
\underset{M}{\text{Max}} \ \{\sum_{i\in M} d_i - \min_{L_1,..L_K}\{\sum_{k=1}^{K} \sum_{j\in P^k(L^k)} c_j^k\}\} & > & \underset{M}{\text{Max}} \ \{\sum_{i\in M} d_i - \min_{L_1,..L_K}\{\sum_{k=1}^{K} \sum_{j\in P^k(L^k)} c_j^k\}\} \\
\text{subject to} & & \text{subject to} \\
\text{(i) } M \subseteq \{\varnothing,1,\ldots,I\} & & \text{(i) } M \subseteq \{\varnothing,1,\ldots,I\} \\
\text{(ii) } L_k \cap L_{k'} = \varnothing \quad \forall k \neq k' & & \text{(ii) } L_k \cap L_{k'} = \varnothing \quad \forall k \neq k' \\
\text{(iii) } \bigcup_{k=1}^{K} L_k = M & & \text{(iii) } \bigcup_{k=1}^{K} L_k = M \\
\text{(iv) More than one } L_k \text{ non-empty} & & \text{(iv) Only one } L_k \text{ non-empty}
\end{array} \tag{2}$$

For a given demand realization, a stage-spanning bottleneck occurs when the supply chain shortfall is strictly greater than the maximum stand-alone stage shortfall. In other words, the supply chain performs more poorly than its weakest stage. This is an inefficiency that one would rather avoid. It would be helpful to have a supply chain configuration that yielded a low probability of occurrence of stage-spanning bottlenecks. The next section develops a measure of this probability.

3.3.1.1 Probability of Occurrence of a Stage-Spanning Bottleneck

In this section I develop an upper bound on the probability of any particular stage-spanning bottleneck.

If the left hand side of (2) is less than or equal to zero, then (2) cannot hold as $M=\{\emptyset\}$ is a valid solution to the right hand side and has an objective value of zero. Let

$$TC_k = \sum_{j=1}^{J_k} c_j^k \text{, and } TC_{\min} = \min_{k=1,\ldots,K}\{TC_k\}$$

i.e. TC_k is the total capacity of stage k and $TC_{\min}$ is the minimum total stage capacity. Although deterministic quantities, I use upper case letters to distinguish stage from plant capacities. Without loss of generality, assume that the stages are numbered such that $TC_1=TC_{\min}$. A valid solution to the right hand side of (2), i.e. a valid single-stage bottleneck, is $M=\{1,\ldots,I\}$ with only L_1 being non-empty. The objective value for this solution is $\sum_{i=1}^{I} d_i\text{-}TC_{\min}$. If the left-hand side of (2) is less than or equal to this, then (2) cannot hold. Therefore, using the two valid right hand side solutions to (2), a necessary, but not sufficient, condition for a stage-spanning bottleneck to occur is

$$\underset{M}{\text{Max}}\ \{\sum_{i\in M} d_i - \min_{L_1,..L_K}\{\sum_{k=1}^{K}\sum_{j\in P^k(L^k)} c_j^k\}\} \quad > \quad \text{Max}\ \{0,\sum_{i=1}^{I} d_i\text{-}TC_{\min}\}$$

subject to

(i) $M \subseteq \{\emptyset, 1, \ldots, I\}$

(ii) $L_k \cap L_{k'} = \emptyset \quad \forall k \neq k'$

(iii) $\bigcup_{k=1}^{K} L_k = M$

(iv) More than one L_k non-empty

(3)

An alternative explanation of this necessary condition is that for a stage-spanning bottleneck to occur, the left-hand side of (2) must be larger than the maximum of the stand-alone stage shortfalls. As each stand-alone stage shortfall is at least as large as that stage's shortfall under total flexibility, the maximum of the stand-alone stage shortfalls is at least as large as the maximum stand-alone stage shortfall under total flexibility. This is given by the right-hand side of (3).

In what follows, I develop an upper bound on the probability that any particular stage-spanning bottleneck occurs.

Let (M_N,L_1,…,L_K) be any feasible solution to **P2**, the set based shortfall formulation. Denote the stages with non-empty L_k by k_1,…,k_N, where N is the number of non-empty L_k. In a stage-spanning bottleneck, there must be at least two non-empty L_k, i.e. N>1. From (3), a necessary condition for (M,L_1,…,L_K) to be a stage-spanning bottleneck for a given demand realization, d_1,…,d_I, is given by,

$$\sum_{i\in M} d_i - \sum_{n=1}^{N}\sum_{j\in P^{k_n}(L_{k_n})} c_j^{k_n} \quad > \quad \text{Max}\ \{0,\sum_{i=1}^{I} d_i\text{-}TC_{\min}\}$$

As this is only a necessary and not a sufficient condition for (M,L_1,…,L_K) to be a stage-spanning bottleneck,

$$\text{Prob}\left[\sum_{i\in M} D_i - \sum_{n=1}^{N} \sum_{j\in P^{k_n}(L_{k_n})} c_j^{k_n} \quad > \quad \text{Max } \{0, \sum_{i=1}^{I} D_i\text{-}TC_{\min}\}\right] \tag{4}$$

is an upper bound on the probability that $(M,L_1,\ldots,L_K)$ is a stage-spanning bottleneck.

Denote this upper bound by, $\Omega_S(M,L_1,\ldots,L_K)$. Let,

$$A_N = \sum_{i\in M} D_i - \sum_{n=1}^{N} \sum_{j\in P^{k_n}(L_{k_n})} c_j^{k_n} \text{ and } B = \sum_{i=1}^{I} D_i\text{-}TC_{\min} - A_n$$

where as noted above, $k_1,\ldots,k_N$, denote the non-empty L_k. Then,

$$\begin{aligned}\Omega_S(M,L_1,\ldots,L_K) &= \text{Prob}\left[\sum_{i\in M} D_i - \sum_{n=1}^{N} \sum_{j\in P^{k}(L_{k_n})} c_j^{k_n} \quad > \quad \text{Max } \{0, \sum_{i=1}^{I} D_i\text{-}TC_{\min}\}\right] \\ &= \text{Prob}\left[A_N > \text{Max}\{0, A_N + B_N\}\right] \\ &= \text{Prob}\left[0 > \text{Max}\{0, A_N + B_N\} - A_N\right] \\ &= \text{Prob}\left[0 > \text{Max}\{-A_N, B_N\}\right] \\ &= \text{Prob}\left[A_N > 0 \text{ and } B_N < 0\right]\end{aligned}$$

If the demands are independent and normally distributed with $D_i \sim N(\mu_i, \sigma_i)$, then,

$$A_N \sim N\left(\sum_{i\in M} \mu_i - \sum_{n=1}^{N} \sum_{P^{k_n}(L_{k_n})} c_j^{k_n}, \sqrt{\sum_{i\in M} \sigma_i}\right)$$

$$B_N \sim N\left(\sum_{i\notin M} \mu_i - TC_{\min} + \sum_{n=1}^{N} \sum_{j\in P^{k_n}(L_{k_n})} c_j^{k_n}, \sqrt{\sum_{i\notin M} \sigma_i}\right)$$

with A_N and B_N being independent. Therefore,

$$\Omega_S(M,L_1,\ldots,L_K) = \text{Prob}\left[A_N > 0\right]\text{Prob}\left[B_N < 0\right] = \left[1 - \Phi(z_1)\right]\Phi(z_2)$$

with $z_1 = -\mu_{A_N} / \sigma_{A_N}$ and $z_2 = -\mu_B / \sigma_{B_N}$. To recap, $\Omega_S(M,L_1,\ldots,L_K)$ is an upper bound on the probability that $(M,L_1,\ldots,L_K)$ is a stage-spanning bottleneck.

If this upper bound is very small for all possible $(M,L_1,\ldots,L_K)$, in other words the probability of occurrence of any particular stage stage-spanning bottleneck is very small, then I conjecture that the probability of occurrence of any of the possible stage-spanning bottleneck is also small.

It is possible to classify stage-spanning bottlenecks by the number, N, of non-empty L_k, i.e. the number of stages that contain bottleneck plants. The upper bound, $\Omega_S(M,L_1,\ldots,L_K)$, depends on N, and so by setting the value of N in $\Omega_S(M,L_1,\ldots,L_K)$ equal to 3 for example, one obtains an upper bound on the probability of a particular 3-stage-spanning bottleneck.

Note that $\text{Max } \{0,\sum_{i=1}^{I} d_i\text{-}TC_{\min}\}$ is the supply chain shortfall under total flexibility. Therefore, from (4), $\Omega_S(M,L_1,\ldots,L_K)$ is the probability that the shortfall of a particular stage-spanning bottleneck exceeds the total flexibility shortfall. The above development is valid even if there is only one stage in the supply chain. Of course in this case, by definition there can be only one non-empty L_k, that is stage 1, and $M=L_1$. $\Omega_S(M)$ can be interpreted as a measure of the probability that the shortfall of the single-stage supply chain exceeds the total flexibility shortfall. In fact, for single-stage supply chains, $\Omega_S(M)=\Pi(M)$, the single-stage measure of Jordan and Graves (1995).

In a sense, $\Omega_S(M,L_1,\ldots,L_K)$ is one possible generalization of the single-stage $\Pi(M)$ measure for multiple-stage supply chains. For single-stage supply chains, $\Pi(M)$ can be interpreted "as the probability of having unfilled demand for the set of products M, while simultaneously having excess capacity in plants that don't build any of the products in M" (Jordan and Graves, 1995). In multiple stage supply chain, $\Omega_S(M,L_1,\ldots,L_K)$ can be interpreted as an upper bound on the probability of $(M, L_1, \ldots, L_K)$ being a stage-spanning bottleneck, when more than one of the L_k are non-empty. Alternatively it can be interpreted as the probability of having unfilled demand for the set of products *M*, due to production of the non-empty subsets, L_k, of *M* being limited by the plant capacities and capabilities at those stages with non-empty subsets, while simultaneously having spare capacity in some plant(s) at every stage.

For multiple-stage supply chains that have low probabilities of stage-spanning bottlenecks, a measure of the probability that a multiple-stage supply chain shortfall exceeds the total flexibility shortfall is developed in Section 3.4.2. This will be a second generalization of $\Pi(M)$ measure, with this one being an equivalent measure, but generalized to multiple-stage supply chains.

3.3.2 Floating Bottlenecks

Even if stage-spanning bottlenecks do not occur in the supply chain for any demand realization, the expected supply chain shortfall is not necessarily equal to the maximum expected stand-alone stage shortfall. Indeed, for most configurations the expected supply chain shortfall will be strictly greater than the maximum expected stand-alone stage shortfall. This is the second supply chain inefficiency, and is termed the floating bottleneck inefficiency.

If stage-spanning bottlenecks do not occur, then for a given demand realization, the supply chain shortfall, *sf*, is given by $sf = \underset{k=1,\ldots,K}{\mathrm{Max}} \{ sf_k \}$ where sf_k is the stand-alone shortfall for stage *k*. In this case,

$$E[SF] = E\left[\underset{k=1,\ldots,K}{\text{Max}}\{SF_k\}\right] \geq \underset{k=1,\ldots,K}{\text{Max}}\{E[SF_k]\}.$$

with the inequality being strict for numerous supply chain configurations. Let $H(K)$ denote the set of stages l=1,…,K for which, $E[SF_l] = \underset{k=1,\ldots,K}{\text{Max}}\{E[SF_k]\}$. Now,

$$E\left[\underset{k=1,\ldots,K}{\text{Max}}\{SF_k\}\right] = \underset{k=1,\ldots,K}{\text{Max}}\{E[SF_k]\} \tag{5}$$

if and only if, for all demand realizations, $\underset{k=1,\ldots,K}{\text{Max}}\{sf_k\} \leq sf_l \ \forall l \in H(K)$, or equivalently,

$$sf_{l_i} = sf_{l_j} \ \forall l_i, l_j \in H(K) \text{ and } sf_k \leq sf_l \ \forall l \in H(K),\ k \notin H(K)$$

To see this, consider a case where $H(K)=\{1\}$, i.e. the first stage is such that $E[SF_l] = \underset{k=1,\ldots,K}{\text{Max}}\{E[SF_k]\}$ and all other stages are such that $E[SF_l] < \underset{k=1,\ldots,K}{\text{Max}}\{E[SF_k]\}$. If for some demand realization, the stand-alone shortfall of some stage is strictly greater than that of stage 1, i.e. $sf_k > sf_1$, $k \neq 1$, then $\underset{k=1,\ldots,K}{\text{Max}}\{sf_k\} > sf_1$ for this demand realization and

$$E\left[\underset{k=1,\ldots,K}{\text{Max}}\{SF_k\}\right] > E[SF_1]$$

Thus, the expected supply chain shortfall will be strictly greater than the maximum expected stand-alone stage shortfall. This phenomenon is referred to as the floating bottleneck inefficiency because it arises when the maximum stand-alone stage shortfall occurs in different stages for different demand realizations. If the maximum stand-alone shortfall always occurred in the same stage, say stage 1, then the expected supply chain shortfall will be equal to the maximum expected stand-alone stage shortfall, assuming stage-spanning bottlenecks do not occur.

Therefore supply chains in which,

$$\text{Prob}\left[SF_{l_i} = SF_{l_j}\ \forall l_i, l_j \in H(K) \text{ and } SF_k \le SF_l \forall l \in H(K), k \notin H(K)\right] \tag{6}$$

is not close to one should be prone to the floating bottleneck inefficiency. Equation (6) is most easily interpreted in the case where $H(K)=\{1\}$. In this case, (6) is the probability that the stand-alone shortfall for all other stages is less than or equal to that of stage 1. If this probability is far less than one, then there are many demand realizations for which the stand-alone stage shortfall of other stages is strictly greater than that of stage 1, that is the bottleneck stage “floats” from one stage to another. If, on the other hand, the probability in equation (6) is close to one, then in most demand realizations, stage 1 is the bottleneck stage. It is therefore reasonable to conjecture that

$$E\left[\underset{k=1,\dots,K}{\text{Max}}\{SF_k\}\right] \approx \underset{k=1,\dots,K}{\text{Max}}\{E[SF_k]\} = E[SF_1] \tag{7}$$

as for most demand realizations $\underset{k=1,\dots,K}{\text{Max}}\{sf_k\} = sf_1$.

It should be noted that having the probability in (6) be close to one does not guarantee that (7) be true as there may be some demand realization for which $\underset{k=1,\dots,K}{\text{Max}}\{sf_k\} >> sf_1$, i.e. the maximum stand-alone stage shortfall is much greater than that of stage 1. However, I will use the probability in equation (6) as a measure of the protection a supply chain configuration provides against the floating bottleneck inefficiency. The closer the probability is to one, the higher the protection. Supply chains in which this probability is high should have a fairly stable and predictable bottleneck stage.

If stage-spanning bottlenecks can occur in the supply chain, then

$$E[SF] > E\left[\underset{k=1,\ldots,K}{\text{Max}}\{SF_k\}\right] \geq \underset{k=1,\ldots,K}{\text{Max}}\{E[SF_k]\}$$

However, as $E\left[\underset{k=1,\ldots,K}{\text{Max}}\{SF_k\}\right]$ still provides a lower bound on the expected supply chain shortfall, one does not want it to be overly large. Therefore, as discussed above, one would like the probability in (6) to be close to one.

Supply chain configurations for which this probability is high and for which the probability of stage-spanning bottlenecks is low, do not suffer significantly from either the floating bottleneck or stage-spanning bottleneck inefficiencies. Such supply chain configurations should perform well, where performance is measured by the expected shortfall, as long as the expected shortfall of the predictable bottleneck stage is small. Of course even if there are no supply chain inefficiencies, if the predictable bottleneck stage's expected shortfall is large, than the supply chain shortfall will be large. More details on performance measures are provided in the next section.

3.4 Performance Measurement

Up to this point, I have been somewhat vague about the performance measure of a configuration, stating simply that a good configuration should result in a low expected supply chain shortfall. What is a low? Different problem instances will result in different shortfalls and so it would be useful to have a consistent measure across problem instances. The next section introduces such a measure.

3.4.1 Configuration Loss

For any given problem instance, a totally flexible supply chain, i.e. one in which every stage has a totally flexible configuration, provides a lower bound on the expected supply chain shortfall for any possible configuration. As such it provides a good baseline against which to measure the performance of other configurations. The following performance measure is used for a configuration **A**,

$$100 \times \left(\frac{\mathrm{E}\left[SF(\mathbf{D},\mathbf{A})\right] - \mathrm{E}\left[SF(\mathbf{D},TC)\right]}{\mathrm{E}\left[SF(\mathbf{D},TC)\right]} \right)$$

where, as defined earlier, E[SF(**D,A**)] is the expected shortfall for a supply chain with configuration **A** and E[SF(**D,**TC)] is the expected shortfall for a totally flexible supply chain.

This measure is defined as the "configuration loss" (or CL) for configuration **A**. It is a measure of the increase in expected shortfall caused by partial flexibility. A configuration loss close to 0% indicates a configuration that performs very well as it approaches the performance of a totally flexible supply chain. The higher the configuration loss the less well the configuration is said to perform. A configuration loss of 100% indicates that the expected shortfall is twice that of total flexibility. The configuration loss can and will be seen to exceed 100% for some configurations.

While a closed form expression, albeit containing the normal cumulative distribution function, for the expected supply chain shortfall under total flexibility can be developed if demands are normally distributed, a closed form expression for E[SF(**D,A**)] does not exist for most configurations. Therefore, one approach to evaluating the performance of configurations is to use Monte Carlo simulation.

For a given supply chain configuration, the product demand vector **d** is randomly generated from the distribution for **D**. Using this demand realization, the shortfall is determined by solving the production planning linear program **P1**. An estimate of E[*SF*(**D,A**)] is then obtained by repeating this process for numerous demand realizations.

In the results presented in this dissertation (Chapter 3), product demands are Normal, $N(\mu,\sigma)$, truncated at μ+/-2σ, as in Jordan and Graves (1995) and Gavish (1994). Demands are identically and independently distributed, with a mean of 100 and a standard deviation of 30, unless otherwise stated. In all cases, the truncation prevented negative demands.

For supply chain simulations in which the number of products was less than or equal to 20, 10,000 demand realizations were used to generate the estimates for the expected shortfall values. The 95% confidence intervals for the expected shortfall estimates were calculated and found to be within +/-3% of the estimates (Law and Kelton, Chapter 4.5). This is comparable to the confidence intervals in Gavish (1994). For supply chains in which the number of products was 30 or 40, the solution time for each demand realization was quite long. Because of this, only 1000 demand realizations were used. In this case the 95% confidence intervals were within +/-9% of the estimates.

3.4.2 Probability that Supply Chain Shortfall Exceeds Total Flexibility Shortfall

The configuration loss should be low if the probability that the supply chain shortfall exceeds that of totally flexible supply chain is low. This provides another measure of supply chain performance. As a closed form expression for this probability does not exist, an alternative measure of this probability is developed.

Let *sf*(**d**,*TF*) denote the supply chain shortfall for a given demand realization, **d**, assuming total flexibility. This is given by,

$$sf(\mathbf{d},TF) = \underset{\mathrm{k=1,\ldots,K}}{\mathrm{Max}}\{\mathrm{Max}\{0,\sum_{i=1}^{I} d_i - \sum_{j=1}^{J_k} c_j^k\}\}$$

For a supply chain with K stages, let Γ_K be the probability that the maximum stand-alone stage shortfall exceeds the supply chain total flexibility shortfall. Γ_K is then given by the following expression,

$$\begin{aligned}\Gamma_K &= \mathrm{Prob}\left[\underset{\mathrm{k=1,\ldots,K}}{\mathrm{Max}}\{\underset{M}{\mathrm{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\}\} > SF(\mathbf{D},TF)\right] \\ &= 1 - \mathrm{Prob}\left[\underset{\mathrm{k=1,\ldots,K}}{\mathrm{Max}}\{\underset{M}{\mathrm{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\}\} \le SF(\mathbf{D},TF)\right] \\ &= 1 - \mathrm{Prob}\left[\underset{M}{\mathrm{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} \le SF(\mathbf{D},TF)\ \ \forall k = 1,\ldots,K\right]\end{aligned} \quad (8)$$

Note that the demand and shortfall random variables are denoted by capital letters.

Under the assumption that the internal probabilities (that a stand-alone stage shortfall is less than or equal to the supply chain total flexibility shortfall) are positively correlated across the different stages, then

$$\text{Prob}\left[\underset{M}{\text{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} \le SF(\mathbf{D},TF)\ \ \forall k=1,\ldots,K\right]$$
$$\ge \prod_{k=1}^{K}\text{Prob}\left[\underset{M}{\text{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} \le SF(\mathbf{D},TF)\right] \tag{9}$$

While this assumption is not proven, simulation indicates it to be valid.

Denote $sf_k(\mathbf{d},TF_k)$ as the stand alone shortfall of stage k for a given demand realization, $\mathbf{d}$, when stage k is totally flexible. The supply chain shortfall under total flexibility, $sf(\mathbf{d},TF)$, will be greater than or equal to the stand-alone shortfall of any stage under total flexibility, $sf_k(\mathbf{d},TF_k)$. Therefore, using this and equation (9),

$$\text{Prob}\left[\underset{M}{\text{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} \le SF(\mathbf{D},TF)\ \ \forall k=1,\ldots,K\right]$$
$$\ge \prod_{k=1}^{K}\text{Prob}\left[\underset{M}{\text{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} \le SF_k(\mathbf{D},TF_k)\right] \tag{10}$$

Substituting equation (10) into equation (8), an upper-bound on Γ_K is given by,

$$\Gamma_K \le 1-\prod_{k=1}^{K}\text{Prob}\left[\underset{M}{\text{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} \le SF_k(\mathbf{D},TF_k)\right]$$
$$=1-\prod_{k=1}^{K}\left(1-\text{Prob}\left[\underset{M}{\text{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} > SF_k(\mathbf{D},TF_k)\right]\right)$$

As in Jordan and Graves(1995), a closed form expression does not exist for

$$\text{Prob}\left[\underset{M}{\text{Max}}\ \{\sum_{i\in M} D_i - \sum_{j\in P^k(M)} c_j^k\} > SF_k(\mathbf{D},TF_k)\right]$$

The authors instead use $\Pi(M^*)$ as a measure for this probability, where

$$\Pi(M) = \text{Prob}\left[\sum_{i \in M} D_i - \sum_{j \in P^k(M)} c_j^k > SF_k(\mathbf{D}, TF_k)\right]$$

and $\Pi(M^*)$ is the maximal such probability for all possible subsets, M.

Likewise, I also use $\Pi(M^*)$ to obtain a measure for this upper bound on Γ_K, the probability that the maximum stand-alone stage shortfall exceeds that of the supply chain total flexibility shortfall. Denote this upper bound by $\Gamma_K(M_1^*,\ldots,M_K^*)$. Then

$$\Gamma_K(M_1^*,\ldots,M_K^*) = 1 - \prod_{k=1}^{K}\left(1 - \Pi_k(M_K^*)\right)$$

If the probability of stage-spanning bottlenecks is low, then for most demand realizations, the supply chain shortfall will equal the maximum stand-alone stage shortfall. As $\Gamma_K(M_1^*,\ldots,M_K^*)$ is a measure of the probability that the maximum stand-alone stage shortfall exceeds the supply chain shortfall under total flexibility, it can be used as a measure of the probability that a K-stage supply chain shortfall exceeds the K-stage supply chain shortfall under total flexibility (in supply chains where the probability of stage-spanning bottlenecks is low). $\Gamma_K(M_1^*,\ldots,M_K^*)$ is a generalization of the $\Pi(M^*)$ measure of Jordan and Graves to multiple-stage supply chains. It has a similar interpretation to the single-stage measure $\Pi(M^*)$; it is a measure of the probability that the supply chain shortfall exceeds the supply chain shortfall under total flexibility, assuming that the probability of stage-spanning bottlenecks is low.

Monte-Carlo simulation is used to estimate the probability of the supply chain shortfall exceeding the total flexibility shortfall.

3.4.3 Configuration Inefficiency

In Section 3.3, it was argued that supply chains that were susceptible to the stage-spanning bottleneck and floating bottleneck inefficiencies would not perform well. A supply chain configuration $\mathbf{A}=\{A_1,\ldots,A_K\}$was said to be inefficient if,

$$\mathrm{E}[SF(\mathbf{D},\mathbf{A})] > \underset{k=1,\ldots,K}{\mathrm{Max}}\{\mathrm{E}[SF_k(\mathbf{D},A_k)]\}$$

That is the expected supply chain shortfall is strictly greater than the maximum expected stand alone stage shortfall. A measure of the supply chain inefficiency caused by the configuration **A**, is then given by,

$$100\times\left(\frac{\mathrm{E}[SF(\mathbf{D},\mathbf{A})] - \underset{k=1,\ldots,K}{\mathrm{Max}}\{\mathrm{E}[SF_k(\mathbf{D},A_k)]\}}{\underset{k=1,\ldots,K}{\mathrm{Max}}\{\mathrm{E}[SF_k(\mathbf{D},A_k)]\}}\right)$$

This measure is called the “configuration inefficiency” (or CI).

As defined, floating bottlenecks occur iff,

$$E\left[\underset{k=1,\ldots,K}{\mathrm{Max}}\{SF_k(\mathbf{D},A_k)\}\right] > \underset{k=1,\ldots,K}{\mathrm{Max}}\{\mathrm{E}[SF_k(\mathbf{D},A_k)]\}$$

A measure of the inefficiency contribution made by floating bottlenecks is then given by,

$$100\times\left(\frac{E\left[\underset{k=1,\ldots,K}{\mathrm{Max}}\{SF_k(\mathbf{D},A_k)\}\right] - \underset{k=1,\ldots,K}{\mathrm{Max}}\{\mathrm{E}[SF_k(\mathbf{D},A_k)]\}}{\underset{k=1,\ldots,K}{\mathrm{Max}}\{\mathrm{E}[SF_k(\mathbf{D},A_k)]\}}\right)$$

This measure is called the “configuration floating inefficiency” (of CFI).

As defined, stage-spanning bottlenecks iff,

$$\mathrm{E}[SF(\mathbf{D},\mathbf{A})] > E\left[\underset{k=1,\ldots,K}{\mathrm{Max}}\{SF_k(\mathbf{D},A_k)\}\right]$$

A measure of the inefficiency contribution made by stage-spanning bottlenecks is then given by,

$$100 \times \left(\frac{\mathrm{E}[SF(\mathbf{D},\mathbf{A})] - E\left[\max_{k=1,\ldots,K} \{SF_k(\mathbf{D},A_k)\}\right]}{E\left[\max_{k=1,\ldots,K} \{SF_k(\mathbf{D},A_k)\}\right]} \right)$$

This measure is called the "configuration spanning inefficiency" (or CSI). Note that,

$$CI = CFI + \left(\frac{E\left[\max_{k=1,\ldots,K} \{SF_k(\mathbf{D},A_k)\}\right]}{\max_{k=1,\ldots,K} \{\mathrm{E}[SF_k(\mathbf{D},A_k)]\}} \right) CSI$$

As for the configuration loss, closed form expressions for these inefficiencies do not exist. The same Monte Carlo simulation used to estimate the configuration loss is also used to estimate these inefficiencies. Estimates of the probability of stage-spanning bottlenecks and floating bottlenecks are also obtained.

A supply chain configuration will not perform well if the configuration inefficiency is large. However, a supply chain might have an inefficiency of zero and still have a large configuration loss, because the configuration is bounded below by the maximum stand-alone stage configuration loss.

A supply chain in which every stage is an exact replica of all other stages, will act like a single-stage supply chain. As such, it will have a configuration inefficiency of zero. A supply chain in which every stage has completely dedicated plants will have a configuration inefficiency of zero, but it will still not perform well.

Therefore in designing a good configuration, care must be taken to ensure that each stage has a configuration that results in a low stand-alone stage configuration loss and also that the stage configurations together give good protection against the stage-spanning bottleneck and floating bottleneck inefficiencies.

3.5 A General Class of Configurations

A totally flexible configuration is one in which every product can be processed in every plant. No other configuration can deliver an expected shortfall less than total flexibility. Stage-spanning bottlenecks cannot occur in a totally flexible configuration, as any bottleneck will comprise all the plants of some stage, namely the stage with minimum total capacity. Floating bottlenecks will also not occur, the bottleneck stage, if there is any, will always be the minimum total capacity stage (this may not be the case if capacity usage varies among products).

Total flexibility for a stage is an extreme configuration in which every non-empty subset of products has the total stage capacity available to it, whereby availability I mean all the stage capacity is capable of processing this subset. This configuration performs well because high demand products can avail of capacity that low demand products are not using. In this section I analyze a general class of flexibility configurations whose members will be distinguished by a lower bound on the fraction of total capacity available to any subset of products. This class is referred to as g-type configurations.

TC_k has been defined earlier as the total capacity of stage k, with $TC_{\min}$ being the minimum total stage capacity. Let $C_k=TC_k/I$, i.e. the total stage capacity divided by the number of products, and $C_{\min}$ be the minimum of C_k, $k=1,\ldots,K$. As before for TC_k, I use upper case letters for C_k to distinguish it from a plant capacity, even though it is deterministic.

The total production of any subset of products is clearly limited by the total capacity available to that subset, where available capacity comprises the capacity of any plant capable of processing a product in the subset. The larger the capacity available to a subset of products, the more likely that the demand for that subset can be met. Of course, the demands of the other products will also play a role in determining how much of this total available capacity is actually allocated to the subset of products.

I now introduce a general classification of flexibility configurations based on the fraction of total stage capacity, TC_k, available to any subset of products. A supply chain stage, k, is said to have a g-type configuration if and only if

$$\sum_{j \in P^k(L_k)} c_j^k \geq TC_k \min\left\{1, \frac{(|L_k| + g - 1)}{I}\right\} \tag{11}$$

where L_k is any non-empty subset of the products $\{1,\ldots,I\}$ and $1 \leq g \leq I$. A stage is defined by the maximum g for which (11) holds.

The left hand side of equation (11) is the total capacity that could possibly be used to process the subset, L_k, of products, i.e. the total available capacity for that subset. The first term on the right hand side is the total stage capacity. The second term is the fraction of this stage capacity that can be used to process products in the subset L_k. This fraction must be less than or equal to one, hence the minimum in the equation. If g is equal to 1, then the fraction of total stage capacity available to a subset of n products is at least (n/I). As g increases, the fraction of total stage capacity available to a subset of n products increases. If g is equal to I, then the fraction is equal to one for all subsets, that is the total stage capacity could be used to process any subset of products; this occurs only in the case of total flexibility.

The class of g-type configurations is quite general. In fact a slight modification to the definition, whereby the $(|L_k|+g-1)$ is replaced by $(\max\{0, (|L_k|+g-1\})$ and $2-I \leq g \leq I$, generates a class that includes all possible stage configurations. This generalization allows for configurations where the fraction of total stage capacity available to a subset n of products is less than (n/I), as the fraction available to any strict subset of products can range from 0 to 1. The lemmas and analysis in this section would all follow through with a slight modification if this more general definition were used. I restrict the definition to (11), but both definitions are equivalent for $g \geq 1$.

Even with the restriction that $g \geq 1$, the configurations allowed are quite numerous. For example, any configuration in which there are I equal capacity plants, with plant i able to process product i (and possibly other products), $i=1,\ldots,I$, is a $g=1$-type configuration. Therefore, equal capacity pair configurations are $g=1$ type, as are any equal capacity n-tuple configurations. The fact that $g=1$ type configurations are so general would suggest that they might not possess very special properties. This will be seen later with reference to stage-spanning bottlenecks.

As the g value increases, the requirement placed on potential configurations becomes quite restrictive. If the plant capacities are equal, a complete h-type chain configuration, defined in Section 3.1, has a g-value equal to h. To see this for $h=2$, remember that any subset, M, of products in a complete h=2 chain configuration has at least one product connected to a plant $j \notin M$ (where the set M also refers to the plants i=1,…,I in the chain).

A supply chain is said to be $g_{\min}$-type, if all stages, k=1,…,K, have g-type configurations with $g_k \geq g_{\min}$. Note that there is no restriction on the specific configuration at each stage other than it be some instance of a g-type, with $g_k \geq g_{\min}$.

Configurations with increasing g-values will be shown to provide better protection against stage-spanning bottlenecks.

3.5.1 Stage-Spanning Bottlenecks in *g*-type Configurations

3.5.1.1 Existence of Stage-Spanning Bottlenecks

Some configurations have the very special property of preventing stage-spanning bottlenecks for all possible demand realizations. In other words, for any demand realization the bottleneck plants are limited to a single stage, although the particular stage may vary from one realization to another.

Lemma 2:

(i) For any subset of products, M, define the problem ***P3(M)*** *as*

$$\underset{L_1,\ldots,L_k}{\text{Min}} \{\sum_{k=1}^{K} \sum_{j \in P^k(L^k)} c_j^k\}$$

subject to

(i) $L_k \cap L_{k'} = \varnothing \quad \forall k \neq k'$

(ii) $\bigcup_{k=1}^{K} L_k = M$

If for every possible M, there exists an optimal solution to ***P3(M)*** *with only one non-empty L_k*, then a stage-spanning bottleneck can never occur.*

(ii) If $A_{\min} \geq \frac{TC_{\max}}{2}$*, where* $A_{\min} = \underset{\substack{i=1,\ldots,I \\ k=1,\ldots,K}}{\text{Min}} \{ \sum_{j \in P^k(i)} c_j^k \}$ *and* $TC_{\max} = \underset{k=1,\ldots,K}{\text{Max}} \{TC_k\}$ *,then a stage-spanning bottleneck can never occur. Note that A_{min} is the minimum total capacity available to any product at any stage and TC_{max} is the maximum total stage capacity across all stages.*

While part (ii) of this lemma places no restrictions on the type of configurations at each stage, the condition is still very restrictive, requiring every product to have at least half of the minimum total stage capacity available to it at all stages. By placing limitations on the flexibility configurations allowed, the restriction on the capacities available to individual products can be significantly relaxed.

Lemma 3

If a supply chain is a $g_{\min}$-type, then a stage-spanning bottleneck can never occur if the total number of products, I, is less than or equal to $2g_{min}$. Furthermore, if at each stage each individual product has the same total capacity available to it, then a stage-spanning bottleneck can never occur if the total number of products, I, is less than or equal to $2(g_{min}+1)$

Note that the requirement in (ii) is not that the same capacity is available at all stages for all products, but rather that the same capacity is available at all stages for each product, this capacity can vary from product to product. It should also be noted that this lemma does not state that stage-spanning bottlenecks occur for larger numbers of products, but rather provides a lower bound on the number of products that guarantees that stage-spanning bottlenecks will not occur. This lower bound is increasing in $g_{\min}$, suggesting that *g*-type configurations provide better protection from stage-spanning bottlenecks as the *g*-value increases.

While some configurations may allow stage-spanning bottlenecks for certain demand realizations, the probability of such bottlenecks may be very low. As such, these configurations also provide good protection. In the next section, I investigate the probability of stage-spanning bottlenecks in $g_{\min}$-type supply chains.

3.5.1.2 Probability of Stage-Spanning Bottlenecks

In Section 3.3.1.1, $\Omega_S(M,L_1,\ldots,L_K)$, an upper bound on the probability that $(M,L_1,\ldots,L_K)$ is a stage-spanning bottleneck was developed.

$$\Omega_S(M,L_1,\ldots,L_K) = [1 - \Phi(z_1)]\Phi(z_2)$$

where

$$z_1 = \frac{\sum_{n=1}^{N} \sum_{j \in P^{k_n}(L_{k_n})} c_j^{k_n} - \sum_{i \in M} \mu_i}{\sqrt{\sum_{i \in M} \sigma_i}} \text{ and } z_2 = \frac{TC_{\min} - \sum_{n=1}^{N} \sum_{P^{k_n}(L_{k_n})} c_j^{k_n} - \sum_{i \notin M} \mu_i}{\sqrt{\sum_{i \notin M} \sigma_i}}$$

and the N stages with non-empty L_k are denoted by $k_1,\ldots,k_N$. Note that N depends on $(M,L_1,\ldots,L_K)$.

As noted in Section 3.3.1.1, if this upper bound is very small for all possible $(M,L_1,\ldots,L_K)$, in other words the probability of occurrence of any particular stage stage-spanning bottleneck is very small, then I conjecture that the probability of occurrence of any of the possible stage-spanning bottleneck is also small.

In this section I restrict attention to supply chains in which each stage has a g-value of at least $g_{\min}$ and in which each stage has a total stage capacity of at least the total expected demand.

$\Omega_S(M,L_1,\ldots,L_K)$ is an upper bound on the probability that a particular $(M,L_1,\ldots,L_K)$ is a stage-spanning bottleneck, and as such depends on $(M,L_1,\ldots,L_K)$. Lemma 4, below, introduces an upper bound on the probability of any stage-spanning bottleneck, $\Omega_S(I,g_{\min})$, that is independent of the particular $(M,L_1,\ldots,L_K)$. That is,

$$\Omega_S(M,L_1,\ldots,L_K) \le \Omega_S(I,g_{\min}) \quad \forall (M,L_1,\ldots,L_K)$$

This upper bound depends on the number of products processed and the supply chain $g_{\min}$ value. It does not depend on the number of stages in the supply chain. By showing that $\Omega_S(I,g_{\min})$ is very small for $g_{\min}$-values greater than or equal to 2, I will demonstrate that $\Omega_S(M,L_1,\ldots,L_K)$ is very small for all possible $(M,L_1,\ldots,L_K)$ if $g_{\min}$ is at least 2. Given this, I conjecture that stage-spanning bottlenecks are rare in supply chains in which $g_{\min}$ is greater than or equal to 2 and the total stage capacities are at least equal to the total expected demand. That is, such supply chains, offer good protection against the stage-spanning inefficiency.

Lemma 4:

If a supply chain is a $g_{\min}$*-type and has the following properties,*

(i) each stage has a total capacity of at least the total expected demand

(ii) the demands for the I products are independent and identically distributed $N(\mu,\sigma)$

then the probability of any particular stage spanning bottleneck is bounded above by $\Omega_S(I,g_{min})$*, where,*

$$\Omega_S(I,g_{\min}) = \Phi\left(\frac{-2(g_{\min}-1)\mu}{\sigma\sqrt{I/2}}\right)^2 \tag{12}$$

Note that this bound does not depend on the number of stages in the supply chain. For a given number of products, I, $\Omega_S(I,g_{min})$ decreases significantly as g_{min} increases. Table 2 below shows the value of $\Omega_S(I,g_{min})$ as the number of products, I, increases for the case of product demands being $N(100,30)$ and g_{min}=1, 2, 3 and 4.

Table 2

	$\Omega_S(I,g_{\min})$			
I	$g_{\min}=1$	$g_{\min}=2$	$g_{\min}=3$	$g_{\min}=4$
10	0.25	2.06E-06	1.55E-18	0.00E+00
15	0.25	5.56E-05	3.17E-13	2.01E-26
20	0.25	3.07E-04	1.54E-10	1.63E-20
25	0.25	8.80E-04	6.60E-09	5.98E-17
30	0.25	1.81E-03	8.30E-08	1.47E-14
35	0.25	3.08E-03	5.16E-07	7.63E-13
40	0.25	4.63E-03	2.06E-06	1.50E-11

If g_{min}=1, then $\Omega_S(I,g_{min}=1)=0.25$, for all I, which is quite large. This suggests that supply chains for which g_{min}=1 may not offer good protection against stage-spanning bottlenecks. Some caution should be applied to this conjecture as a large upper bound does not necessarily imply a large probability. However, in Section 3.6.1 a certain g_{min}=1 type configuration, namely a pairs configuration, is shown to have a large probability of stage-spanning bottlenecks.

As can be seen from Table 2, $\Omega_S(I, g_{min}=2)$ is practically negligible for I less than 20. However as the number of products increases to 20 and beyond, $\Omega_S(I, g_{min}=2)$ stops being negligible, although it remains small. This suggests that supply chains with g_{min}=2 should give reasonable protection against the occurrence of stage-spanning bottlenecks unless the number of products is very large.

As can be seen, $\Omega_S(I, g_{min}$=3 or 4) is practically negligible even for I=40. This suggests that supply chains with $g_{min} \geq 3$ should give very good protection against the occurrence of stage-spanning bottlenecks even if the number of products is very large.

3.5.2 Probability of Shortfall Exceeding Total Flexibility in g_{min}-type Supply Chains

In Section 3.4.2, a measure for the upper bound on the probability of the supply chain shortfall exceeding the total flexibility shortfall was developed for configurations that have a low probability of stage-spanning bottlenecks. This measure is given by,

$$\Gamma_K(M_1^*,\ldots,M_K^*) = 1 - \prod_{k=1}^{K} \left(1 - \Pi_k(M_K^*)\right)$$

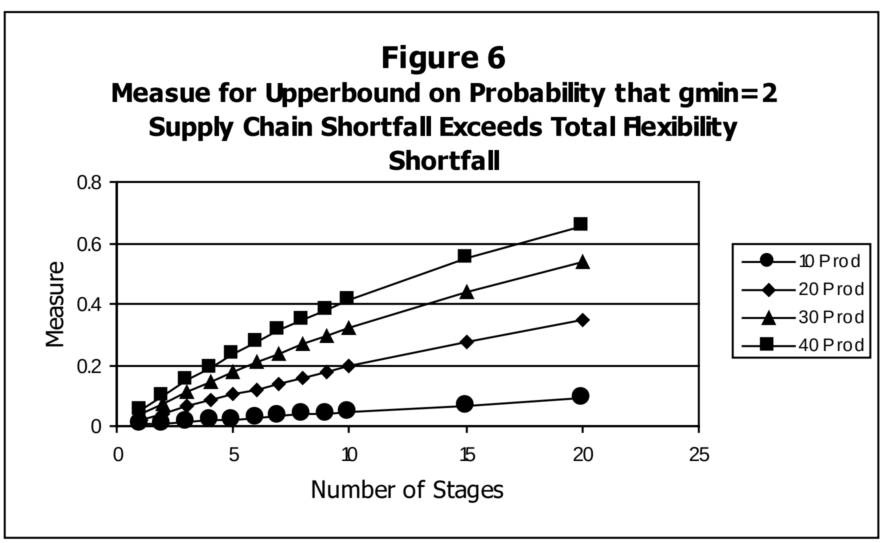
where $\Pi_k(M^*)$ is the measure of Jordan and Graves (1995). From Section 3.3.1.1, $\Pi_k(M^*)=\Omega_S(M^*,L_1^*)$. Using the derivation of Lemma 4, but setting N=1, it can be shown that,

$$\Pi_k(M^*) = \Omega_S(M^*, L_1^*) = \Phi\left(\frac{-(g_{\min}-1)\mu}{\sigma\sqrt{I/2}}\right)^2$$

Note that for a h-type chain configuration, for which $g_{\min}=h$ (see Section 3.5), this is the same expression as equation (10) in Jordan and Graves (1995).

For $g_{\min}=1$, stage-spanning bottlenecks may occur with a high probability, but even if they didn't, $\Pi_k(M^*)=0.25$ and so $\Gamma_K(M_1^*,\ldots,M_K^*)=1-(0.75)^K$. This is quite large even for small K, e.g. 0.25, 0.76, 0.94 for K=1,5,10 respectively. So even in the absence of stage-spanning bottlenecks, the probability that the supply chain shortfall exceeds the total flexibility shortfall may be quite large for $g_{\min}=1$-type supply chains. In Section 3.6, a certain $g_{\min}=1$-type supply chain, namely a pairs configuration, is shown to have a large probability of the supply chain shortfall exceeding the total flexibility shortfall.

For $g_{\min}=2$ or 3, Section 3.5.1.2 suggested that the probability of stage-spanning bottlenecks should be very low. Figure 6 shows $\Gamma_K(M_1^*,\ldots,M_K^*)$ for $g_{\min}=2$ and iid $N(100,30)$ demands.



This measure is quite large for systems with large numbers of products or stages. However, if g_{min}=3, then $\Gamma_K(M_1*,\ldots,M_K*)\leq 0.003$ for any system with no more that 40 products and 20 stages. This suggests that g_{min}=3-type supply chains should perform well even for systems with a very large numbers of products and stages. Supply chains with g_{min}=2- should perform well unless the numbers of products and stages is large. Performance here refers to the probability that the supply chain shortfall exceeds the shortfall of the supply chain under total flexibility.

3.6 Flexibility Configured in Pairs

In the introduction, pairs flexibility was defined as a configuration in which there were I plants with plants i and j being able to process products i and j only. Pairs configurations can therefore be defined by the product pairings, e.g. {1,2},{3,6} and {4,5} for a six-product case. Note that a pairs configuration is equivalent to a configuration in which there are $I/2$ plants in which every plant can process two products and no two plants can process the same products.

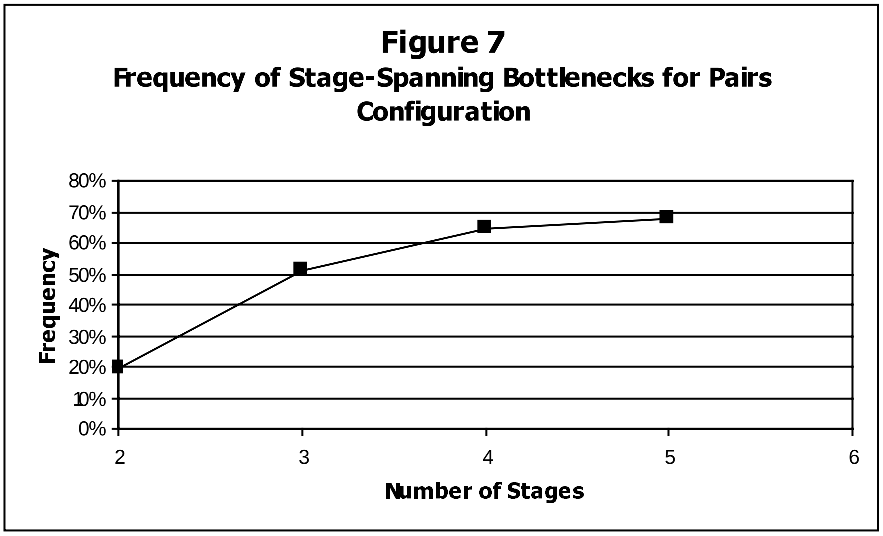
As a pairs flexibility configuration does not perform very well for single-stage supply chains (Jordan and Graves, 1995), it might be expected to perform even more poorly in multiple-stage supply chains. However, it provides a good baseline against which to compare other configurations. It is also illustrative of the effect that supply chain inefficiencies have on performance.

In this section I analyze supply chains in which every stage has I equal capacity plants and in which every stage is configured with some instance of a pairs configuration. There is no restriction on the particular pairs configuration used at any stage. Product demands are iid. In this case, the expected stand-alone stage shortfall is the same for all stages.

The simulation results presented are all based on a ten-product supply chain with product demands being iid $N(100,30)$ unless otherwise stated. The particular pairs configuration at each stage is different from all other stages. All plants have a capacity of 100.

3.6.1 Supply Chain Inefficiencies

A stage with I plants and a pairs configuration has g-value of 1. This can be seen by considering the subsets of paired product. The fraction of total stage capacity available to one of these pairs must be less than or equal to $2/I$, and therefore from equation (11) the g-value is 1. As a supply chain with a pairs configuration at each stage is a g_{min}=1-type supply chain, the upper bound on the probability of any particular stage-spanning bottleneck is 0.25 (see Section 3.5.1). This suggests that stage-spanning bottleneck might occur frequently. Figure 7 below, which shows the frequency of stage-spanning bottlenecks in a simulated ten-product supply chain with each stage having a different pairs configuration, confirms this conjecture. The probability of stage-spanning bottlenecks is large even for two-stage supply chains. For supply chains with five stages, the frequency is 68%.



Because the expected stand-alone shortfall is the same for all stages, the measure for floating bottlenecks developed in Section 3.3.2, equation (6), becomes,

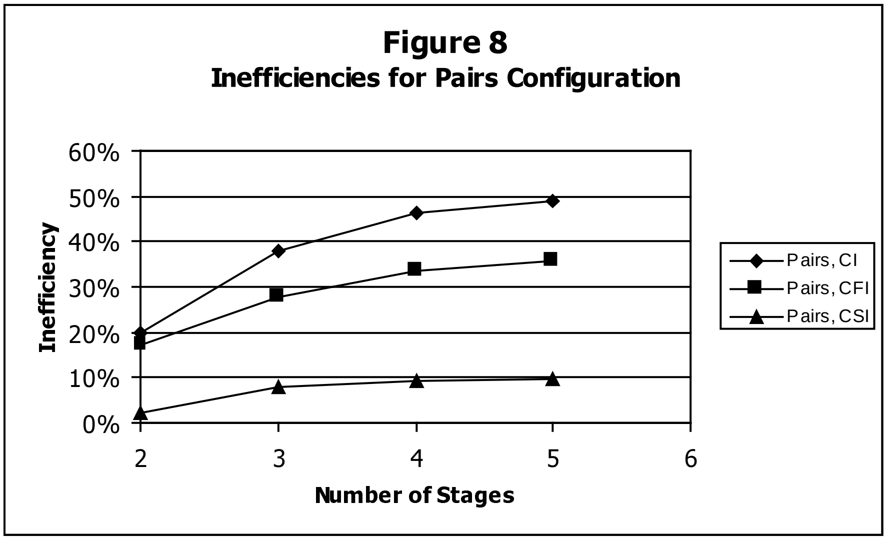
$$\text{Prob}\left[SF_k = SF_l \ \ \forall \text{stages } l, k = 1, \ldots, K\right]$$

This probability needing to be close to 1 if floating bottlenecks are to occur with a low probability (Section 3.3.2).

Consider a two-stage four-product (iid demands) supply chain with stage 1 configured as {1,2},{3,4} and stage 2 as {1,3},{2,4}, with all plants having capacity equal to the mean demand, μ. What is the probability that the stand-alone shortfall of each stage is different? The stand-alone shortfalls of the two stages will be the same if the optimal set of products, *M**, in the single-stage version of shortfall formulation, **P2**, is the same for both stages. The shortfalls will be different if the optimal *M** are different. The possible optimal sets for stage 1 are {∅}, {1,2}, {3,4} and {1,2,3,4}. For stage 2, they are {∅}, {1,3}, {2,4} and {1,2,3,4}. The optimal sets for these two stages, and hence the stand-alone shortfalls, will be different if either {1,2} or {3,4} is the optimal *M** set for stage 1 as these cannot be the optimal sets of stage 2. This will occur if the total demand for products 1 and 2 is greater (less) than 2μ and the total demand for products 3 and 4 is less (greater) than 2μ. This event has a probability of 0.5. So, the probability in (6), i.e. the probability that the stand-alone shortfalls of the two stages are the same, is less than or equal to 0.5, which is significantly less than 1. Therefore, even for this small supply chain, the probability of floating bottlenecks is large.

A five-stage ten-product supply chain in which each stage had a different pairs configuration was simulated to determine the probability that the stand-alone shortfall of stage k, k=1,…,5, was the same as that of stage l, l=1,…,5, $l≠k$. The further this probability is from one, the larger the floating bottleneck inefficiency is likely to be (Section 3.3.2). The probability that the stand-alone shortfall of stage k is the same as that of stage l was estimated by the frequency of this event in a 10,000 run simulation. The simulation yielded $P[SF_k=SF_l, k \neq l]=0.066$ on average across the twenty possible (l,k) stage combination, with a maximum of 0.153 and a minimum of 0.033. Therefore, the floating bottleneck inefficiency should be quite severe for this configuration.

Figure 8 shows the configuration inefficiencies for the pairs configuration as the number of stages increases from 2 to 5. As expected from above, both the stage-spanning (as measured by CSI) and floating inefficiencies (as measured by CFI) are large even for small numbers of stages.

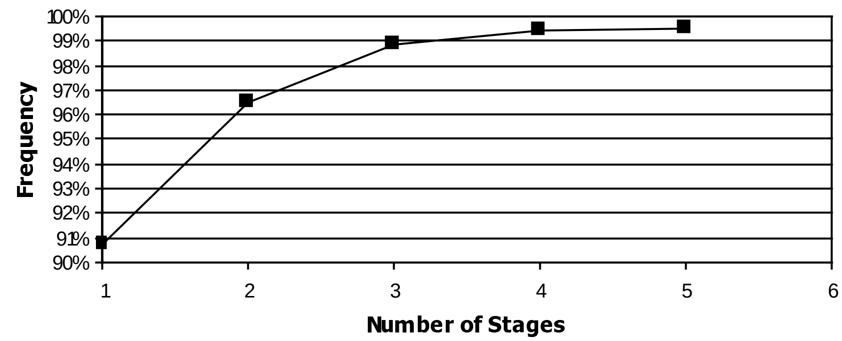


The susceptibility of the pairs configuration to the supply chain inefficiencies coupled with the fact that it does not perform very well even in single-stage supply chains should ensure that the pairs configuration performs poorly in multiple-stage supply chains.

3.6.2 Configuration Performance

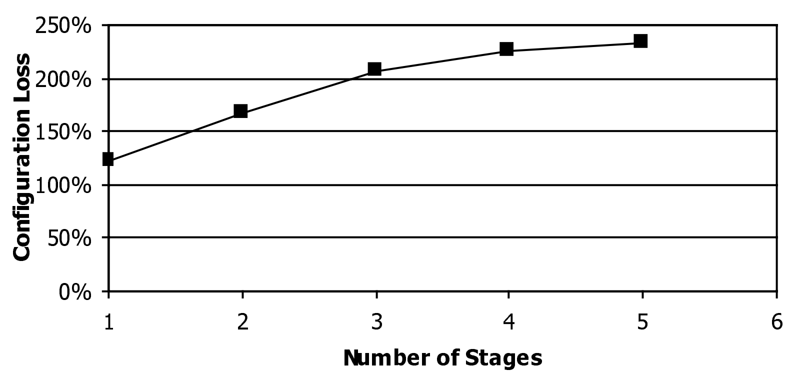
The upper bound on the probability of a $g_{\min}$=1 supply chain shortfall exceeding the total flexibility shortfall was shown to be large in Section 3.5, suggesting that a pairs configuration would have a large probability of this event. The simulation results (for the ten-product system) presented in Figure 9 below testify to this. The probability that the shortfall exceeds that of total flexibility is above 90% even for a single-stage supply chain and approaches 100% in supply chains with only a small number of stages.

Figure 9
Frequency of Shortfall Exceeding Total Flexibility Shortfallfor Pairs Configuration



With the probability of the shortfall exceeding that of total flexibility so large, and supply chain inefficiencies so frequent, the configuration loss should be large for supply chains using a pairs configuration. Figure 10 shows that the configuration loss in a ten-product system is very large even for supply chains with a small number of stages.

Figure 10
Configuration Loss for Pairs Configuration



A pairs configuration does not perform very well for single-stage systems. Supply chain inefficiencies cause its performance to become significantly worse in multiple-stage supply chains. The next section investigates a supply chain configuration that performs significantly better than a pairs configuration.

3.7 Flexibility Configured in Chains

As discussed in Section 3.5, a supply chain will perform well only if the individual stages perform well and the overall supply chain does not suffer from either the stage-spanning bottleneck or floating bottleneck inefficiencies.

Jordan and Graves (1995) showed that for single-stage supply chains, a chaining configuration, defined in the Section 3.1, performs remarkably well. An h-type chain is a chain in which every product can be produced in h plants. For complete chains with h=2, the expected shortfall was very close to that of total flexibility. In other words the configuration loss was very low. If chaining strategies also provide good protection against the supply chain inefficiencies, then supply chains in which each stage uses a chain configuration should perform well.

In this section I analyze supply chains in which every stage has I equal capacity plants and in which every stage is configured with some instance of a complete h-type chain configuration, where the h-value is the same for all stages. There is no restriction on the particular chain configuration at any stage. In this case, the expected stand-alone stage shortfall is the same for all stages.

The simulation results presented are all based on an I product supply chain with product demands being iid N(100,30) unless otherwise stated. The particular h-type chain configuration at each stage was randomly generated. All plants have a capacity of 100. In this case the expected stand-alone stage shortfall is the same for all stages.

3.7.1 Supply Chain Inefficiencies

A stage with I equal capacity plants and an h-type chain configuration has g value equal to h (Section 3.5). An h-type chain has a total of Ih product-plant links. From equation (11), for a configuration to have a g value equal to h, the fraction of total stage capacity available to any single product must be at least h/I. Therefore, a configuration with a g value equal to h must have at least Ih product-plant links. This can be seen by considering each individual product; each product must be connected to h plants if the fraction of total stage capacity available to it is to be at least h/I. Because an h-type chain requires only Ih links and a configuration with a g value equal to h requires at least Ih links, an h-type chain uses the minimum possible links required for a g-type configuration. If flexibility investment is measured in terms of the number product-plant links, no other configuration is more cost effective than an h-type chain in delivering a g value of h.

A supply chain with complete h-type chains at all stages is a $g_{\min}=h$-type supply chain. From Lemma 3(b), stage-spanning bottlenecks never occur in such supply chains if the number of products is less than or equal to $2(h+1)$. Even for supply chains with more products than this, the upper bound on the probability of any particular stage-spanning bottleneck is extremely low (Section 3.5.1.2). This suggests that stage-spanning bottleneck occur with a very low probability.

For the case of ten products and five stages, with *h*=2 and *h*=3, simulation gave a frequency of 0%; that is in 10,000 demand realizations, a stage-spanning bottleneck did not occur once. This compares to a frequency of 68% for a five-stage supply chain with a pairs configuration. In the case of a standard deviation of 50, the frequency was 0.67% for *h*=2 type chains and again 0% for *h*=3 type chains. So a chaining configuration offers very good protection against the stage-spanning bottleneck inefficiency.

How does it perform with respect to the floating bottleneck inefficiency? Because the expected stand-alone shortfall is the same for all stages, the measure for floating bottlenecks developed in Section 3.3.2 (equation (6)) becomes,

$$\text{Prob}\left[SF_k = SF_l \quad \forall \text{stages } l, k = 1, \ldots, K\right]$$

with this probability needing to be close to 1 if floating bottlenecks are to occur with a low probability.

Lemma 7

For a two-stage four-product supply chain with each stage having four plants and a type h=2 chain configuration, and all plant capacities being equal (=c), if the product demands are iid N(μ,σ), then the probability that the stand alone shortfalls for the two stages are the same is greater than or equal to,

$$1-\left(8\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\left[\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)+\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\right]+8\Phi\left(\frac{\mu-2c}{\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu-c}{\sigma}\right)\right)$$

For a mean product demand of 100, Table 3 shows that this lower bound (LB) is close to 1 for plant capacities greater than or equal to 90. Simulation results, also shown in

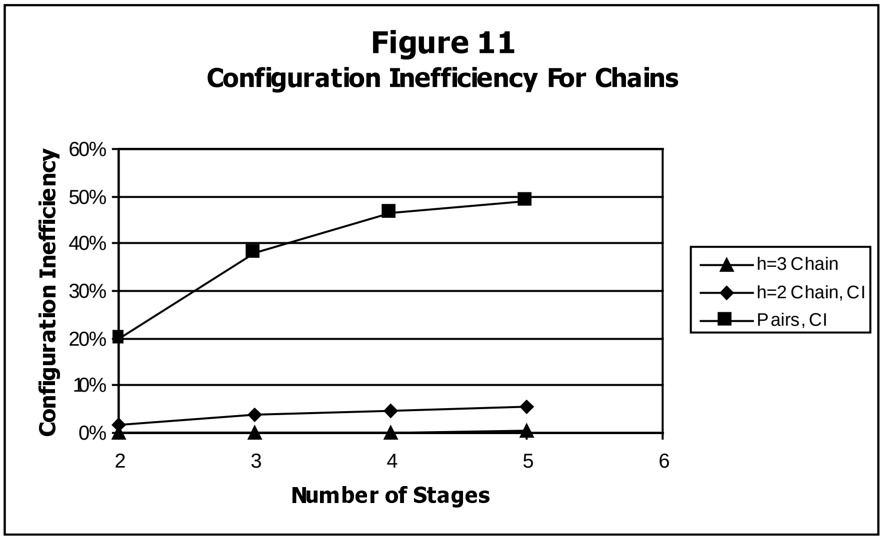
Table 3, show that the probability is very close to 1 for plant capacities greater than or equal to 70.

Table 3

	σ = 30		σ = 40		σ = 50	
Capacity	LB	SIM	LB	SIM	LB	SIM
70	0.9948	1.0000	0.3495	0.9894	0.7719	0.9554
80	0.9957	1.0000	0.4790	0.9870	0.8143	0.9554
90	0.9977	0.9998	0.9666	0.9909	0.8687	0.9629
100	0.9992	0.9999	0.9832	0.9940	0.9202	0.9754
110	0.9998	1.0000	0.9935	0.9982	0.9586	0.9872
120	1.0000	1.0000	0.9981	1.0000	0.9818	0.9952
130	1.0000	1.0000	0.9996	1.0000	0.9933	0.9991

In order to investigate whether the probability remains close to 1 as the number of products increases, a five-stage ten-product supply chain was simulated. Each stage had a different, randomly selected, chain configuration. The probability that the stand-alone shortfall of stage k, k=1,…,5, was the same as that of stage l, l=1,…,5, $l \neq k$, was estimated by the frequency of this event in a 10,000 run simulation. The closer this probability is to one, the smaller is the floating bottleneck inefficiency (Section 3.3.2). The simulation for h=2 chains, yielded $P[SF_k=SF_l,\ k \neq l]$=0.894 on average across the twenty possible (l,k) stage combination, with a maximum of 0.909 and a minimum of 0.875. For h=3, the average was 0.995. This compares to an average of 0.066 for a pairs configuration, see Section 3.6.1. Therefore, chain configuration would appear to provide good protection against the floating bottleneck inefficiency, especially for h=3.

Figure 11 below shows the total configuration inefficiencies for the pairs configuration and chain configurations as the number of stages increases from 2 to 5. As expected from above, the inefficiencies are very low for the chain configurations, with a maximum inefficiency of 0.32% for h=3-type chains. As stage-spanning bottlenecks never occurred for the chains, the inefficiency is all caused by the floating bottleneck inefficiency in this case.



3.7.2 Configuration Performance

The upper bound on the probability of a g_{min}=2 supply chain shortfall exceeding the total flexibility shortfall is small unless the number of products or the number of stages is large(Section 3.5.2). For g_{min}=3 the upper bound is extremely small even for supply chains with a very large number of products and a very large number of stages. This suggests that for a chain configuration the probability of the supply chain shortfall exceeding the total flexibility shortfall would have be very small. For ten-product supply chains, the simulation results presented in Figure 12 testify to this. For the h=3 type chains the maximum probability is 1.07%.

With the probability of the shortfall exceeding that of total flexibility low, and supply chain inefficiencies very small, the configuration loss should be small for supply chains using a chain configuration. As defined in Section 3.4.1, the configuration loss compares the configuration shortfall to the shortfall of a totally flexible system. Figure 13 below shows the configuration loss for *h*=2 and *h*=3 chains. The maximum configuration loss for *h*=3 was 0.34%. Figure 14 provides a comparison with the pairs configuration.

Figure 12
Frequency of Shortfall Exceeding Total Flexibility Shortfall for Chain Configuration

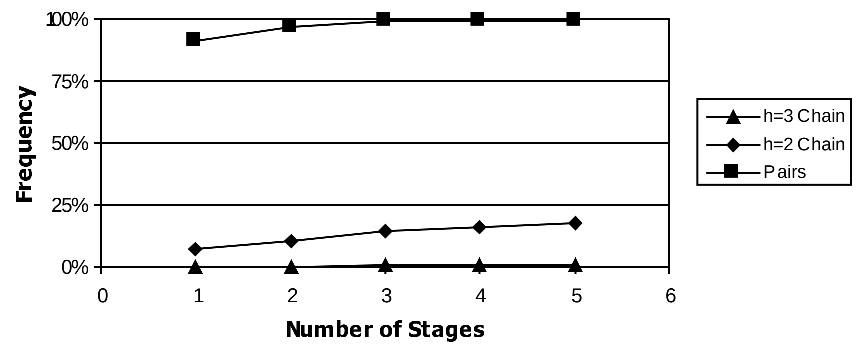


Figure 13
Configuration Loss for Chain Configuration

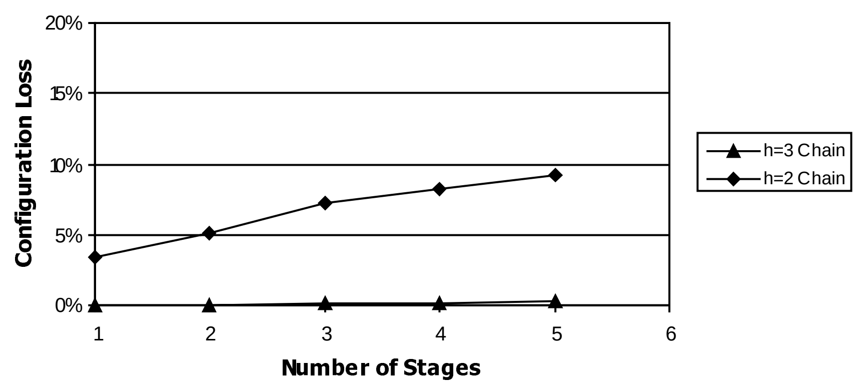
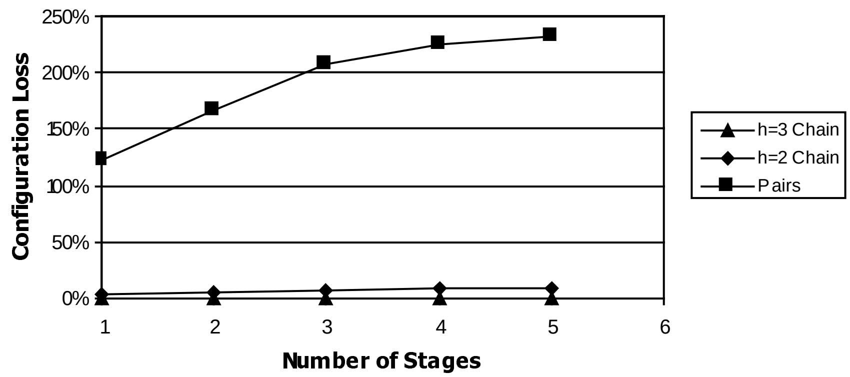
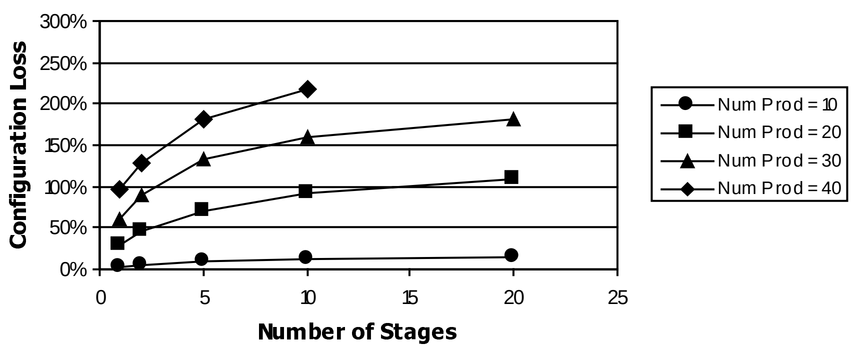


Figure 14
Configuration Loss for Chain Configuration



The *h*=2 chaining strategy performs quite well but its performance decreases as the number of stages increases. The *h*=3 configuration is less susceptible to this as the probability of the floating bottleneck inefficiency is much lower (Section 3.6.1). As the number of product increases, the performance of the *h*=2-type chain decreases. Figure 15 below shows the configuration loss for the *h*=2-type chain as the number of products increases. As the simulation time increases rapidly in the number of products and stages, only 1000 runs were made for the 30 and 40-product cases.

Figure 15
Configuration Loss as the Number of Stages and the Number of Products Increase

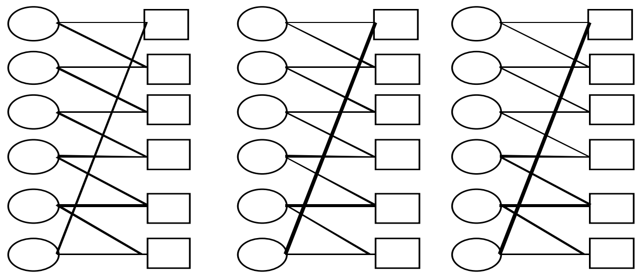


A similar simulation for h=3-type chains resulted in a maximum configuration loss of 1% over all test cases, showing that the h=3-type supply chain configuration performs almost identically to a totally flexible supply chain even for very large numbers of products and stages. An h=2-type supply chain configuration performs almost identically to a totally flexible supply chains as long as the number of stages or products is not too large.

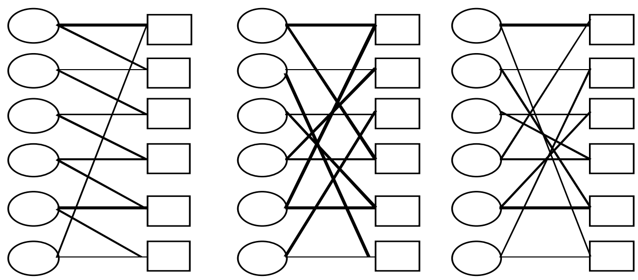
Jordan and Graves (1995) showed that an h=2-type chaining strategy works well for single-stage supply chains. This dissertation shows that a similar strategy works well in multiple-stage supply chains, but an extra "layer" of flexibility may need to be added if the number of products or stages is large. Chains perform well because they provide good protection from both the stage-spanning and floating bottleneck inefficiencies and because they perform well on an individual stage basis.

3.7.3 Random Chains versus Replicated Configurations

A supply chain in which every stage is an exact replication of the first stage will perform identically to a single-stage supply chain. In other words, replication avoids any supply chain inefficiency. By replication, I mean that each stage has the exact same number and type (in capacity terms) of plants and the same flexibility configuration. Figure 16 below shows a three-stage supply chain in which each stage is a replication of the first stage and another three-stage supply chain in which there is a randomly selected chain at each stage.



(a) replicated chain at each of the three stages



(b) random chain at each of the three stages

Figure 16: A three-stage supply chain with (a) replicated chain configurations and (b) random chain configurations

One strategy would then be to simply configure every stage to be identical. However this may be infeasible for a number of reasons; it may be technically impossible, prohibitively expensive or the various stages of the supply chain may be under the control of different decision-makers. Such a replication strategy does however provide a good baseline against which to measure other strategies.

A random chaining strategy was shown to perform well in the Section 3.7.2. How does it compare to a replication strategy? Figures 17 and 18 show the performance of random chaining strategy (*h*=2 and *h*=3) versus a replicated pairs and a replicated *h*=2 chain strategy, respectively. The random chaining strategy significantly outperforms the replicated pairs strategy. The random *h*=2 chain strategy is outperformed by a replicated *h*=2 chain strategy, where the performance gap grows with the number of stages. Nevertheless, the random *h*=2 strategy is still quite good. This is because of the low probability of supply chain inefficiency in chain configurations. The random h=3 chain strategy performs better than a replicated h=2 chain strategy.

Figure 17
Random Chain versus Replicated Pairs Strategy

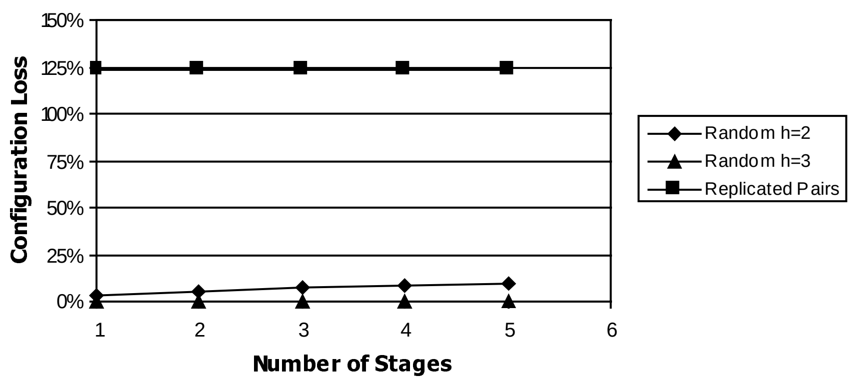
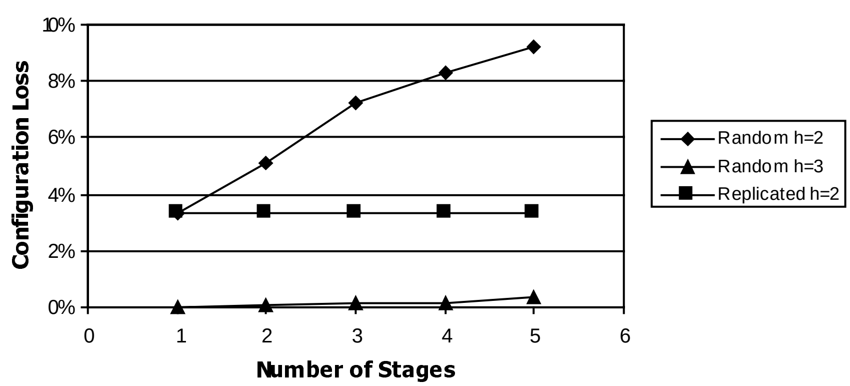


Figure 18
Random Chain versus Replicated h=2 Chain Strategy



3.7.4 Unequal Capacity Usage

As mentioned in the model discussion in Section 3.2, this dissertation assumes an equal capacity usage for all products that a plant can process. This assumption is made so as to focus on the effect that configurations have on performance without having the added complication of unequal capacity usage. Unequal capacity usage brings a new dimension to the problem. Product-plant allocations should also consider the amount of capacity needed per unit of product in a plant. The previous results for chains and pairs configurations should still hold when capacity usage is product dependent. However, when choosing the chain configuration for a stage, a random choice is probably not the best policy, rather some attempt should be made to avoid assigning a product to a plant in which it has a high per unit capacity usage.

A ten-product five-stage supply chain in which the capacity usage's were randomly generated was tested with both a random pairs and a random chain strategy. For each possible product-plant link, the capacity usage parameter was randomly generated; with the values being in the range [0.85,1.15]. For these capacity usages, the expected shortfall in supply chains configured with either random pairs, random h=2 chains or random h=3 chains was estimated using simulation. The chains (h=2 and 3) again significantly outperformed the pairs strategy, yielding expected shortfalls of more than a factor of three less than the pairs configuration. As noted above, a random strategy could be improved upon by a more judicious choice of plant-product allocations. However the value of a chaining strategy should still be high.

18 Conclusion

Most existing flexibility and capacity research (e.g. Fine and Freund, 1990, Li and Tirupati, 1994,1995,1997, Jordan and Graves, 1995) has focused on single-stage systems, in effect assuming that the bottleneck stage was demand independent. This stream of flexibility research has developed our understanding of the role of process flexibility in coping with demand uncertainty but leaves open the question of flexibility investments in multiple-stage systems.

This dissertation has highlighted two potential inefficiencies in multiple-product multiple-stage supply chains that affect the performance of flexibility configurations; namely floating bottlenecks and stage-spanning bottlenecks. The existence of floating bottlenecks has been previously noted in the literature, but stage-spanning bottlenecks have not been identified to my knowledge. Stage-spanning bottlenecks are an interesting inefficiency in that they can arise even when demand is deterministic. A flexibility configuration must not only perform well on an individual stage basis but also provide protection against these possible supply chain inefficiencies.

A chaining strategy that has already been shown to perform well in single-stage systems is shown to help prevent, or at least mitigate, the two supply chain inefficiencies. As such, they perform well in multiple-stage supply chains.

In practice supply chains will be more complicated than those analyzed in this research. It is unlikely that one could implement the idealized complete chain configuration at all stages. However it is still possible to infer guidelines for supply chain flexibility deployment from this research.

For single-stage supply chains, the guidelines of Jordan and Graves (1995) should be followed. These are (i) try to create chains that encompass as many plants and products as possible (ideally all plants and products would be part of one single chain), (ii) try to equalize the number of plants (measured in total units of capacity) to which each product in the chain is directly connected and (iii) try to equalize the number of products (measured in total units of expected demand) to which each plant in the chain is directly connected.

As noted by Jordan and Graves (1995), in single-stage systems there are rapidly diminishing benefits to adding more flexibility once an h=2 chain has been formed. However, in multiple-stage systems, if the supply chain produces many products or comprises numerous stages, then an h=2 chaining strategy may not adequately protect against the supply chain inefficiencies identified in this dissertation. In the idealized case, where there is an equal number of products and plants at each stage, adding flexibility to go from an h=2 chain to an h=3 chain at each stage significantly increases the supply chain performance. The benefits of flexibility do not decrease as rapidly in multiple-stage systems as in single-stage systems. This is a result of the potential inefficiencies in multiple-stage supply chains. Additional flexibility provides protection against inefficiencies. Note that in going from an h=2 chain to an h=3 chain, one is essentially overlaying another h=2 chain onto the original h=2 chain at each stage. Once this extra layer of flexibility, again configured as a chain, has been added, the performance of the supply chain is almost identical to that of total flexibility, indicating that further benefits to additional flexibility rapidly decrease.

For multiple-stage systems, at a minimum the single-stage guidelines, (i)-(iii) above, of Jordan and Graves (1995) should be followed to create a chain structure for each of the supply chain stages. In supply chains with a large number of products or stages, additional flexibility is advisable, especially for stages in which the capacity is not much greater than the expected demand. This extra layer of flexibility should again be added in accordance with the above guidelines to create another chain structure overlaying the initial chain structure.

4 Capacity Decisions in Multiple-Product Supply Chains

4.1 Introduction

Demand for new products is frequently uncertain. Even existing products may face demand variability that cannot be predicted. Capacity, inventory and flexibility are mechanisms that firms can use to cope with this uncertainty. By investing is sufficient capacity, a firm can fulfill the product demand with a high probability. Capacity investment is typically expensive and a balance should be struck between capacity and demand fulfillment.

Capacity decisions can be complicated by the nature of the associated product supply chain. A product may require processing at more than one stage. In such cases, the capacity investment decision must consider multiple stages. However if the supply chain processes a single product only, and stage capacities are deterministic and reliable, then stage capacities should be equal, where capacities are expressed in terms of final product units. The supply chain capacity decision can be mapped into a single-stage capacity decision in such situations.

Supply chains often process more than one product. This may arise in situations where a firm produces variants of a common underlying product; both computer and car manufacturers offer feature lists from which a product can be customized to an individual's requirement. Alternatively a firm may produce products for different markets but the products may share a common processing technology. Job shops produce a wide range of products using a collection of different processing resources; different products require different resources for their production. Multiple product supply chains are subject to floating bottlenecks, in which the set of stages that limit throughput is dependent on the product demand realizations. Floating bottlenecks have been recognized in theory (e.g. Hopp and Spearman, 1997) and practice (e.g. Alcalde, 1997). One possible formalization of floating bottlenecks is given in this dissertation. Capacity decisions in multiple-product multiple-stage supply chains are complicated by the possibility of floating bottlenecks. It is insufficient to focus on the capacity of one stage only; the capacities of the various stages need to be considered.

Demand uncertainty is not the only factor in capacity decisions. Even if future demands are deterministic, the timing and quantity of capacity decisions can be complicated by the capacity cost function, which typically exhibits economies of scale.

Capacity expansion problems have been studied in the operations literature since the late 1950's, with the vast majority of work focusing on the timing and quantity of investments given deterministic future demand for a single product. In single product problems, there is only one bottleneck stage. A capacity planning model that focuses only on the most expensive, in terms of capacity, stage is thus sufficient. The early work of Manne (1967) focused on single-product problems and this focus dominated until the 1980's. As single-product single-stage capacity planning is not the subject of this dissertation, I make no attempt to provide a detailed survey of the literature in this field. A comprehensive survey can be found in Luss (1982).

As recognized by Manne, the assumption of deterministic future demand is strong, and single-product capacity planning under uncertainty has been studied in a number of papers (Davis, Dempster, Sethi and Vermes, 1987, Paraskevopoulos, Karakitsos and

Rustem, 1991, Bean, Higle and Smith, 1992). In all of these papers, only one stage is considered, but this is reasonable due to the single product focus of the papers.

Eppen, Martin, and Schrage (1988) describe a multiple-product single-stage multiple-plant capacity planning model developed for General Motors. By single-stage multiple-plant, I mean that the products require processing at one stage only, but that there are multiple plants at the stage that can provide the required processing. The model is quite comprehensive in its treatment of a single-stage multiple-plant problem. The model considers a five year planning horizon and determines the capacity configuration that each plant will take on in each of the five years. Uncertainty is modeled by allowing each product to have three possible demand scenarios in each of the five years. The overall model is formulated as an integer stochastic program with recourse. Due to the relatively complex recourse function, allowing more than three demand scenarios severely affects computation time. The model incorporates twenty products, seven plants, and thirty possible capacity configurations in total across the seven plants.

Multiple-stage investment has only recently become a focus of attention (Eberly and Van Mieghem, 1997 and Harrison and Van Mieghem, 1999). Motivated by the focus in economics and operations on single-factor (or stage) investment, Eberly and Van Mieghem develop a quite general model of multiple-factor investment under uncertainty. Allowing for investment and disinvestment in factors, concave operating profits and a general form of uncertainty, the authors determine the optimal multi-period investment policy structure under the assumption that capacity investment and disinvestment costs are convex. It should be noted that capacity investment costs are more likely to be concave as they often exhibit economies of scale. The optimal policy partitions the possible investment space into various regions. Depending on the region in which the current capacity vector falls, investment or disinvestment in a factor is optimal. There is also a "continuation region" in which it is optimal to make no adjustments to the various factors. When outside this region, it is optimal to move to the boundary of the continuation region. While this paper establishes the structure of the optimal policy under general conditions, determining the actual optimal investment levels for more specific problems is still a challenge.

Harrison and Van Mieghem (1999) build on the work of Eberly and Van Mieghem (1997) by studying a more specific problem formulation. Studying a multiple-product, multiple-stage capacity planning problem in which a product requires processing at a subset of the stages, the authors prove that the optimal capacity investment involves some hedging that would never be optimal if the demands were known. In other words, there may be no demand scenario for which all stages operate at capacity. Capacity decisions are made when demands are uncertain; production planning, which maximizes profit subject to capacity constraints, occurs after demands become known. The product demand vector is assumed to be independent and identically distributed from period to period and, as in Eppen, Martin, and Schrage (1988), inventory cannot be carried from one period to the next. This enables the authors to formulate the investment decision as a single period problem. An additional implied assumption is that the stages are totally flexible. Investment and disinvestment in the various capacities are allowed subject to linear costs instead of the more general convex cost allowed by Eberly and Van Mieghem (1997). For sufficiently small problems, the continuation region can be calculated. A two product three stage problem is discussed in the paper. As noted by the authors, the optimal solution “requires solving the characteristic equations simultaneously using the multivariate demand distribution”, a requirement that is likely to be prohibitive for many real problems. Indeed, the purpose of this paper is really to provide insight into the notion that multiple-stage capacity planning under uncertainty involves hedging rather than to provide an efficient solution method.

Inventory decisions are closely related to capacity decisions in that the firm is investing in a buffer to cope with uncertain demand. Inventory theory has used both expected shortfall costs and service level metrics in developing inventory policies. The multiple-product multiple-stage capacity literature has focused solely on the expected cost metric. Capacity decisions based on a service level criterion have not been studied in multiple-product multiple-stage supply chains although Li and Tirupati (1994, 1997) use this metric in a single-stage multiple-plant model.

Chapter 4 studies the capacity decision in multiple-product multiple-stage supply chains and develops problem formulations that can be solved using Excel Solver. Unlike other commercial optimization packages, such as CPLEX, Excel Solver is widely available to operations managers. It is therefore beneficial to have Excel based solution techniques. Section 4.2 introduces the supply chain model used. Section 4.3 studies the capacity decision using a service level criterion while Section 4.4 focuses on an expected shortfall criterion. Concluding remarks are presented in Section 4.5. Proofs of all lemmas can be found in Chapter 5 (Appendix 3).

4.2 The Model

The supply chain consists of K different stages, k=1,…,K. The numbering of the stages does not imply any specific processing sequence and the supply chain need not be serial in nature. Any general multiple-stage production system in which there are K distinct operations, where an operation is distinct if it requires a different processing resource from all other operations, is allowed.

The supply chain produces I different products, i=1,…,I, with product i requiring processing at a subset $Q(i)$ of stages. Stage k must process the set of products $P(k)$. Each stage k comprises processing resources that are totally flexible. As such, each stage can be modeled as a single resource with capacity c_k. One unit of product i requires β_{ik} units of stage k's capacity. β_{ik} is referred to as the production coefficient for product i at stage k.

As in Harrison and Van Mieghem (1999), the product demand vector is assumed to be independent and identically distributed from period to period, no inventory is carried from one period to the next and production planning occurs after demand is realized. The demand distribution and inventory assumptions allow the capacity decision to be mapped into a single period problem.

Capacity decisions must be made before the product demands are realized. Capacity costs are linear, with stage k capacity costing p_k per unit. This ignores any fixed cost element of capacity acquisition. However these fixed costs can be ignored if all stages require a positive capacity acquisition. All stages will require a positive capacity if the investment is being made for a system in which no stage has existing capacity, for example a completely new plant. However, if some capacity already exists, this linear cost assumption may not hold, and the models presented would need to be modified to incorporate the fixed costs.

Certain problem parameters are stochastic, while others are deterministic. Deterministic parameters, such as plant capacities, are written in lower case. Random variables are written in upper case, and actual realizations in lower case. Thus the product demand random vector is denoted by $\mathbf{D}=\{D_1,\ldots,D_I\}$ and a demand realization vector by $\mathbf{d}=\{d_1,\ldots,d_I\}$. Demands are assumed to be normally distributed.

It is assumed that the supply chain is optimized by one central decision-maker, referred to as the firm, as opposed to multiple independent agents as considered in Chapter 2.

This supply chain model is almost identical to the model presented in Chapter 3. The key difference is that stages are assumed to be totally flexible and products only require processing at a subset of stages.

Alcalde (1997) defines a flow path as a set of products requiring processing at the same set of stages. In other words, products i and j are in the same flow path, f, if and only if $Q(i)=Q(j)$. Flow paths allow products to be aggregated in such a way as to simplify the supply chain description while still keeping the key characteristics. To ensure that the aggregation does not alter the optimal solution in the models presented in Sections 4.3 and 4.4, I add the restriction that $\beta_{ik}=\beta_{fk}$ for all products in flow path f. Let F be the set of flow paths for the products processed by the supply chain. It is assumed that if $Q(i)=Q(j)$ and $\beta_{ik}=\beta_{jk}\ \forall k\in Q(i)$, then both i and j are part of the same flow path (i.e. F is the set of flow paths with minimum cardinality). Let $T(f)$ be the set of products in flow path f. The demand for flow path f, D_f, is then given by $D_f = \sum_{i \in T(f)} D_i$. From here on, $Q(f)$ and $P(k)$ refer to flow paths rather than products. Flow paths are denoted by the set of stages they pass through.

4.2.1 Nomenclature

For ease of reference, the nomenclature used in this chapter is presented.

<u>Scalars</u>

c_k:	The capacity of stage k
d_i:	The demand realization for product i
D_i:	The demand random variable for product i
d_f:	The demand realization for flow path f
D_f:	The demand random variable for flow path f
F:	The total number of flow paths (f=1,…,F)
g:	A lower bound on the service level
I:	The number of products (i=1,…,I)
K:	The number of stages in the supply chain (k=1,…,K)
p_k:	The per unit capacity cost of stage k
S:	The number of demand scenarios (s=1,…,S)
sf_f:	The shortfall for flow path f
sf_f^s:	The shortfall for flow path f in scenario s
sf_{SC}:	The supply chain shortfall realization
SF_{SC}:	The supply chain shortfall random variable
T:	An upper bound on the expected shortfall
X_k:	The demand for stage k capacity $= \sum_{f \in P(k)} \beta_{fk} D_f$
Z_k:	The protection level for stage k $= (c_k - \mu_{X_k}) / \sigma_{X_k}$

β_{ik}: The production coefficient for product i at stage k (the amount of capacity of stage k used per unit production of product i)

μ_i: The expectation of demand for product i

μ_f: The expectation of demand for flow path f

μ_{Xk}: The expectation of demand for stage k capacity

σ_i: The standard deviation of demand for product i

σ_f: The standard deviation of demand for flow path f

σ_{Xk}: The standard deviation of demand for stage k capacity

ρ_{kl}: The correlation coefficient for stage k and stage l capacity demands

Vectors/Matrices:

d: The product demand realization vector $=\{d_1,\ldots,d_I\}$

D: The product demand random vector $=\{D_1,\ldots,D_I\}$

R: The correlation coefficient matrix for the stage capacity demands

Z: The protection level vector $=\{Z_1,\ldots,Z_K\}$

Sets:

$P(k)$: The set of products that require stage k

$Q(i)$: The set of stages k that process product i

Λ: A subset of the stages $\{1,\ldots,K\}$

4.3 Service Level Criterion

When investing in capacity, a firm may decide that it requires sufficient capacity at each stage in the supply chain so as to ensure that the total demand for all products is met with a certain probability in each period. This is analogous to the service level criterion common in inventory theory. This section develops a formulation for and insight into this service level problem. The motivating supply chain used is described next.

4.3.1 Work Center A

Graham (1998) studies an actual production system of a large US manufacturer. This production system is referred to as Work Center A. Figure 1 depicts the work center.

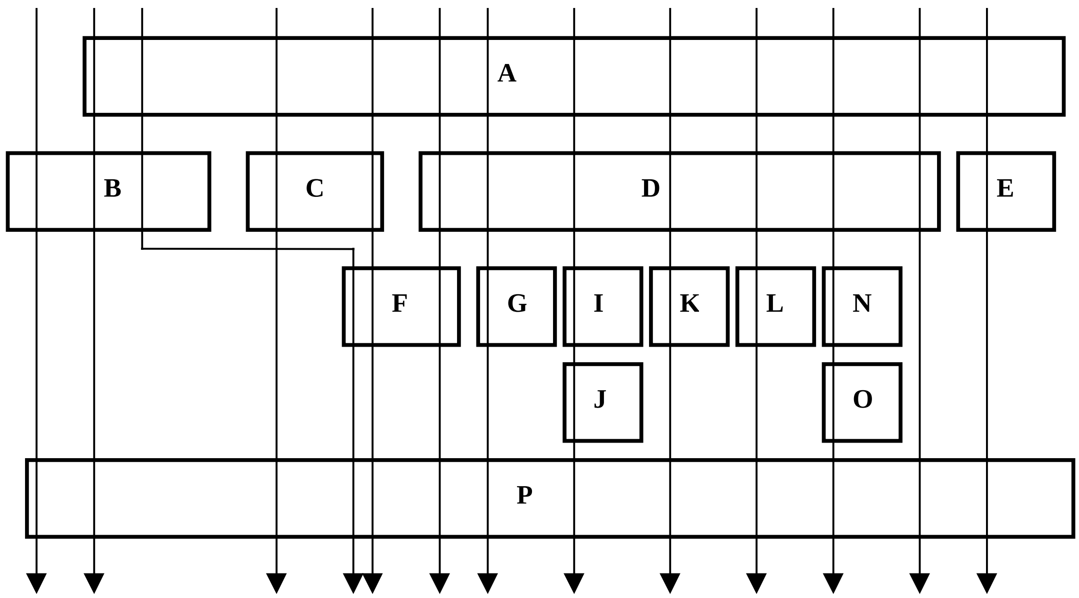


Figure 1: Work Center A

Boxes denotes production stages and lines denote flow paths – a flow path is defined by the stages the line passes through.

Daily production requirements are uncertain and the schedule for a particular day is set two days in advance. However, material properties dictate that Work Center A be operated with no inventory carried over from one day to the next ("There is no planned inventory in the work center. While some work in process does 'pile up' in front of some processes as the day goes on, at the end of each day, no in-process material is left in the work area. All orders are started and finished on the same day" Graham, 1997). Each day starts with zero inventory and a known demand to be completed that day. This system maps well into the supply chain model introduced in Section 4.2.

Daily demand data for the product flows (d_f) are presented in Table 1. The production coefficients (β_{fk}) for the flow paths are presented in Table 2.

Table 1

Flow Path	Mean	St. Dev.
BP	63.38	52.98
ABP	32.86	54.02
ABFP	3.43	6.35
ACP	300.52	214.37
ACFP	411.19	149.30
ADFP	165.00	133.45
ADGP	0.95	4.36
ADIJP	372.81	256.29
ADKP	42.40	70.77
ADLP	4.76	21.82
ADNOP	4.10	9.71
ADP	1216.19	564.68
AEP	25.48	25.53

Table 2

	A	***B***	***C***	***D***	***E***	***F***	***G***	***I***	***J***	***K***	***L***	***N***	***O***	***P***
BP		0.90												0.77
ABP	0.56	1.05												0.77
ABFP	0.45	1.05				1.00								0.77
ACP	0.73		1.08											1.33
ACFP	0.75		0.92			1.00								0.77
ADFP	1.21			1.25		1.00								0.77
ADGP	0.41			0.81			1.00							0.95
ADIJP	2.01			1.25				1.00	1.00					1.42
ADKP	0.90			0.81						1.00				0.65
ADLP	1.76			1.25							1.00			0.71
ADNOP	0.86			0.74								1.00	1.00	1.18
ADP	0.94			0.89										1.48
AEP	1.41				1.00									1.42

It should be noted that stages *I* and *J* both process flow path *ADIJP* only. As such, the same capacity, in terms of *ADIJP* units, should be installed for both stages. Therefore one can aggregate these two stages into a single stage with a capacity cost equal to c_I+c_J. This new stage is denoted by *I*. Likewise stages *N* and *O* can be aggregated and this aggregated stage is denoted by *N*.

Graham (1997) does not provide these data, rather these data were obtained directly from the author. I would like to thank Jud Graham for kindly providing these data.

4.3.2 The Formulation

Given a service level target, one wants to minimize the cost of capacity needed to obtain this service level. A service level target of *g* means that the probability that all flow path demands are met must be at least *g*. A shortfall will occur if the total demand for any stage capacity exceeds that stage's installed capacity.

The total demand for stage capacity at stage *k* is referred to as stage *k* demand, X_k, which is given by,

$$X_k = \sum_{f \in P(k)} \beta_{fk} D_f$$

If individual product demands, and hence flow path demands, are assumed to be normally distributed, X_k is normally distributed. Assuming that the flow path demands are independent, the mean and variance of X_k are given by,

$$\mu_{X_k} = \sum_{f \in P(k)} \beta_{fk} \mu_{D_f}$$

$$\sigma^2_{X_k} = \sum_{f \in P(k)} \beta^2_{fk} \sigma^2_{D_f}$$

The correlation coefficient, ρ_{kl}, for stage *k* and stage *l* demands, is given by,

$$\rho_{lk} = \frac{\sum_{f \in [P(k) \cap P(l)]} \beta_{fk} \beta_{fl} \sigma_{D_f}^2}{\sqrt{\sum_{f \in P(k)} \beta_{fk}^2 \sigma_{D_f}^2} \sqrt{\sum_{f \in P(l)} \beta_{fl}^2 \sigma_{D_f}^2}}$$

with ρ_{lk}=0 if there is no flow path that visits *l* and *k*. Note this correlation coefficient is non-negative; that is stage demands are positively correlated if flow path demands are independent.

If flow path demands were correlated, then the above variance and correlation expressions would need to be modified to incorporate this dependence.

The stage demands can then be expressed in vector form where **X** is the vector of stage demands, $\mathbf{X}=\{X_1,\ldots,X_K\}$. **X** has a multivariate normal distribution with a mean vector, variance vector and correlation matrix specified by the elements given above.

If a given realization of stage *k*'s demand exceeds stage *k* capacity, i.e. $x_k>c_k$, then there is a shortfall. Therefore the probability that all demand is met, given a capacity vector of $\mathbf{C}=\{c_1,\ldots,c_K\}$, is $P[\mathbf{X}\leq\mathbf{C}]$. This can be expressed in terms of the standard unit multinormal distribution as $P[\mathbf{X}\leq\mathbf{C}]=\Phi[\mathbf{Z},\mathbf{R}]$. $\mathbf{Z}=\{Z_1,\ldots,Z_K\}$ where,

$$Z_k = \frac{c_k - \mu_{X_k}}{\sigma_{X_k}}$$

and **R** is the correlation matrix specified by the ρ_{kl}. The probability that stage *k* demand will exceed its actual capacity is given by $\Phi[Z_k]$ where $\Phi[\bullet]$ is the standard normal cumulative distribution function. Z_k is referred to as the protection level for stage *k*.

The problem of determining the least cost capacity investment vector subject to a service level of *g*, is then given by **P1**,

$$\text{Min}\ \{\sum_{k=1}^{K} p_k c_k\}$$

subject to

$$\mathbf{\Phi[Z,R]} \geq \text{g}$$

The service level is assumed to be high enough so that a positive capacity is required at each stage. **P1** cannot be solved as formulated because closed form expressions for multivariate normal probabilities do not exist. I develop two different approaches solving this problem. The first involves approximating $\Phi[\mathbf{Z},\mathbf{R}]$ by ignoring the correlation structure of the random variables. The second involves obtaining a lower bound for $\Phi[\mathbf{Z},\mathbf{R}]$ in which some of the correlation structure is captured.

4.3.3 Independent Stage Demands

By assuming stage demands to be independent, I hope to develop insight into the optimal capacity investment in the absence of correlation. This will then provide a benchmark against which to compare the capacity investment vector when correlation is considered.

Assuming the stage demands to be independent, then $\Phi[\mathbf{Z},\mathbf{R}]$ is given by,

$$\mathbf{\Phi[Z,R]} = \prod_{k=1}^{K} \Phi[Z_k]$$

and **P1** becomes,

$$\text{Min}\ \{\sum_{k=1}^{K} p_k c_k\}$$

subject to

$$\prod_{k=1}^{K} \Phi[Z_k] \geq \text{g}$$

This problem is referred to as **P1,Ind** and can be thought of as an approximation to **P1** in which the correlation is ignored.

Note that if correlation matrix **R** has the special structure in which $\rho_{ij}=\lambda_i\lambda_j \forall i\neq j$ and $\lambda_i \geq 0$, $i,j=1,\ldots,K$, then a lower bound for $\Phi[\mathbf{Z},\mathbf{R}]$ is given by $\prod_{k=1}^{K}\Phi[Z_k]$ (Tong, 1980). However, this will not be true in general and so the constraint in **P1,Ind** is an approximation rather than a lower bound.

The Lagrangian relaxation of **P1,Ind** is given by,

$$L(\mathbf{C},\lambda) = \sum_{k=1}^{K} p_k c_k + \lambda\left(g - \prod_{k=1}^{K}\Phi[Z_k]\right)$$

and the partial derivatives with respect to c_k is

$$\frac{\partial L(\mathbf{C},\lambda)}{\partial c_k} = p_k - \lambda\left(\frac{\partial \Phi[Z_k]}{\partial c_k}\right)\prod_{l=1, l\neq k}^{K}\Phi[Z_l] \quad k = 1,\ldots,K$$

Now,

$$Z_k = \frac{c_k - \mu_{X_k}}{\sigma_{X_k}} \text{ and } \frac{\partial \Phi[x]}{\partial x} = \phi[x]$$

Therefore,

$$\frac{\partial \Phi[Z_k]}{\partial c_k} = \frac{\phi[Z_k]}{\sigma_{X_k}}$$

and,

$$\frac{\partial L(\mathbf{C},\lambda)}{\partial c_k} = p_k - \lambda\left(\frac{\phi[Z_k]}{\sigma_{X_k}}\right)\prod_{l=1,l\neq k}^{K}\Phi[Z_l] \qquad k = 1,\ldots,K$$

The partial derivatives with respect to λ is

$$\frac{\partial L(\mathbf{C},\lambda)}{\partial \lambda} = \mathrm{g} - \prod_{k=1}^{K}\Phi[Z_k]$$

An optimal solution will satisfy the service level constraint with equality. To see this, consider a solution in which the service level constraint is not binding, i.e. the inequality is strict. The objective function value can be strictly decreased while still satisfying the constraint by decreasing the capacity of one of the stages. As the inequality is continuous and monotonically increasing in each of the stage capacities, there will be solutions for which the bound holds with equality. Therefore the objective function value of any non-binding solution is strictly greater than the objective function of some binding solution.

An optimal solution will satisfy the following set of equations,

$$\left.\frac{\partial L(\mathbf{C},\lambda)}{\partial c_k}\right|_{\mathbf{C}=\mathbf{C}^*,\lambda=\lambda^*} = p_k - \lambda^*\left(\frac{\phi[Z_k^*]}{\sigma_{X_k}}\right)\prod_{l=1,l\neq k}^{K}\Phi[Z_k^*] = 0 \qquad k = 1,\ldots,K \tag{1}$$

$$\left.\frac{\partial L(\mathbf{C},\lambda)}{\partial \lambda}\right|_{\mathbf{C}=\mathbf{C}^*,\lambda=\lambda^*} = g - \prod_{k=1}^{K}\Phi[Z_k^*] = 0 \tag{2}$$

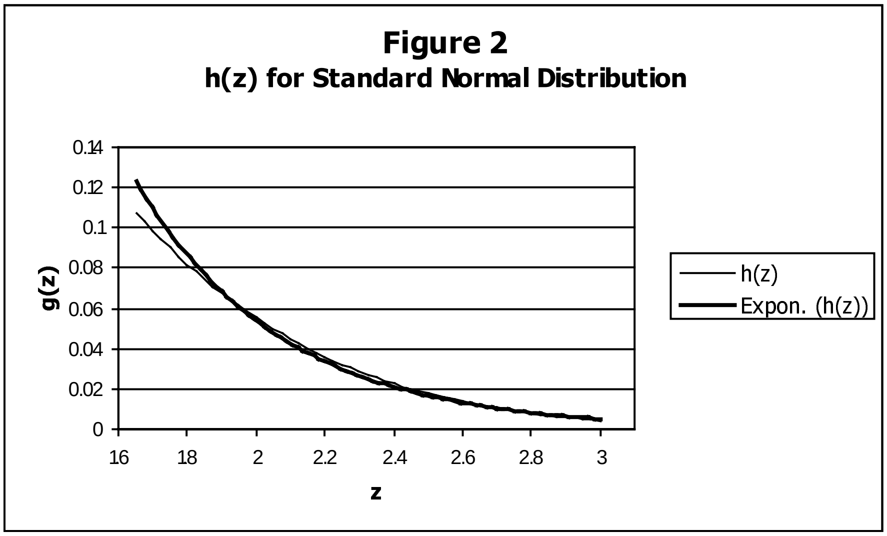
where,

$$Z_k^* = \frac{c_k^* - \mu_{X_k}}{\sigma_{X_k}}$$

Dividing through equation (1) by $\prod_{k=1}^{K}\Phi[Z_k^*]$ and applying (2), one obtains,

$$h(Z_k^*) = \frac{\sigma_{X_k} p_k}{\lambda^* g} \qquad k = 1,\ldots,K \quad \text{where } h(z) = \left(\frac{\phi[z]}{\Phi[z]}\right)$$

Figure 2 shows $h(z)$ for z greater than 1.65 (Φ[z]>0.95). It also shows an exponential function, $f(z)=ae^{-dz}$ with a=6.0322 and d=2.3579, curve fitted to $h(z)$. This curve was generated by Excel (the least squares fit) and has an R^2 value of 0.995.



As can be seen $h(z)$ is reasonably well approximated by the exponential function for large z values. If the desired service level is reasonably high, then the Z_k values will be high. Approximating $h(z)$ with the exponential function, $h(z)=ae^{-dz}$, then

$$ae^{-dZ_k^*} = \frac{\sigma_{X_k} p_k}{g\lambda^*}$$

$$\Rightarrow -dZ_k^* = \ln\left(\frac{\sigma_{X_k} p_k}{a\lambda^* g}\right) = -\ln\left(a\lambda^* g\right) + \ln\left(\sigma_{X_k} p_k\right)$$

$$\Rightarrow Z_k^* = \left(\frac{1}{d}\right)\ln\left(a\lambda^* g\right) - \left(\frac{1}{d}\right)\ln\left(\sigma_{X_k} p_k\right) \qquad k = 1,\ldots,K \tag{3}$$

and so the optimal protection levels, Z_k*, should be (approximately) a decreasing linear function of the natural log of the product of the stage demand standard deviation and capacity cost.

Substituting (3) into (2), one obtains the following equation for the optimal Lagrange multiplier,

$$\prod_{k=1}^{K}\Phi[Z_k^*]=\prod_{k=1}^{K}\Phi\left[\left(\frac{1}{d}\right)\ln\left(a\lambda^* g\right)-\left(\frac{1}{d}\right)\ln\left(\sigma_{X_k} p_k\right)\right]=g$$

The only unknown in this equation is λ* and as such the equation can be solved by a simple binary search over λ*. After solving for λ*, one can use (3) to obtain the approximately optimal stage service levels for **P1,Ind**. Note that the solutions are approximate as *h(z)* was approximated by an exponential function.

Alternatively, **P1,Ind** can be solved using mathematical programming software, such as Excel Solver. Φ[•] does not have a closed form solution and so an approximation (Lin, 1989) is used. This approximation is given by,

$$\Phi[z]=1-0.5*\exp\left(-0.717z-0.416z^2\right)\ ,\ z>0 \tag{4}$$

To ensure that z is greater than 0, **P1,Ind** is adapted slightly. Existing capacities are set equal to the mean stage demands rather than zero and the decision variables are the amount of capacity to add to each stage. This modification does not change the optimal value as a service level greater than 0.5 requires every stage to have a capacity greater than its mean demand.

Results for the actual Work Center A demand data are presented in Section 4.3.5. Eleven other test problem instances were generated for **P1,Ind**. The same stages as Work

Center A were used. Cost data (p_k) were generated from a truncated (non-negative) normal distribution with a mean of 5 and standard deviation of 2. Stage demand standard deviations were generated from a truncated (non-negative) normal distribution with a mean of 500 and standard deviation of 200, except for example 1 in which all standard deviations were set equal to 300. Mean stage demands were obtained by generating coefficient of variation data from a truncated (non-negative) normal distribution with a mean of 0.5 and standard deviation of 0.2. The data for the eleven examples can be found in Table 3. A service level of 0.95 was used for all eleven examples.

For each of the eleven examples, two problems were solved; one in which all twelve stages were considered and another in which only six stages (*A*-*F*) were considered. The second problem, in which only six stages were considered, was solved so to demonstrate that systems with fewer stages will have lower individual stage protection levels. The problems were solved using Excel Solver. Solution time was in the order of seconds. The optimal protection level results can be found in Table 5 in Section 4.3.4.

Table 3

		A	B	C	D	E	F	G	I	K	L	N	P
1	**mean**	548.5	992.9	458.9	377.6	1744.7	487.4	411.2	509.1	475.8	423.9	371.3	430.5
	std	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
	cost	4.4	2.4	5.5	7.6	7.4	8.5	0.6	4.5	7.2	2.8	3.6	1.6
2	**mean**	726.1	458.5	1987.9	535.4	5074.0	525.3	833.3	558.6	2452.7	975.9	1457.8	735.8
	std	290.3	315.8	519.8	291.4	539.4	107.1	480.8	432.2	712.2	611.4	733.2	546.4
	cost	5.5	3.0	6.5	7.9	1.7	6.2	7.3	5.9	6.3	7.1	8.1	7.0
3	**mean**	1478.7	1168.3	760.8	1678.6	255.1	644.9	214.5	555.7	1669.0	748.4	571.6	444.5
	std	804.5	598.6	582.4	933.0	70.6	365.3	116.9	428.6	523.6	447.2	356.6	339.9
	cost	4.0	6.9	2.6	5.4	1.1	2.0	5.8	7.7	2.9	6.3	5.0	7.4
4	**mean**	1394.2	1422.6	1321.7	1151.8	2304.2	2457.3	1459.4	1800.8	6019.6	2505.2	1112.8	365.6
	std	589.6	470.9	636.2	462.8	618.1	939.4	769.1	882.0	700.7	744.8	678.3	346.8
	cost	5.6	6.2	3.4	5.9	6.0	4.5	5.3	3.3	6.9	3.0	4.5	5.7
5	**mean**	385.9	1391.1	1092.2	1211.8	1974.5	2118.3	1646.9	628.8	626.7	731.4	1479.7	792.3
	std	313.8	590.6	526.0	723.5	434.8	256.6	672.3	492.2	312.8	449.8	752.9	322.4
	cost	2.9	3.2	5.2	2.9	5.4	1.1	4.8	4.3	7.1	6.1	7.3	5.5
6	**mean**	814.8	740.2	679.4	921.9	489.3	195.9	211.3	2280.3	333.7	1464.7	1280.5	909.8
	std	453.4	460.4	232.7	545.9	295.8	87.7	111.7	741.4	230.0	433.3	580.5	514.3
	cost	8.0	6.0	5.8	9.3	0.7	3.7	1.2	4.3	5.2	4.5	3.6	3.4
7	**mean**	169.4	1184.7	1303.7	676.7	378.8	1277.4	1032.8	626.1	501.4	1087.7	2081.9	600.4
	std	128.0	444.5	707.9	360.7	233.0	827.0	359.5	172.3	338.5	516.2	909.4	475.2
	cost	5.4	5.1	7.7	5.6	2.8	5.7	5.4	7.7	3.1	6.0	6.2	7.6
8	**mean**	1133.8	932.6	1189.0	2014.2	350.3	2113.9	875.7	3538.7	481.0	1381.4	1210.8	797.6
	std	876.2	643.8	539.5	383.5	144.0	655.6	321.3	504.3	244.1	576.5	510.2	441.9
	cost	4.2	3.3	4.8	4.0	2.7	3.8	5.3	4.9	1.2	3.0	6.1	9.5
9	**mean**	2043.8	739.8	398.9	290.4	861.4	837.1	1879.7	1241.1	638.4	1321.9	1049.2	5663.6
	std	493.0	342.0	288.5	210.6	458.7	562.1	776.2	883.6	509.1	572.0	660.5	774.1
	cost	5.9	4.7	6.4	4.6	6.2	9.4	7.7	8.8	7.0	7.4	6.8	3.5
10	**mean**	2239.7	1270.0	406.1	898.4	1418.1	449.4	1761.2	823.9	2376.6	2062.4	893.0	1447.1
	std	972.8	532.1	276.1	284.0	647.6	154.9	711.1	596.5	461.6	776.2	554.8	573.0
	cost	3.1	5.9	5.3	7.2	4.3	2.6	6.7	4.9	3.1	4.5	7.5	3.2
11	**mean**	1610.2	656.6	907.4	369.8	385.7	1204.9	880.0	1770.4	1569.8	2891.3	911.4	716.6
	std	688.0	548.4	673.5	286.6	281.9	763.6	532.4	516.3	668.9	910.8	492.6	424.7
	cost	8.1	4.2	4.8	6.0	2.2	1.2	4.1	7.8	5.0	6.1	5.1	4.1

Graphical results for Examples 5 are shown in Figure 3. The linear regression line for the individual stage protection level as a function of the natural log of cost times standard deviation is also shown.

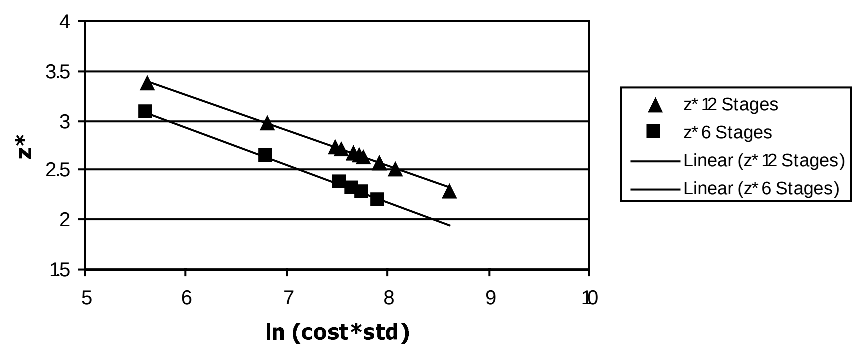


Figure 3
Protection Level as a function of Cost and Standard Deviation (Example 5)

As can be seen from Figure 3, the optimal protection levels are indeed linear in the natural log of capacity cost times standard deviation. The other examples all showed similar linear relationships. The R^2 values were above 0.99 for all problems. I conjecture that they do not equal 1.0 because of the approximation (eqn. (4)) used for the normal cumulative distribution function. When only six stages are considered, the protection levels are lower. This is expected as all demand is met only if all stages have sufficient capacity. The probability of there being sufficient capacity will decrease as the number of stages increases.

4.3.4 A Bonferroni Inequality

In the previous section, no use was made of the correlation between individual stage demands – they were assumed to be independent. Stage demands, as noted earlier, are in fact correlated. This correlation should alter the optimal capacity investments.

Consider a supply chain with K stages, each with a protection level of Z. If the stage demands are independent, then the service level is given by $(\Phi[Z])^K$. However if the stage demands are all perfectly positively correlated, then the service level is given by $\Phi[Z]$ which is larger than the service level if the stages are independent. It should therefore be expected that a positive correlation between two stage demands should lead to a reduction in the capacity required at these stages. This reduction should be most pronounced for highly positively correlated stages. As noted earlier, if flow path demand are independent (or positively correlated), then the stage demands will be positively correlated.

Closed form expressions do not exist for general multivariate normal probabilities. A Bonferroni inequality is used to obtain a lower bound on the probability constraint in **P1**. Theorem 7.1.4 of Tong (1980) gives the following inequality,

$$1-\sum_{k=1}^{K}\left(1-\Phi[Z_k]\right)+\max_{1\le l\le K}\left\{\sum_{k=1,k\neq l}^{K}L[Z_k,Z_l,\rho_{kl}]\right\}\le \mathbf{\Phi[Z,R]} \tag{5}$$

where $L[Z_k,Z_l,\rho_{kl}]$ is the probability that both Z_k and Z_l are greater than 0, a bivariate probability. The correlation coefficient is given by ρ_{kl}.

As a closed form expression for $L[Z_k,Z_l,\rho_{kl}]$ does not exist, an approximation (Cox and Wermuth, 1991) is used. This approximation is given by,

$$L[Z_k,Z_l,\rho_{kl}]=\Phi[-Z_l]\Phi[w(Z_k,Z_l,\rho_{kl})]$$

where,

$$w(Z_k, Z_l, \rho_{kl}) = \frac{\rho_{kl}\left(\phi(Z_l) \big/ (1-\Phi[Z_l])\right) - Z_k}{\sqrt{1-\rho_{kl}^2}}$$

So the lower bound in (5) can be approximated by,

$$1 - \sum_{k=1}^{K} (1-\Phi[Z_k]) + \Phi[-Z_K] \sum_{k=1}^{K-1} \Phi[w(Z_k, Z_K, \rho_{kK})] \leq \mathbf{\Phi[Z,R]} \tag{6}$$

where the maximum expression in equation (5) has been replaced by the K^{th} term. One could choose to replace the maximum with any of the K terms.

Using (6), an alternative formulation to **P1** can then be rewritten as

$$\text{Min } \{\sum_{k=1}^{K} p_k c_k\}$$

subject to

$$1 - \sum_{k=1}^{K} (1-\Phi[Z_k]) + \Phi[-Z_K] \sum_{k=1}^{K-1} \Phi[w(Z_k, Z_K, \rho_{kK})] \geq \text{g}$$

This formulation is referred to as **P1,Bonf** . Unlike **P1,Ind**, this formulation uses some correlation information. As noted above, there are K-1 other possible bounds that could be constructed by replacing the maximum expression in (5) with some $k<K$. Each of the possible bounds, k=1,…,K, uses the correlation coefficients of stage k with all other stages, i.e. $\rho_{1k},\ldots,\rho_{Kk}$. One could solve K different versions of **P1,Bonf**, with each version using a different bound, and then select the solution with the lowest capacity cost.

Using Lin's approximation for Φ[•], given in equation (4), Excel Solver can be used to solve **P1,Bonf**.

The same eleven test problems generated for **P1,Ind** are used with stage P being arbitrarily chosen as the K^{th} stage. As **P1,Bonf** requires correlation coefficients for the

K^{th} stage, the coefficients for stage *P*'s demand and other stage demands were generated from a discrete distribution with values 0.1, 0.3 and 0.9 having probabilities of 0.5, 0.4 and 0.1 respectively. The correlation coefficient data can be found in Table 4.

Table 4

		A	**B**	**C**	**D**	**E**	**F**	**G**	**I**	**K**	**L**	**N**	**P**
1	ρ_{xP}	0.3	0.1	0.3	0.9	0.1	0.3	0.3	0.3	0.3	0.3	0.9	1.0
2	ρ_{xP}	0.1	0.3	0.1	0.3	0.1	0.1	0.3	0.9	0.1	0.3	0.3	1.0
3	ρ_{xP}	0.1	0.1	0.1	0.3	0.1	0.1	0.3	0.1	0.3	0.1	0.9	1.0
4	ρ_{xP}	0.9	0.3	0.3	0.1	0.1	0.3	0.1	0.3	0.1	0.3	0.3	1.0
5	ρ_{xP}	0.9	0.3	0.3	0.9	0.1	0.1	0.1	0.1	0.3	0.1	0.1	1.0
6	ρ_{xP}	0.3	0.3	0.9	0.3	0.1	0.3	0.3	0.9	0.1	0.3	0.3	1.0
7	ρ_{xP}	0.3	0.1	0.3	0.1	0.3	0.9	0.9	0.9	0.3	0.3	0.3	1.0
8	ρ_{xP}	0.9	0.1	0.1	0.3	0.1	0.1	0.1	0.9	0.1	0.3	0.3	1.0
9	ρ_{xP}	0.1	0.1	0.3	0.1	0.1	0.9	0.1	0.1	0.1	0.3	0.9	1.0
10	ρ_{xP}	0.9	0.1	0.3	0.1	0.1	0.1	0.3	0.1	0.1	0.1	0.1	1.0
11	ρ_{xP}	0.9	0.3	0.1	0.9	0.1	0.3	0.3	0.1	0.1	0.3	0.3	1.0

Graphical results for Example 5 are presented in Figure 4. To illustrate the effect of demand correlation, the graph shows the protection levels for **P1,Ind**, **P1,Bonf**, and the linear regression for **P1,Ind**. The results for **P1,Ind** are the same as those shown in Figure 3 for the 12-stage problem.

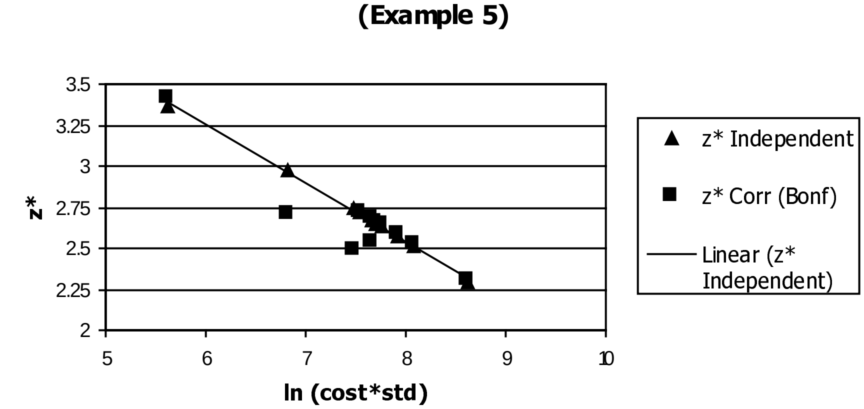


Figure 4
Effect of Correlation on Protection Level (Example 5)

As can be seen from Figure 4, the optimal protection levels for **P1,Bonf** are similar to those for **P1,Ind** except that some of the protection levels are lower. It turns out that stage *P* and the stages with high correlation coefficients (0.9) are those with the lower protection levels.

Table 5 presents the results for all eleven examples. 12 (or 6) (Indep) refers to the results obtained for **P1,Ind**. The number 12 (or 6) refers to how many stages were considered in the problem (see previous section). 12 (Bonf) refers to the results obtained for **P1,Bonf.** For the problems with twelve stages considered, I have highlighted (in bold typeface) the protection levels for those stages that had a correlation coefficient of 0.9 with stage *P* . The possible correlation coefficients are 0.1, 0.3 and 0.9. I have also highlighted the protection levels of stage *P*.

As can be seen, the greatest change in protection levels occur either for stage *P* or for those stages most highly correlated with stage *P*. On average over the eleven examples, the protection level of stage *P* decrease by 11.5% with a maximum of 18.9% and a minimum of 5.8%. For the stages with a correlation coefficient of 0.9, the average decrease is 6.7% with a maximum of 13.3% and a minimum of 2.4%. For all other stages, the average change in protection level is 0.6% with a maximum of 1.5% and a minimum of 0.0 %. Note that the protection levels both increase and decrease slightly for the stages with a correlation coefficient of 0.1 or 0.3. Therefore the percentages quoted are based on the absolute change in protection level to avoid positive and negative numbers canceling.

Table 5

		A	B	C	D	E	F	G	I	K	L	N	P
1	**12 (Indep)**	2.66	2.88	2.58	**2.45**	2.46	2.40	3.31	2.65	2.47	2.82	2.73	**3.02**
	12 (Bonf)	2.67	2.91	2.59	**2.39**	2.48	2.41	3.36	2.66	2.49	2.78	2.55	**2.45**
	6 (Indep)	2.51	2.74	2.42	2.29	2.30	2.24						
2	**12 (Indep)**	2.86	3.03	2.57	2.72	3.06	3.15	2.56	**2.68**	2.46	2.48	2.35	**2.53**
	12 (Bonf)	2.87	3.03	2.59	2.72	3.09	3.17	2.57	**2.49**	2.48	2.48	2.35	**2.37**
	6 (Indep)	2.41	2.61	2.08	2.25	2.62	2.75						
3	**12 (Indep)**	2.51	2.41	2.79	2.32	3.73	3.05	3.07	2.49	2.79	2.56	**2.73**	**2.60**
	12 (Bonf)	2.52	2.43	2.80	2.33	3.74	3.05	3.07	2.51	2.79	2.57	**2.56**	**2.45**
	6 (Indep)	2.29	2.18	2.59	2.09	3.58	2.86						
4	**12 (Indep)**	**2.62**	2.66	2.77	2.68	2.57	2.52	2.54	2.67	2.47	2.77	2.65	**2.81**
	12 (Bonf)	**2.52**	2.67	2.78	2.70	2.59	2.53	2.55	2.67	2.48	2.77	2.65	**2.55**
	6 (Indep)	2.37	2.42	2.54	2.44	2.32	2.26						
5	**12 (Indep)**	**2.97**	2.72	2.57	**2.67**	2.63	3.37	2.51	2.67	2.65	2.57	2.29	**2.74**
	12 (Bonf)	**2.71**	2.73	2.58	**2.54**	2.65	3.41	2.53	2.69	2.66	2.59	2.31	**2.48**
	6 (Indep)	2.64	2.36	2.19	2.30	2.26	3.06						
6	**12 (Indep)**	2.40	2.52	**2.78**	2.26	3.42	3.27	3.53	**2.46**	2.82	2.65	2.62	**2.69**
	12 (Bonf)	2.42	2.53	**2.52**	2.27	3.38	3.28	3.52	**2.34**	2.85	2.66	2.64	**2.32**
	6 (Indep)	2.19	2.31	2.60	2.04	3.23	3.11	3.11					
7	**12 (Indep)**	3.11	2.70	2.36	2.75	3.14	**2.42**	2.76	**2.90**	2.98	2.59	2.34	**2.53**
	12 (Bonf)	3.12	2.74	2.38	2.78	3.15	**2.28**	2.45	**2.51**	2.98	2.61	2.36	**2.20**
	6 (Indep)	2.90	2.46	2.08	2.51	2.92	2.15						
8	**12 (Indep)**	2.43	2.65	2.57	2.77	3.24	2.59	2.73	**2.59**	3.33	2.73	2.50	**2.38**
	12 (Bonf)	2.25	2.67	2.60	2.78	3.26	2.61	2.76	**2.32**	3.35	2.73	2.51	**2.13**
	6 (Indep)	2.16	2.39	2.31	2.52	3.03	2.33						
9	**12 (Indep)**	2.73	2.93	2.89	3.11	2.73	**2.50**	2.45	2.34	2.65	2.58	**2.56**	**2.75**
	12 (Bonf)	2.75	2.95	2.90	3.13	2.75	**2.37**	2.47	2.36	2.67	2.59	**2.41**	**2.34**
	6 (Indep)	2.34	2.58	2.53	2.77	2.35	2.08						
10	**12 (Indep)**	**2.57**	2.57	2.85	2.73	2.61	3.28	2.40	2.60	2.86	2.53	2.46	**2.77**
	12 (Bonf)	**2.48**	2.58	2.86	2.75	2.62	3.30	2.41	2.61	2.87	2.54	2.47	**2.53**
	6 (Indep)	2.24	2.23	2.55	2.41	2.28	3.02						
11	**12 (Indep)**	**2.37**	2.71	2.58	2.82	3.16	3.01	2.74	2.50	2.57	2.36	2.68	**2.82**
	12 (Bonf)	**2.30**	2.72	2.61	2.57	3.18	3.01	2.74	2.52	2.60	2.38	2.69	**2.37**
	6 (Indep)	2.02	2.41	2.27	2.53	2.91	2.77						

Stages with highly correlated demands can afford to have lower protection levels. This agrees with the hypothesized correlation effect discussed at the start of this section.

Two different approaches to solving the capacity problem **P1** have been developed. **P1,Ind** involves an approximation the service level probability bound by assuming independence. **P1,Bonf** uses some correlation information to replace the actual service level with a lower bound. I now use each of these formulations to analyze Work Center A. In doing so I compare the effectiveness of the two approaches.

4.3.5 Work Center A Results

The two formulations, **P1,Ind** and **P1,Bonf** developed above were used to study the capacity investment problem using the actual flow path demand data for Work Center A. Two problem instances were studied. In the first, Example 12, capacity costs were set equal to one and the objective was the minimization of total capacity. In the second, Example 13, a random set of capacity costs was generated. Table 6 contains the capacity cost data for both examples.

Table 6

Example	A	B	C	D	E	F	G	I	K	L	N	P
12	1	1	1	1	1	1	1	1	1	1	1	1
13	4.40	2.44	5.49	7.55	7.40	8.47	0.63	4.53	7.19	2.83	3.62	1.62

For each example, **P1,Ind** and **P1,Bonf** were both solved using a service level of 0.95 and a service level of 0.99 (95% and 99%).

In solving **P1,Bonf**, stage *P* is chosen as the reference stage in the lower bound (6) for the service level. This is not an arbitrary choice. As can be seen from Figure 1, all flow paths visit stage *P*. The demand on stage *P* should be correlated with demands for all other stages. Stages, such as *A*, which share a large number of flow paths with *P* will be highly correlated. To capture this correlation information, stage *P* is chosen. One could instead choose another stage such as *C* to be the reference stage. However, as *C* only shares flow paths with *A* and *P*, stage *C* demand is not correlated with most other stage demands. Therefore a lot of the system correlation information is not captured if stage *C* is selected. The effect of choosing stage *C* is discussed at the end of this section.

Table 7 presents the optimal protection levels for the different problems and the percentage difference (reduction) in service level from **P1,Ind** to **P1,Bonf**. It also presents the objective values and the percentage difference (reduction) in objective values.

Table 7

	Service Level		A	B	C	D	E	F	G	I	K	L	N	P	*Obj. Value*
12	**95%**	**Ind**	2.20	3.09	2.64	2.30	3.43	2.75	3.91	2.67	3.14	3.47	3.69	2.22	*17620.8*
		Bonf	2.01	3.15	2.64	2.03	3.46	2.79	3.96	2.66	3.16	3.52	3.74	1.86	*16962.8*
		% diff	8.7%	-1.9%	-0.1%	11.6%	-0.8%	-1.4%	-1.1%	0.2%	-0.5%	-1.3%	-1.3%	16.4%	*3.7%*
	99%	**Ind**	2.77	3.58	3.15	2.86	3.84	3.24	4.29	3.16	3.56	3.88	4.14	2.84	*19471.3*
		Bonf	2.65	3.65	3.20	2.68	3.95	3.33	4.37	3.20	3.60	4.03	4.22	2.53	*19026.7*
		% diff	4.4%	-2.0%	-1.5%	6.4%	-3.1%	-2.8%	-1.8%	-1.5%	-1.4%	-3.9%	-2.0%	10.7%	*2.3%*
13	**95%**	**Ind**	2.25	3.31	2.60	2.11	3.30	2.54	4.43	2.69	2.98	3.63	3.79	2.57	*78331.8*
		Bonf	2.07	3.35	2.60	2.00	3.33	2.57	4.43	2.67	3.02	3.67	3.83	1.95	*76287.6*
		% diff	7.9%	-1.1%	0.0%	5.1%	-0.9%	-1.0%	-0.1%	0.5%	-1.2%	-1.0%	-0.9%	24.2%	*2.6%*
	99%	**Ind**	2.82	3.74	3.12	2.72	3.73	3.07	4.78	3.20	3.45	4.04	4.18	3.09	*86688.8*
		Bonf	2.69	3.78	3.15	2.64	3.77	3.12	4.83	3.23	3.49	4.07	4.22	2.64	*85375.1*
		% diff	4.5%	-0.9%	-1.2%	2.8%	-0.9%	-1.4%	-1.1%	-0.9%	-1.2%	-0.9%	-0.9%	14.7%	*1.5%*

The protection levels for the highly correlated stages, A, D and P, are significantly reduced when the correlation is considered. This is consistent with the results obtained for the eleven examples in the previous section. The protection levels for the other stages increase slightly. I conjecture that this slight increase is an effect of the lower bounds and approximations used.

These results are presented in graphical form in Figures 5-8.

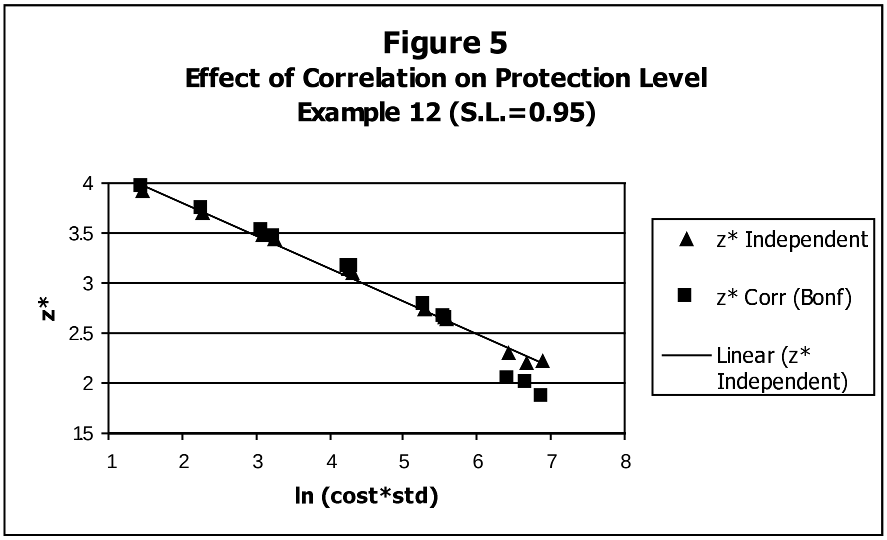


Figure 5
Effect of Correlation on Protection Level
Example 12 (S.L.=0.95)

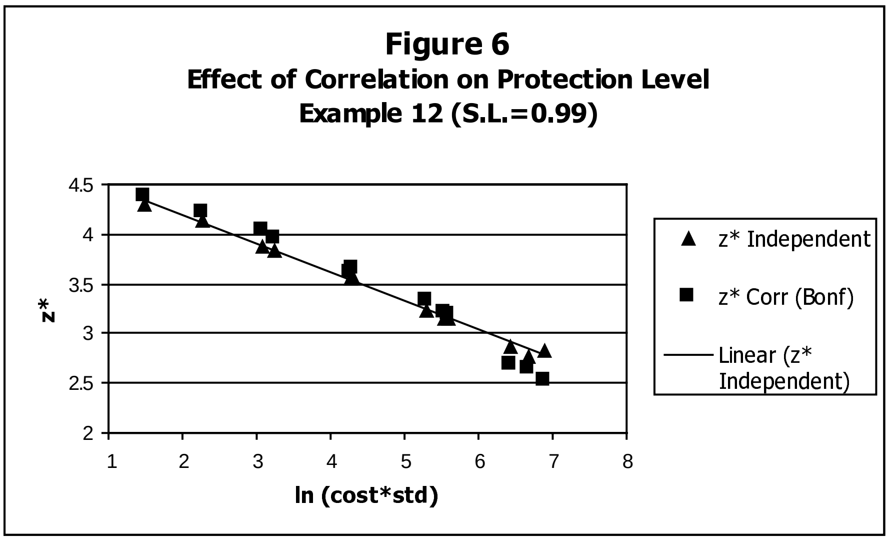


Figure 6
Effect of Correlation on Protection Level
Example 12 (S.L.=0.99)

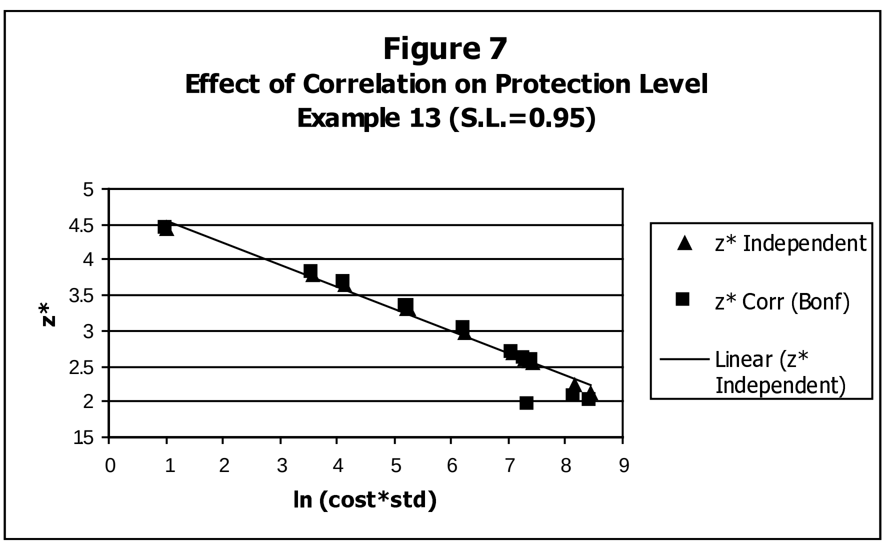
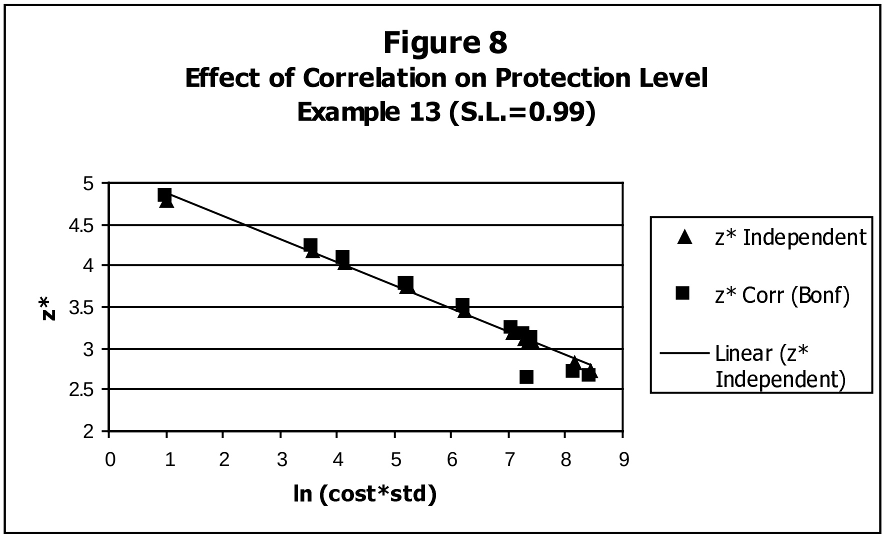


Figure 7
Effect of Correlation on Protection Level
Example 13 (S.L.=0.95)



Again one sees the approximately linear relationship between the protection level and the natural log of cost times standard deviation for **P1,Ind**. As mentioned above, when correlation is considered the protection levels are altered somewhat, with the protection level for highly correlated stages being reduced.

One possible capacity investment policy is to assume that stage demands are independent and provide an equal protection level for all stages. If the service level target is *g*, then the protection level for each stage is given by,

$$Z_k^* = \Phi^{-1}\left[g^{1/K}\right] \quad k = 1,\ldots,K$$

This policy is used as a benchmark against which to compare the results for **P1,Ind** and **P1,Bonf**. Table 8 presents the results for 90%, 95% and 99% service levels. As can be seen, both formulations outperform the equal protection policy, and **P1,Bonf** outperforms **P1,Ind**.

Table 8

Example	Service Level		Simulated Service Level	% Reduction in Total Cost over Equal Protection Policy	Actual Reduction in Total Cost over Equal Protection Policy
12	90%	**Equal**	91.9%	-	-
		Ind	94.4%	4.9%	864.1
		Bonf	89.8%	9.3%	1626.3
	95%	**Equal**	95.6%	-	-
		Ind	97.2%	4.2%	778.5
		Bonf	95.5%	7.8%	1436.5
	99%	**Equal**	99.4%	-	-
		Ind	99.3%	3.1%	626.0
		Bonf	99.2%	5.3%	1070.6
13	90%	**Equal**	91.9%	-	-
		Ind	94.1%	4.9%	3811.8
		Bonf	90.5%	8.0%	6244.5
	95%	**Equal**	95.6%	-	-
		Ind	97.1%	4.2%	3459.6
		Bonf	95.7%	6.7%	5503.8
	99%	**Equal**	99.4%	-	-
		Ind	99.4%	3.1%	2815.4
		Bonf	99.2%	4.6%	4129.1

The service level lower bound in **P1,Bonf** uses correlation data that is not used in the service level approximation in **P1,Ind**. As such one might expect the actual service level to be more closely attained by **P1,Bonf**. Simulation is used to estimate the actual service levels obtained by the various policies. Using Excel, twenty thousand flow path demand vector realizations were generated. Each vector comprised of a realization for each flow path. The frequency of shortfall was calculated and a simulated service level obtained. In all cases the simulated service levels are equal to or higher than the target service level(1).

(1) A simulated service level of 89.8% is obtained for the **P1,Bonf** solution for the 90% service level target.

P1,Bonf gives simulated service levels closer to the target than **P1,Ind**. This is one possible reason for the objective value being lower in **P1,Bonf**.

The simulated service levels for **P1,Bonf** are quite close to those obtained by an equal fractile policy; yet the capacity costs is significantly lower. This suggests that an equal protection policy is not a very effective one. A better policy can be obtained by considering the capacity costs, demand standard deviations and demand correlations.

The percentage reduction in total cost over the equal fractile policy increases as the service level decreases. One possible explanation for this is that the improvement (or cost reduction) values are quoted relative to the capacity cost required by an equal fractile policy. As the service level decreases from 99%, the capacity cost required by an equal fractile policy decreases rapidly. Thus even if the actual reduction in capacity cost decreases, the relative reduction may increase.

However, Table 8 shows that the actual reduction in capacity cost increases as the service level decreases. Therefore the increasing relative improvement over the equal fractile policy as the service level decreases is not an artifact of a decreasing denominator.

As noted earlier, stage *P* was selected as the reference stage for the lower bound (6) in **P1,Bonf**. The Work Center A examples were re-solved using stage *C* as the reference stage in the lower bound. Stage *C* demand is correlated only with the demands on stage *A*, *F* and *P*, whereas stage *P* demand is correlated with all stage demands. Given that **P1,Bonf** uses the correlation information for the reference stage, one might expect **P1,Bonf** to perform better when *P* is chosen as the reference stage. Table 9 shows this to be true.

Table 9

Example	Service Level	Reference Stage	Simulated Service Level	% Reduction in Total Cost over Equal Protection Policy
12	**90%**	**P**	89.8%	9.3%
		C	95.0%	4.8%
	95%	**P**	95.5%	7.8%
		C	97.4%	4.2%
	99%	**P**	99.2%	5.3%
		C	99.2%	3.2%
13	**90%**	**P**	90.5%	8.0%
		C	94.0%	4.8%
	95%	**P**	95.7%	6.7%
		C	97.4%	4.2%
	99%	**P**	99.2%	4.6%
		C	99.4%	3.2%

In every case, the percentage reduction in total cost relative to an equal fractile policy is greater when stage *P* is chosen as the reference than when stage *C* is chosen. One reason this improvement is that choosing stage *P* appears to give a tighter bound on the service level. As can be seen from Table 9, the simulated service levels are nearer the target when *P* is chosen. This does not explain all of the improvement. For the 99% target case in Example 12, the simulated service level is 99.2% for both *P* and *C* but *P* attains this service level at a lower total cost.

It is therefore recommended that when using **P1,Bonf,** one selects the reference stage to be one whose demand is correlated with a large number of other stage demands.

4.4 Expected Shortfall Criterion

When investing in capacity, a firm may decide to use an expected shortfall criterion rather than a service level criterion. That is, the capacity at each stage should be such that the expected shortfall in a period is less than or equal to some target T.

This criterion supposes that a production planning objective of minimizing total supply chain shortfall is used. The supply chain shortfall is the sum of the individual flow path shortfalls. Production planning is assumed to occur after demand is realized and the production planning problem, for a given capacity vector **C**, is given by the following linear program, **P2(C)**,

$$\text{Min}\ \{sf_{SC} = \sum_{f=1}^{F} sf_f\}$$

subject to

$$1.\ y_f + sf_f \geq d_f \qquad f \in F$$

$$2.\ \sum_{f \in P(k)} \beta_{fk} y_f \leq c_k \quad k = 1,\ldots,K$$

$$3.\ \mathbf{y}, \mathbf{sf} \geq \mathbf{0}$$

where y_f is the production of flow path *f* and sf_f is the shortfall of flow path *f*, and sf_{SC} is the total supply chain shortfall.

While all parameters are deterministic for the production planning problem, the demands are stochastic when the capacity vector is being determined. So for a given capacity vector, **C**, the expected total shortfall, $E_{\mathbf{D}}[SF_{SC}(\mathbf{D},\mathbf{C})]$, is given by the expected objective value of the production planning problem.

The capacity problem can be formulated as a two stage stochastic program with recourse, where the recourse function for a given capacity vector is the production planning problem, **P2(C)**. The stochastic program, **SP1**, is then given by,

$$\text{Min}\,\{\sum_{k=1}^{K} p_k c_k\}$$
subject to
$$E[SF_{SC}] \le \text{T}$$

The expected shortfall upper bound is assumed to be sufficiently low as to require a positive capacity at all stages.

The Lagrangian relaxation of this problem can be mapped into the multiple-resource investment problem formulated by Harrison and Van Mieghem(1999). Using Lemma 1 in that paper, the optimal capacity addition vector must satisfy,

$$c_k{*} > 0 \Rightarrow \nabla_k E[SF_{SC}(\mathbf{D},\mathbf{C*})] = \lambda * p_k$$

where $\nabla_k E[SF_{SC}(\mathbf{D},\mathbf{C*})]$ is the derivative of the expected total shortfall with respect to stage k capacity at the optimal capacity vector and λ* is the optimal Lagrange multiplier.

This tells us that an optimum capacity addition vector is such that the marginal values of increasing capacity, defined as the change in expected shortfall divided by the capacity cost, at each stage should be equalized. In some sense, the supply chain capacities should be "balanced" where "balance" is defined in terms of marginal capacity values.

While it is important to understand the structure of the optimal policy it is also useful to be able to solve the capacity problem, **SP1**, for actual problem instances. Harrison and Van Mieghem (1999) provide an exact solution technique suitable for very small problems and give a two-product three-stage example. However the method is not suitable for even moderately sized problems as it requires the evaluation of multivariate normal probabilities which is prohibitively difficult as the problem sizes grow in the number of stages.

This section develops an alternative solution method suitable for certain supply chains. The supply chain used for problem instances in this section is described next.

4.4.1 The Alcalde Job Shop

Alcalde (1997) describes a job shop shown in Figure 9 below. This job shop is referred to as the Alcalde job shop in this dissertation.

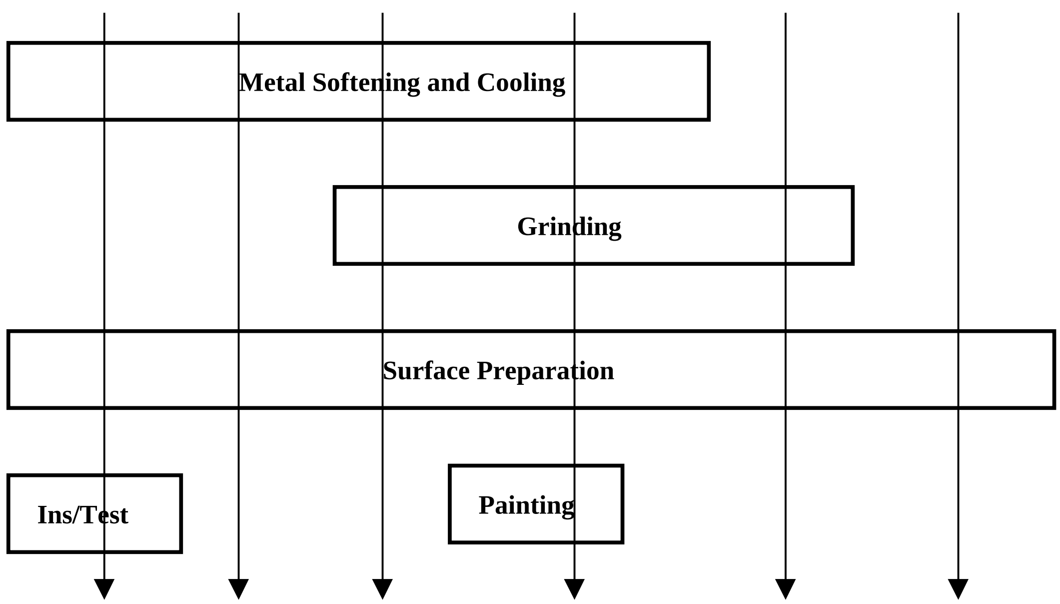


Figure 9: Alcalde Job Shop

As before, boxes denote production stages and lines denote flow paths – a flow path is defined by the stages the line passes through.

Demand data were not available for the Alcalde job shop so the flow path demands presented in Table 10 were chosen and used for all the examples solved in this section.

Table 10

Flow Path	Mean	St. Dev
MSC,SP,I/T	650	195
MSC,SP	2500	750
MSC,G,SP	900	270
MSC,G,SP,P	600	180
G,SP	300	90
SP	1200	360

As will be discussed in Section 4.4.3, the stochastic program developed is not useful for the Work Center A supply chain.

4.4.2 Independent Stage Demands

As discussed in Section 4.3, a pair of stage demands is correlated if there is one or more flow paths that require processing at both stages. If each stage processes only products unique to that stage, then the stage demands would be independent. As in Section 4.3, I first assume that stage demands are independent to develop some insight into the expected shortfall criterion. The results and formulations developed in this section are applicable to all supply chains characterized by the general model described in Section 4.2.

If each stage processes a unique set of products, then the production planning problem is greatly simplified, each stage produces the minimum of its total demand and capacity. For a capacity of c_k, the expected shortfall for stage k is then given by,

$$\int_{c_k}^{\infty}(x_k - c_k)h_k(x_k)dx_k$$

where $h_k(x_k)$ is the demand distribution for stage k. The stochastic program, **SP1**, is then given by,

$$\text{Min}\ \{\sum_{k=1}^{K} p_k c_k\}$$

subject to

$$\sum_{k=1}^{K}\left[\int_{c_k}^{\infty}(x_k - c_k)h_k(x_k)dx_k\right] \le \mathrm{T}$$

This formulation is referred to as **SP1,IND**. The Lagrangian relaxation of this problem is given by,

$$L(\mathbf{C},\lambda) = \sum_{k=1}^{K} p_k c_k + \lambda\left(T - \sum_{k=1}^{K}\left[\int_{c_k}^{\infty}(x_k - c_k)h_k(x_k)dx_k\right]\right)$$

and the partial derivatives with respect to λ and c_k are,

$$\frac{\partial L(\mathbf{C},\lambda)}{\partial \lambda} = T - \sum_{k=1}^{K}\left[\int_{c_k}^{\infty}(x_k - c_k)h_k(x_k)dx_k\right]$$

$$\begin{aligned}\frac{\partial L(\mathbf{C},\lambda)}{\partial c_k} &= p_k + \lambda\int_{c_k}^{\infty} h_{X_k}(x_k)dx_k \\ &= p_k + \lambda(1 - H_k[c_k]) \quad k = 1,\ldots,K\end{aligned}$$

where $H_k[\bullet]$ is the cumulative distribution function of stage k demand.

An optimal solution will satisfy the expected shortfall constraint with equality. To see this, consider a solution in which the service level constraint is not binding, i.e. the inequality is strict. The objective function value can be strictly decreased by decreasing the capacity of one of the stages while still satisfying the constraint. As the inequality is continuous and monotonically increasing in each of the stage capacities, there will be solutions for which the bound holds with equality. Therefore the objective function value of any non-binding solution is strictly greater than the objective function of some binding solution.

Therefore the partial derivatives will equal zero in an optimal solution and the solution will satisfy the following set of equations,

$$\sum_{k=1}^{K}\left[\int_{c_k^*}^{\infty}(x_k - c_k^*)h_k(x_k)dx_k\right] = T \tag{7}$$

$$H_k[c_k^*] = 1 + \frac{p_k^*}{\lambda^*} \quad k = 1,\ldots,K \tag{8}$$

So at an optimal capacity vector, the probability that a stage's demand exceeds its capacity is a linear function of the capacity cost. As λ^* is negative, the probability is a decreasing function of cost.

SP1,Ind can be solved using mathematical programming software, such as Excel Solver. If stage demands are normally distributed, then $H_k[c_k]=\Phi[Z_k]$ where,

$$Z_k = \frac{c_k - \mu_{X_k}}{\sigma_{X_k}}$$

and,

$$\begin{aligned}\int_{c_k}^{\infty}(x_k - c_k)h_k(x_k)dx_k &= \sigma_{X_k}\phi(Z_k) + (\mu_k - c_k)(1 - \Phi[Z_k]) \\ &= \sigma_{X_k}\phi(Z_k) - \sigma_{X_k}Z_k(1 - \Phi[Z_k])\end{aligned} \tag{9}$$

where (9) is obtained using equation (4.15) of Hadley and Whitin (1963). **SP1,IND** can then be expressed using equation (9) in the constraint. As in Section 4.3, Lin's approximation, equation (4), is used for $\Phi[Z_k]$.

It should be noted that in the case of equal capacity costs, the optimal solution to **SP1,IND** can be obtained without using an optimization package. From equation (8), the probability that a stage's demand exceeds its capacity will be the same for all stages. That is $\Phi[Z_k*]$ is the same for all k=1,…,K, and therefore $Z_k*=Z*$ for all k=1,…,K. Using equations (7) and (9), then

$$\left[\phi(Z^*) - Z_k^*\left(1 - \Phi[Z^*]\right)\right]\sum_{k=1}^{K}\sigma_{X_k} = T \tag{10}$$

Equation (10), and thus equal capacity cost problems, can be solved using a simple line search technique.

Five examples for the Alcalde job shop were analyzed. In each example the demand data presented earlier in Table 10 are used. The cost data for the five examples are presented in Table 11.

Table 11

Example	MS&C	Grinding	Surf. Prep.	Ins/Test	Painting
1	1.00	1.00	1.00	1.00	1.00
2	0.88	0.60	1.10	1.40	1.38
3	1.12	1.46	1.46	1.09	0.74
4	0.96	0.94	1.41	0.84	1.29
5	1.00	1.00	3.00	1.00	1.00

Example 1 has all the capacity costs equal. Example 2,3 and 4 have the costs randomly generated from a uniform distribution over the range [0.5,1.5]. In Example 5, all the stages except for Surface Preparation have equal capacity costs. Surface Preparation has a significantly higher cost.

Each problem example was solved using Excel Solver for six different expected shortfall bounds, ranging from 100 to 600. Solution time was in the order of seconds.

As noted above, because capacity costs are equal, the optimal solution to Example 1 should be such that Φ[Z*] are equal for all stages. Table 12 presents the Φ[Z*] results for Example 1. It also presents the analytical Φ[Z*] obtained by solving equation (10) above. As can be seen, the Solver results agree very closely with the analytical solution.

Table 12

	Excel Solver Φ[*Z**]					Analytical Φ[*Z**]
Expected Shortfall Upper Bound	**MS&C**	**Grinding**	**Surf. Prep.**	**Ins/Test**	**Painting**	
100	0.924	0.924	0.925	0.924	0.924	**0.925**
200	0.864	0.864	0.864	0.864	0.864	**0.867**
300	0.812	0.812	0.812	0.812	0.812	**0.816**
400	0.765	0.765	0.765	0.765	0.765	**0.769**
500	0.723	0.723	0.723	0.723	0.723	**0.727**
600	0.684	0.684	0.684	0.684	0.684	**0.688**

Table 13 presents the Excel Solver Φ[Z*] results for examples 2-5. The linear relationship between Φ[Z*] and the capacity cost is shown in Figures 10-13. Stages with higher capacity costs are given lower protection levels, with the probability of demand exceeding capacity decreasing linearly in the capacity cost.

Table 13

		Φ(z*)				
Example	**Expected Shortfall Upper Bound**	**MS&C**	**Grinding**	**Surf. Prep.**	**Ins/Test**	**Painting**
2	100	0.933	0.955	0.917	0.894	0.895
	200	0.879	0.918	0.851	0.812	0.814
	300	0.833	0.884	0.795	0.742	0.745
	400	0.792	0.855	0.745	0.679	0.683
	500	0.755	0.829	0.700	0.621	0.625
	600	0.722	0.806	0.659	0.564	0.569
3	100	0.935	0.915	0.915	0.937	0.958
	200	0.882	0.847	0.848	0.885	0.922
	300	0.837	0.789	0.790	0.841	0.891
	400	0.796	0.738	0.739	0.801	0.863
	500	0.760	0.691	0.692	0.766	0.838
	600	0.727	0.648	0.649	0.734	0.815
4	100	0.938	0.940	0.909	0.946	0.916
	200	0.887	0.890	0.837	0.901	0.850
	300	0.844	0.847	0.775	0.862	0.793
	400	0.805	0.809	0.721	0.827	0.742
	500	0.770	0.775	0.671	0.797	0.697
	600	0.739	0.745	0.624	0.769	0.655
5	100	0.961	0.961	0.881	0.961	0.961
	200	0.929	0.929	0.792	0.929	0.929
	300	0.900	0.900	0.717	0.900	0.901
	400	0.876	0.876	0.651	0.876	0.876
	500	0.855	0.856	0.590	0.855	0.855
	600	0.838	0.838	0.534	0.838	0.838

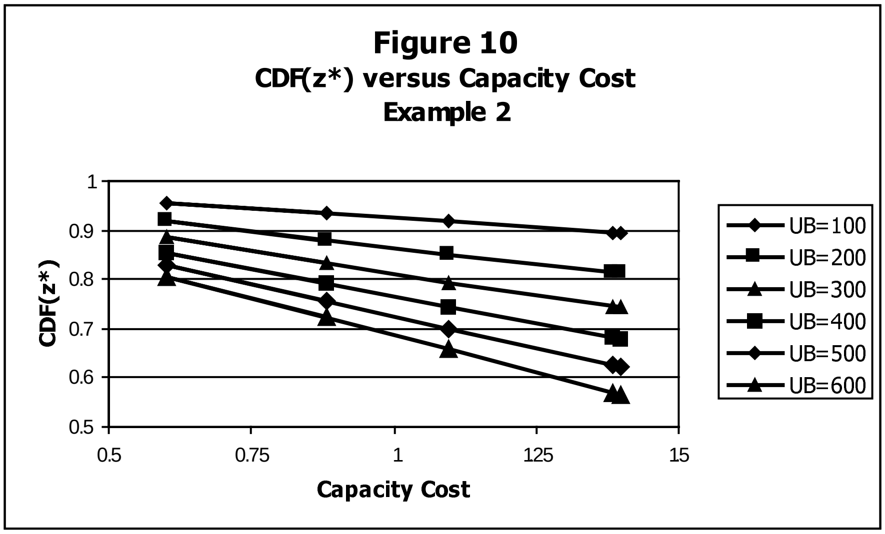


Figure 10
CDF(z*) versus Capacity Cost
Example 2

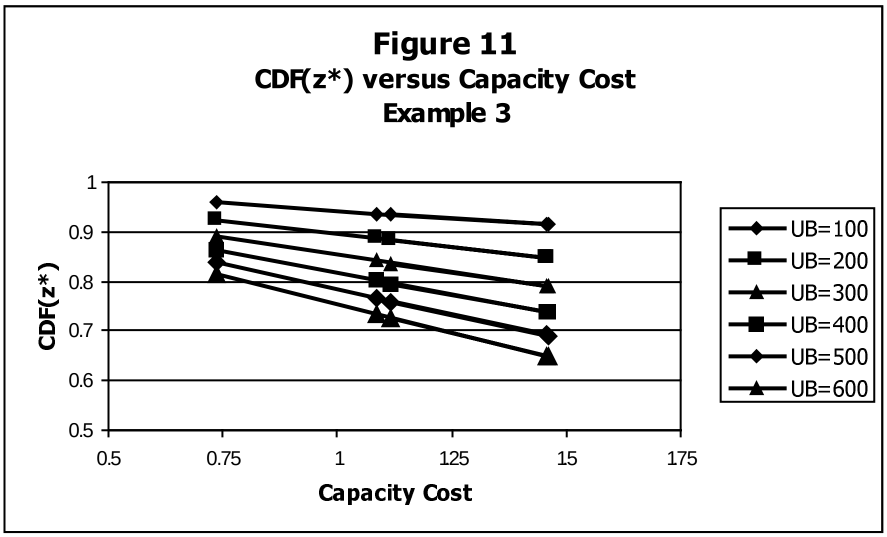


Figure 11
CDF(z*) versus Capacity Cost
Example 3

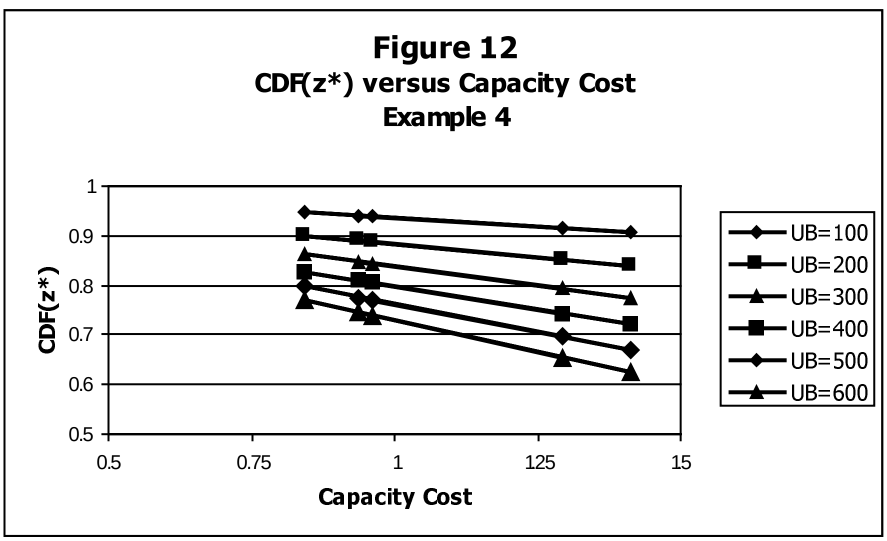
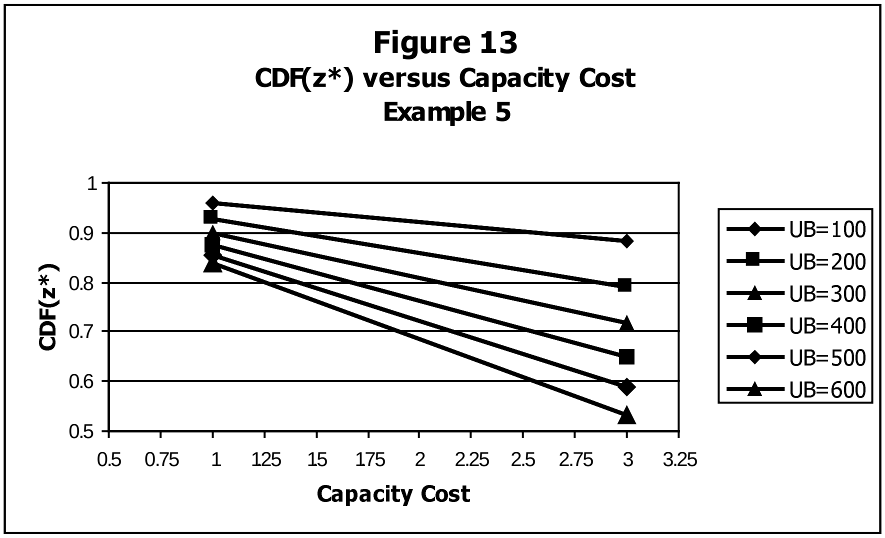


Figure 12
CDF(z*) versus Capacity Cost
Example 4



4.4.3 A Scenario Based Stochastic Program

In the previous section, stage demands were assumed to be independent. However, this is not the case. As stages can process similar flow paths, the stage demands will be correlated. In this section, a scenario based stochastic program that accounts for this correlation is developed.

The idea behind the formulation is to use a set of scenarios to specify the flow path demand distributions. Clearly the demand distributions will be better modeled as the number of scenarios increases. For certain supply chains, a formulation is developed that enables a large number of demand scenarios to be considered.

A scenario based version of the stochastic program **SP1**, can be obtained by simply defining a set of scenarios, with each scenario, s, $s=1,\ldots,S$ specifying a possible vector of flow path demands. The stochastic program is then given by,

$$\text{Min}\ \{\sum_{k=1}^{K} p_k c_k\}$$

subject to

$$1.\ (1/\text{S})\sum_{s=1}^{S}\sum_{f\in F} sf_f^s \le T$$

$$2.\ y_f^s + sf_f^s \ge d_f \qquad f \in F \quad s = 1,\ldots,S$$

$$3.\ \sum_{f\in P(k)} \beta_{fk} y_f^s \le c_k \quad k = 1,\ldots,K \quad s = 1,\ldots,S$$

$$4.\ \mathbf{c}, \mathbf{y}, \mathbf{sf} \ge \mathbf{0}$$

The first constraint ensures that the average total shortfall, across the S scenarios, is less than or equal to the upper bound. Constraints 2 and 3 are those given in the production planning linear program **P2**. This stochastic program is thus simply a scenario based program in which the production planning linear program **P2** is used as the recourse function for each scenario. This program is referred to as **SP1(P2)**.

The number of decision variables used in **SP1(P2)** is equal to K+2$S|F|$, where $|F|$ is the number of flow paths. For the Alcalde job shop, K=5 and $|F|$=6. Therefore the number of variables is 5+12S. Excel Solver can only accept 200 decision variables. Therefore, the maximum number of scenarios that can be used with Excel Solver is 16. As the number of flow paths increases, the number of scenarios that Excel Solver can handle becomes quite small and the ability to model the actual probability distributions decreases. An alternative stochastic program is developed for certain supply chain types in which the number of scenarios allowed is significantly increased.

The key to this alternative stochastic program lies in an alternative formulation to the production planning linear program **P2** that can be developed if all the capacity usages β_{fk} are equal for each stage. That is, for each stage k, $\beta_{fk}=\beta_k\ \forall f\in P(k)$. If this is the case, the stage capacities can be written as $c_k^{\text{new}}=c_k/\beta_k$. I assume from here on that all capacities are expressed in this manner and suppress the “new” superscript. The linear program **P2** is then given by,

$$\text{Min}\,\{sf_{SC} = \sum_{f=1}^{F} sf_f\}$$

subject to

$$1.\; y_f + sf_f \geq d_f \qquad f \in F$$

$$2.\; \sum_{f \in P(k)} y_f \leq c_k \quad k = 1,\ldots,K$$

$$3.\; \mathbf{y}, \mathbf{sf} \geq \mathbf{0}$$

This linear program is referred to as **P3**. A lower bound on the minimum total shortfall is given in Lemma 1 below. This lemma also gives a sufficient condition on the supply chain so that the minimum shortfall actually equals this lower bound.

Define the path-stage matrix **B** of a supply chain to be the matrix in which there is a row for each flow path f, a column for each stage k and b_{fk}=1 if $k \in Q(f)$ and b_{fk}=0 if $k \notin Q(f)$. In other words, b_{fk}=1 if flow path f requires stage k and 0 otherwise. Let Λ denote a subset of the stages k=1,…,K. Let $P(\Lambda)$ denote the set of flow paths processed by any stage $k \in \Lambda$.

Lemma 1:

(i) A lower bound on the minimum total shortfall in problem ***P3*** *is given by*

$$\max_{\Lambda \subseteq \{1,\ldots,K\}} \{ \sum_{f \in P(\Lambda)} d_f - \sum_{k \in \Lambda} c_k \}$$

(ii) If the path-stage matrix ***B*** *is totally unimodular (TU), then the minimum total shortfall in problem* ***P3*** *equals the lower bound in (i).*

The lower bound in (i) does not equal the actual minimum total shortfall in all supply chains. Consider the supply chain represented by the following path stage matrix

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

In this supply chain[(2)], the flow path *A* requires processing at stages 1 and 2, flow path *B* at stages 2 and 3, and flow path *C* at stages 1 and 3. Let each stage have a capacity of c. If the supply chain produces c units of flow path *A*, then it can not produce any units of the flow paths *B* and *C*. Stage 3's capacity is not used at all. A similar phenomenon occurs if the supply chain produces c units of either flow path *B* or *C*. The maximum total production that this supply chain can achieve is $3c/2$ and occurs when the supply chain produces $c/2$ units of each product. In this case, all three stages produce to capacity. Let the demand for each of the products be c, then the lower bound of Lemma 1(i) is c. The minimum total shortfall is $3c/2$ as the maximum total production of $3c/2$ can be attained by producing $c/2$ units of each product. Therefore, for this supply chain the lower bound in Lemma (i) is strictly less than the minimum total shortfall.

Lemma 1(ii) gives a sufficient condition on the supply chain path-stage matrix for the lower bound to equal the minimum total shortfall, namely that the matrix be totally unimodular (TU). Polynomial time algorithms exist for testing total unimodularity of a

[(2)] This supply chain is closely related the supply chain of Figure 3 in Chapter 3. This is not coincidental as the classes of supply chains being considered are closely related. Chapter 3 considers supply chains in which every product visits every stage with stages being partially flexible. Chapter 4 considers supply chains in which a product visits a subset of stages with stages being totally flexible. Lemma 1 in Chapter 4 can be thought of as analogous to Lemma 1 of Chapter 3.

matrix (Schrijver, 1987). The question arises as to whether this condition is so restrictive that very few supply chains satisfy it. The following lemma suggests that this is not the case.

Lemma 2:

The path-stage matrices for the Alcalde Job Shop and for Work Center A are both totally unimodular (TU).

Therefore both supply chains studied in Chapter 4 satisfy the TU condition. Obviously one can not generalize from these two supply chains to state that most real world supply chains satisfy the TU condition. However given Lemma 2, one might conjecture that there is a reasonable number of supply chains satisfying the TU condition.

In the remainder of this section, I restrict attention to those supply chains with totally unimodular path-stage matrices. For these supply chains, the minimum total shortfall is given by the expression

$$\max_{\Lambda \subseteq \{1,\ldots,K\}} \{ \sum_{f \in P(\Lambda)} d_f - \sum_{k \in \Lambda} c_k \}$$

The subsets over which this maximization is evaluated can be restricted as shown be the following lemma.

Lemma 3:

Any subset Λ that can be partitioned into two disjoint subsets Λ_m and Λ_n such that $P(\Lambda_m) \subseteq P(\Lambda_n)$, can be omitted from the set of subsets over which the maximum in Lemma 1 is evaluated.

So, one can restrict the maximization in Lemma 1 to the stage subsets Λ that cannot be partitioned in this manner. This maximization is referred to as **P4**.

For a simple illustration, consider a system with two stages, *A* and *B*, and two flow paths, *A* and *AB*. The complete set of stage subsets is $\{\{\varnothing\},\{A\},\{B\},\{A,B\}\}$. However the subset $\{A,B\}$ cannot be the maximizing set in Lemma 1, as

$$d_A + d_{AB} - c_A - c_B < d_A + d_{AB} - c_A$$

In some sense the stage subset $\{AB\}$ double counts the capacity available for flow path *AB* as this flow path cannot use the sum of the available capacity in *A* and *B* but only the maximum of the available capacity of *A* and *B*. Therefore the maximum is never achieved by the stage subset $\{AB\}$.

Denoting the possible maximizing sets in **P4** as $\Lambda_1,\ldots,\Lambda_L$, the optimum value to **P3** can then be expressed as,

$$\max_{l=1,\ldots,L} \{ \sum_{f \in P(\Lambda_l)} d_f - \sum_{k \in \Lambda_l} c_k \}$$

An algorithm for determining the sets $\Lambda_1,\ldots,\Lambda_L$ can be found in Appendix 3.

This is a simple maximization over a set of numbers and thus can be evaluated much more simply than **P3** can be solved. A linear program package is not even required. For the Alcalde job shop the set of possible stage subsets is {Metal Softening and Cooling (MS&C), Grinding, Surface Preparation (SP), Inspect/Test (I/T), Painting, MS&C and Grinding, Grinding and I/T, I/T and Painting, Null Set} and L=9.

This new expression for the optimal shortfall for a given demand realization can now be used in place of **P2** in the recourse function for the scenario based version of the stochastic program **SP1**. This stochastic program, denoted by **SP1(P4),** is given by,

$$\text{Min}\,\{\sum_{k=1}^{K} p_k c_k\}$$

subject to

$$1.\ (1/\text{S})\sum_{s=1}^{S}\sum_{f\in F} sf^{\,s} \le T$$

$$2.\ sf^{\,s} \ge \sum_{f\in P(\Lambda_l)} d_f - \sum_{k\in \Lambda_l} c_k \qquad l = 1,\ldots,L \quad s = 1,\ldots,S$$

$$3.\ \mathbf{c}, \mathbf{sf} \ge \mathbf{0}$$

The number of constraints is *LS*+1. Note that the null set can be removed from the possible set of stage subsets as the right hand side of constraint 2 is zero for the null set. The number of decision variables used in **SP1(P4)** is equal to *K*+*S* as opposed to *K*+2*S*|*F*|, for **SP1(P2)**. For the Alcalde job shop, *K*=5 and |*F*|=6. Therefore the number of variables is 5+S as opposed to 5+12*S*. Excel Solver can only accept 200 decision variables. Therefore, the maximum number of scenarios that could be used with Excel Solver is now 195 instead of 16. This is a very significant increase in the number of scenarios allowed and as such should enable a much better modeling of the demand distributions.

Two very important points should be noted. The first one is that this formulation assumes that capacity usages are equal for all flow paths at a stage. This may be a reasonable approximation for some supply chains but may be poor for others. Secondly, while the number of variables decreases, the number of constraints can increase significantly.

The number of constraints is *LS*+1 and therefore increases as the number of possible stage subsets to consider, *L*, increases. In the Alcalde job shop, *L*=8 when the null set is removed and a problem of this size is easily handled. However, if *L* is large than this formulation may be very hard to work with. *L* will be large if the number of stages is large and the possible subsets cannot be pruned significantly using Lemma 2. The Work Center A supply chain contains 12 stages with numerous stages having no overlap in the flow paths processed. *L* is very large for this supply chain and the stochastic formulation, **SP1(P4)** developed is thus not very useful.

4.4.3.1 Results for the Alcalde Job

The stochastic program formulation, **SP1(P4)** was used for the five problem examples detailed for the Alcalde job shop in Section 4.4.2. Again, each example was solved for a range of expected shortfall bounds, from 100 to 600. Frontline's Excel Premium Solver was used and in each case 600 demand scenarios were used. Flow path demand scenarios were randomly generated from a normal distribution given the mean and standard deviations specified in Table 10 in Section 4.4.1. This problem comprised 605 variables and 4801 constraints. Solution time was on the order of 20 minutes.

The solution time can be reduced significantly if the number of demand scenarios is reduced. This stochastic program approximates the actual expected shortfall by essentially sampling the demand distribution. As the number of scenarios used decreases, the demand distribution approximation deteriorates. The confidence interval about the expected shortfall will therefore be larger. The shortfall for each demand scenario is given in the optimal solution to **SP1(P4)** and thus the confidence interval can be determined following Law and Kelton (1991, Chapter 4.5). This was done for Example 5 for an expected shortfall upper bound of 400. To investigate the sensitivity of the confidence interval to the number of demand scenarios used, Example 5 was solved for a number of cases, with each case having a different number of demand scenarios. The confidence interval was determined for each case and the results are presented in Table 14.

Table 14

Number of Scenarios	**95% Confidence Interval**
16	400+/-239.2
100	400+/-108.1
195	400+/-79.0
300	400+/-68.1
600	400+/-45.5

As can be seen from Table 14, the 95% confidence interval for the expected shortfall is [160.8,639.2] when only sixteen demand scenarios are used. This is very wide considering the expected shortfall upper bound is 400. As noted in Section 4.4.3, sixteen is the maximum number of demand scenarios allowed by the standard Excel Solver when solving the stochastic program **SP1(P2),** i.e. the program that uses the original production planning formulation. This large confidence interval suggests that using the **SP1(P2**) program would not be advisable.

For 195 scenarios, the confidence interval is a lot narrower [291.9,508.1], but is still significant. 195 corresponds to the maximum number of scenarios allowed by the standard Excel Solver for the stochastic program **SP1(P4)**. This formulation thus gives a tighter confidence interval because more demand scenarios are allowed. However, by increasing the number of demand scenarios, the confidence interval can be significantly tightened. At 600 scenarios, the interval is [354.5,445.5]. It is therefore recommended that the number of demand scenarios used be as large as possible given the optimization package and solution time available.

In Section 4.3, when a service level criterion is used to determine the capacity investments, it was shown that positive correlation leads to a reduction in the protection level required as compared to the case in which correlation is ignored. Again, it is reasonable to conjecture that ignoring correlation will lead to an overestimation of the required capacity when an expected shortfall criterion is used.

The scenario based stochastic program, **SP1(P4)**, does take into account the correlation between stage demands. To compare the capacity investments when correlation is considered to the case in which it is ignored, Table 15 presents the decrease in $\Phi[Z_k*]$ for both the independent formulation, **SP1,Ind**, and **SP1(P4)**. A positive value implies that $\Phi[Z_k*]$ decreases going from **SP1,Ind** to **SP1(P4)**. The percentage reduction in capacity cost obtained by using **SP1(P4)** instead of **SP1,Ind** is also presented. Tables 16 and 17 present the actual $\Phi[Z_k*]$ results.

In Section 4.4.2, the optimal capacity investment for **SP1,Ind** was shown to be such that $\Phi[Z_k*]$ was a decreasing linear function of the capacity cost, where Z_k* is the protection level for stage k, defined earlier. Comparing the results **SP1(P4)** and **SP1,Ind,** in nearly all cases the $\Phi[Z_k*]$ decreases for the MS&C, Grinding, Inspect/test and Painting stages. In other words, the protection level is decreased when correlation is considered. The $\Phi[Z_k*]$ for the Surface Preparation are slightly increased.

The Surface Preparation stage processes all flow paths, while the other stages only process a strict subset of the flow paths. The other stage demands are positively correlated with the Surface Preparation stage demand. If the Surface Preparation stage's capacity is not exceeded, the probability that another stage's capacity is exceeded is lower than this probability if the stage demands are independent. The solution appears to account for this by lowering the protection levels of the other stages.

As can be seen from Table 15, the reduction in capacity costs increases as the shortfall bound increases from 100 to 600.

Table 15

Example	Expected Shortfall Upper Bound	% Decrease in Total Cost	Decrease in Φ (z*) from SP1,Ind to SP1(P4)				
			MS&C	Grinding	Surf. Prep.	Ins/Test	Painting
1	**100**	0.1%	0.01	-0.03	0.01	0.03	0.00
	200	0.9%	0.02	0.02	0.01	0.06	-0.02
	300	1.7%	0.06	0.06	0.00	0.09	0.01
	400	2.5%	0.11	0.08	0.00	0.09	0.06
	500	3.2%	0.14	0.12	0.00	0.14	0.11
	600	4.0%	0.15	0.17	0.01	0.13	0.18
2	**100**	0.4%	0.02	-0.02	0.00	0.07	-0.02
	200	1.0%	0.04	0.01	0.00	0.06	0.01
	300	1.5%	0.09	0.00	-0.01	0.15	0.00
	400	2.2%	0.12	0.10	-0.02	0.12	0.05
	500	2.9%	0.17	0.10	-0.02	0.12	0.12
	600	3.6%	0.18	0.18	0.02	0.06	0.07
3	**100**	-0.1%	0.02	-0.03	0.00	-0.01	0.03
	200	0.8%	0.02	0.03	0.00	0.06	0.01
	300	1.5%	0.10	0.06	-0.01	0.08	-0.01
	400	2.1%	0.12	0.09	-0.01	0.11	0.04
	500	2.8%	0.17	0.10	0.00	0.11	0.07
	600	3.5%	0.18	0.13	0.01	0.17	0.12
4	**100**	0.0%	0.02	-0.02	-0.01	0.05	0.00
	200	0.8%	0.04	-0.01	-0.01	0.07	0.03
	300	1.2%	0.10	0.10	-0.03	0.06	-0.01
	400	1.9%	0.13	0.09	-0.02	0.11	0.04
	500	2.6%	0.15	0.14	-0.01	0.08	0.12
	600	3.3%	0.17	0.18	-0.01	0.18	0.15
5	**100**	-0.4%	0.05	-0.02	-0.03	0.06	0.03
	200	0.1%	0.06	0.03	-0.03	0.10	0.02
	300	0.1%	0.10	0.10	-0.05	0.09	0.00
	400	0.8%	0.16	0.12	-0.04	0.07	-0.02
	500	1.2%	0.19	0.13	-0.04	0.14	0.02
	600	1.7%	0.25	0.18	-0.04	0.18	0.03

Table 16

Example	Expected Shortfall Upper Bound	SP1,Ind Φ (z*)				
		MS&C	Grinding	Surf. Prep.	Ins/Test	Painting
1	**100**	0.92	0.92	0.92	0.92	0.92
	200	0.86	0.86	0.86	0.86	0.86
	300	0.81	0.81	0.81	0.81	0.81
	400	0.77	0.76	0.76	0.76	0.77
	500	0.72	0.72	0.72	0.72	0.72
	600	0.68	0.68	0.68	0.68	0.68
2	**100**	0.93	0.96	0.92	0.89	0.90
	200	0.88	0.92	0.85	0.81	0.81
	300	0.83	0.88	0.79	0.74	0.74
	400	0.79	0.85	0.74	0.68	0.68
	500	0.76	0.83	0.70	0.62	0.62
	600	0.72	0.81	0.66	0.56	0.57
3	**100**	0.94	0.91	0.92	0.94	0.96
	200	0.88	0.85	0.85	0.89	0.92
	300	0.84	0.79	0.79	0.84	0.89
	400	0.80	0.74	0.74	0.80	0.86
	500	0.76	0.69	0.69	0.77	0.84
	600	0.73	0.65	0.65	0.73	0.82
4	**100**	0.94	0.94	0.91	0.95	0.92
	200	0.89	0.89	0.84	0.90	0.85
	300	0.84	0.85	0.78	0.86	0.79
	400	0.80	0.81	0.72	0.83	0.74
	500	0.77	0.78	0.67	0.80	0.70
	600	0.74	0.75	0.62	0.77	0.65
5	**100**	0.96	0.96	0.88	0.96	0.96
	200	0.93	0.93	0.79	0.93	0.93
	300	0.90	0.90	0.72	0.90	0.90
	400	0.88	0.88	0.65	0.88	0.88
	500	0.86	0.86	0.59	0.86	0.86
	600	0.84	0.84	0.53	0.84	0.84

Table 17

Example	Expected Shortfall Upper Bound	SP1(P4) Φ (z*)				
		MS&C	Grinding	Surf. Prep.	Ins/Test	Painting
1	**100**	0.92	0.96	0.92	0.90	0.93
	200	0.84	0.85	0.86	0.81	0.89
	300	0.75	0.75	0.81	0.72	0.80
	400	0.66	0.69	0.76	0.67	0.71
	500	0.59	0.60	0.72	0.59	0.62
	600	0.53	0.52	0.67	0.55	0.50
2	**100**	0.91	0.98	0.91	0.82	0.91
	200	0.84	0.91	0.85	0.76	0.80
	300	0.74	0.88	0.81	0.59	0.74
	400	0.67	0.75	0.76	0.55	0.63
	500	0.59	0.73	0.72	0.50	0.50
	600	0.55	0.62	0.64	0.50	0.50
3	**100**	0.92	0.94	0.92	0.94	0.93
	200	0.86	0.81	0.85	0.83	0.92
	300	0.74	0.73	0.80	0.76	0.90
	400	0.67	0.65	0.75	0.70	0.82
	500	0.59	0.59	0.70	0.66	0.77
	600	0.55	0.52	0.64	0.57	0.70
4	**100**	0.92	0.96	0.92	0.90	0.92
	200	0.84	0.90	0.84	0.83	0.82
	300	0.74	0.75	0.81	0.81	0.80
	400	0.68	0.72	0.74	0.72	0.70
	500	0.62	0.64	0.68	0.72	0.57
	600	0.57	0.57	0.63	0.59	0.50
5	**100**	0.92	0.98	0.91	0.90	0.93
	200	0.86	0.90	0.83	0.83	0.91
	300	0.80	0.80	0.77	0.81	0.90
	400	0.72	0.75	0.69	0.81	0.90
	500	0.67	0.73	0.63	0.72	0.83
	600	0.59	0.66	0.58	0.66	0.81

4.5 Conclusion

Most existing literature on capacity planning under uncertainty focuses on either single-product or single-stage supply chains. In multiple-product supply chains, floating bottlenecks can occur and thus it is necessary to consider the multiple stages of the supply chain. Eberly and Van Mieghem (1997) and Harrison and van Mieghem (1999) study this problem when an expected profit criterion is used to determine the optimal capacity investment.

Chapter 4 studies the supply chain capacity problem using both a service level and an expected shortfall criterion. It develops some insight into both problems and provides Excel based solution techniques.

For the service level criterion, an equal protection level policy is easily improved upon. If the stage demands are independent, the optimal protection level is approximately a decreasing function of the log of the product of stage demand standard deviation and capacity cost. Ignoring correlation overestimates the capacity required. Stages that are highly positively correlated can have significantly lower protection levels. A non-linear formulation is developed using a Bonferroni inequality and approximations to single and bivariate normal probabilities.

For the expected shortfall criterion, it is shown that the optimal capacities should be such that, for a small increase in a stage capacity, the decrease in shortfall divided by the capacity cost should be the same for all stages. This follows directly from the work of Harrison and Van Mieghem (1999). If stage demands are independent, then the optimal protection levels are such that the probability of a stage demand exceeding its capacity is a linear function of the capacity cost. A scenario based stochastic program that enable a very large number of scenarios to be considered is developed for certain supply chain types.

As in Eberly and Van Mieghem (1997) and Harrison and van Mieghem (1999), Chapter 4 assumes that stages are totally flexible when capacity decisions are being made. This is quite a strong assumption. Chapter 3 studies the flexibility decision assuming capacities are fixed. Developing a unified strategy that considers both partial flexibility and capacity simultaneously is an area that has yet to be considered for multiple-stage multiple-product supply chains.

5 References and Appendices

5.1 Appendix 1

This appendix contains proofs of the lemmas from Chapter 2.

Lemma 1:

The optimal capacity choice, y_I*, for a central decision-maker is given by,

$$y_I^* = F_X^{-1}\left(\frac{m_I - c_I}{m_I - v_I}\right) = F_X^{-1}\left(\frac{r - p_M - p_S - c_M - c_S}{r - p_M - p_S - v_M - v_S}\right)$$

Proof:

The total channel profit as a function of y_I, $\Pi(y_I)$ is given by,

$$\Pi(y_I) = -c_I y_I + m_I \int_a^{y_I} x f_X(x)dx + m_I y_I [1 - F_X(y_I)] + v_I \int_a^{y_I} [y_I - x] f_X(x)dx$$

Taking the derivative with respect to y_I, $\frac{d\Pi(y_I)}{dy_I} = -c_I + m_I - (m_I - v_I)F_X(y_I)$. The second derivative with respect to y_I, is $\frac{d^2\Pi(y_I)}{dy_I^{\ 2}} = -(m_I - v_I) f_X(y_I)$

The retail price is restricted to $r>p_M+p_S+c_M+c_S$. The salvage values are also strictly less than the capacity costs. Therefore m_I-v_I>0. The density function is non-negative, therefore,

$$\frac{d^2\Pi(y_I)}{dy_I^{\ 2}} \le 0$$

and $\Pi(y_I)$ is a concave function of y_I. Therefore, the first order condition is sufficient for y_I* to be the profit maximizing capacity. So,

$$\left.\frac{d\Pi(y_I)}{dy_I}\right|_{y_I = y_I^*} = 0 \ \Rightarrow y_I^* = F_X^{-1}\left(\frac{m_I - c_I}{m_I - v_I}\right)$$

Lemma 2:

For a given wholesale price, w, the optimal capacity choice, $y_M(w)$, for the manufacturer, assuming the supplier has infinite capacity, is given by,

$$y_M(w) = F_X^{-1}\left(\frac{m_M - c_M}{m_M - v_M}\right) = F_X^{-1}\left(\frac{r - w - p_M - c_M}{r - w - p_M - v_M}\right)$$

Proof:

As for Lemma 1 above, but replacing capacity, salvage and unit margin parameters as appropriate.

Lemma 3:

For a given wholesale price, w, the optimal capacity choice, $y_S(w)$, for the supplier, assuming the manufacturer has infinite capacity, is given by,

$$y_S(w) = F_X^{-1}\left(\frac{m_S - c_S}{m_S - v_S}\right) = F_X^{-1}\left(\frac{w - p_S - c_S}{w - p_S - v_S}\right)$$

Proof:

As for Lemma 1 above, but replacing capacity, salvage and unit margin parameters as appropriate.

Lemma 4:

(i) $y_M(w)$ is strictly decreasing in w

(ii) $y_S(w)$ is strictly increasing in w

(Note that $F_X(x)$ is assumed to be continuous and differentiable)

Proof:

(i)

$$\frac{dy_M(w)}{dw} = \left(\frac{1}{f_X(y_M(w))}\right)\left(\frac{v_M - c_M}{(r - w - p_M - v_M)^2}\right) < 0$$

(ii)

$$\frac{dy_S(w)}{dw} = \left(\frac{1}{f_X(y_S(w))}\right)\left(\frac{c_S - v_S}{(w - p_S - v_S)^2}\right) > 0$$

Lemma 5:

(i) There is a unique wholesale price, w, such that $y_M(w)=y_S(w)$

(ii) This unique wholesale price, w_{crit}, is given by,

$$w_{crit}=\frac{(r-p_M)(c_S-v_S)+p_S(c_M-v_M)+c_Mv_S-c_Sv_M}{c_S-v_S+c_M-v_M}$$

(iii) At this wholesale price, w_{crit}, $y_M(w_{crit})=y_S(w_{crit})=y_I*$

(iv) If $w=w_{crit}$, then the total channel profits equals the channel profits obtained by a central decision-maker.

Proof:

(i) and (ii)From Lemma 4, $y_M(w)$ is strictly decreasing in w and $y_S(w)$ is strictly increasing in w. Therefore $y_M(w)$ and $y_S(w)$ can cross each other at most once.

$$y_M(w)=F_X^{-1}\left(\frac{m_M-c_M}{m_M-v_M}\right)\qquad y_S(w)=F_X^{-1}\left(\frac{m_S-c_S}{m_S-v_S}\right)$$

So, $y_M(w)=y_S(w)$ iff,

$$\left(\frac{m_M-c_M}{m_M-v_M}\right)=\left(\frac{m_S-c_S}{m_S-v_S}\right)$$

$m_M=r\text{-}w\text{-}p_M$ and $m_S=w\text{-}p_S$. So, $y_M(w)=y_S(w)$ only if,

$$w=\frac{(r-p_M)(c_S-v_S)+p_S(c_M-v_M)+c_Mv_S-c_Sv_M}{(c_S-v_S+c_M-v_M)}$$

(iii) $y_I*=y_S(w)$ iff,

$$\left(\frac{m_I-c_I}{m_I-v_I}\right)=\left(\frac{m_S-c_S}{m_S-v_S}\right)$$

$m_I=r\text{-}p_M\text{-}p_S$, $c_I=c_M+c_S$, $v_I=v_M+v_S$ and So, $y_I*=y_S(w)$ iff,

$$w=\frac{(r-p_M)(c_S-v_S)+p_S(c_M-v_M)+c_Mv_S-c_Sv_M}{(c_S-v_S+c_M-v_M)}$$

which is the same condition for $y_M(w)=y_S(w)$ so at w_{crit}, $y_M(w_{crit})=y_S(w_{crit})=y_I*$

(iv) The total channel profits depend only on the supplier and manufacturer capacity. Since, $y_M(w_{crit})=y_S(w_{crit})=y_I*$, the capacities are the same as those chosen by a central decision-maker and therefore the total channel profits are the same.

Lemma 6:

The necessary and sufficient condition on the demand distribution, $F_X(x)$, for $y_S(w)$ to be a concave function of w is given by,

$$\frac{f'_X(x)[1-F_X(x)]}{[f_X(x)]^2} \geq -2 \quad \forall\, x \in [a,b]$$

For strict concavity, the inequality needs to be strict. This condition is referred to as the (strict) concavity condition in the rest of the proofs. Note that this condition is equivalent to requiring that $g'(x) \geq 0\ \forall$ x, where,

$$g(x) = \frac{f_X(x)}{[1-F_X(x)]^2}$$

Proof:

$$y_S(w) = F_X^{-1}\left(\frac{m_S - c_S}{m_S - v_S}\right) = F_X^{-1}\left(\frac{w - p_S - c_S}{w - p_S - v_S}\right)$$

Therefore,

$$\frac{d^2 y_S(w)}{dw^2} = \left(\frac{1}{f_X(y_S(w))}\right)\left(\frac{(1 - F_X[y_S(w)])^3}{(c_S - v_S)^2}\right)\left(-2 - \left(\frac{f'_X(y_S(w))(1 - F_X[y_S(w)])}{[f_X(y_S(w))]^2}\right)\right)$$

The first bracketed term is > 0. Now, because $v_S < c_S$, i.e. the salvage value is strictly less than capacity cost,

$$F_X[y_S(w)] = \left(\frac{w - p_S - c_S}{w - p_S - v_S}\right) < 1$$

Therefore, the second bracketed term is > 0. So, for

$$\frac{d^2 y_S(w)}{dw^2} \leq 0$$

the third bracketed term needs to be ≤ 0, and for strict concavity this term needs to be < 0. Indeed this condition is sufficient. So, the necessary and sufficient condition for $y_S(w)$ to be a concave function of w is then given by,

$$\frac{f'_X(y_S(w))[1-F_X(y_S(w))]}{[f_X(y_S(w))]^2} \geq -2$$

Since $y_S(w)$ is strictly increasing in w and since for every $x \in [a,b]$ there is wholesale price w such that $y_S(w)=x$, then this condition is satisfied iff

$$\frac{f_X'(x)[1-F_X(x)]}{\left[f_X(x)\right]^2} \geq -2 \quad \forall\, x \in [a,b]$$

Lemma 7:

(i) If $F_X(x)$ is an increasing failure rate distribution (IFR), then,

$$\frac{f_X'(x)[1-F_X(x)]}{\left[f_X(x)\right]^2} > 1 \quad \forall\, x$$

and (ii) the strict concavity condition is satisfied so $y_S(w)$ is a strictly concave function.

Proof:

(i)The failure rate function, $h(x)$, for a distribution $F_X(x)$, is defined as,

$$h(x) = \frac{f_X(x)}{1-F_X(x)}$$

A distribution is said to have an increasing failure rate function (IFR) if $h'(x) > 0 \quad \forall\, x$

$$h'(x) = \frac{[1-F_X(x)]\, f_X'(x) - [f_X(x)]^2}{[1-F_X(x)]^2}$$

If $F_X(x)$ is an IFR distribution, then

$$h'(x) > 0 \quad \forall\, x \Rightarrow \frac{[1-F_X(x)]\, f_X'(x) - [f_X(x)]^2}{[1-F_X(x)]^2} > 0 \quad \forall\, x$$

$$\Rightarrow \frac{[1-F_X(x)]\, f_X'(x)}{[f_X(x)]^2} > 1 \quad \forall\, x$$

(ii) Follows directly from part (i) and Lemma 6 above

Lemma 8:

In *Game A*, where the wholesale price is exogenous and the manufacturer is the Stackleberg leader, the manufacturer and supplier choose their capacities such that $y_M^* = y_S^* = \min\{y_S(w), y_M(w)\}$, where $y_M(w)$ and $y_S(w)$ are given by Lemma 2 and Lemma 3 respectively.

Proof:

The supplier never chooses a capacity larger than the manufacturer's capacity choice, as its total sales are limited by the manufacturer's capacity. From Lemma 3, the supplier's profit,

$\Pi_S(y)$, is concave with the maximum achieved at $y_S(w)$. For $y<y_S(w)$, $\Pi_S(y)$ is increasing in y. Therefore, if the manufacturer announces a capacity of y, the supplier chooses a capacity of $min\{y,y_S(w)\}$. Likewise, the manufacturer never announces a capacity larger than the supplier would choose, as its total sales are limited by the supplier's capacity choice. From Lemma 2, the manufacturer's profit, $\Pi_M(y)$, is concave with the maximum achieved at $y_M(w)$. For $y<y_M(w)$, $\Pi_M(y)$ is increasing in y. Therefore, the manufacturer chooses its capacity, $y_M^*=\min\{y_S(w),y_M(w)\}$. The supplier chooses a capacity of $min\{y_M^*,y_S(w)\}=min\{y_S(w),y_M(w)\}$.

Lemma 9:

In *Game B*, where the manufacturer chooses the wholesale price, the manufacturer never chooses a wholesale price, w, such that $w>w_{crit}$.

Proof:

Let the manufacturer choose a wholesale price, $w^*>w_{\mathrm{crit}}$. The channel capacity is then given by $\min\{y_S(w^*),y_M(w^*)\}$. For $w>w_{crit}$, $y_S(w)>y_M(w)$, so the channel capacity is $y_M(w^*)$. However, there exists a $w^{**}<w_{crit}$ such that $y_S(w^{**})=y_M(w^*)$. [For every $x\in[a,b]$ there is a wholesale price w_S such that that $y_S(w_S)=x$. Likewise there is a unique wholesale price w_M such that $y_M(w_M)=x$. $y_S(w)$ is strictly increasing in w, $y_M(w)$ is strictly decreasing in w and there is a unique wholesale price w_{crit} for which $y_S(w_{crit})=y_M(w_{crit})$.]

The manufacturer's expected total sales depend only on the channel capacity. Thus the manufacturer's expected sales revenue is the same for w^* and w^{**}, as is the manufacturer's capacity cost and expected salvage revenue. The cost per unit sold is strictly less if w^{**} is chosen. Therefore the manufacturer's profit is strictly greater at w^{**} and so the manufacturer never chooses a $w>w_{crit}$.

Therefore, the manufacturer's wholesale price choice can be restricted to $w\leq w_{crit}$. For the supplier to invest in capacity, the wholesale price, w, must be larger than the sum of unit capacity cost and unit marginal cost ($w\geq p_S+c_S$). So the manufacturer's optimal wholesale price falls within $p_S+c_S<w\leq w_{crit.}$

Lemma 10:

If $F_X(x)$ satisfies the concavity condition, then the manufacturer's profit, $\Pi_M(w)$, restricted to $p_S+c_S<w\leq w_{crit}$, is a strictly concave function.

Proof:

In this range the manufacturer's profit, $\Pi_M(w)$, is given by,

$$\Pi_M(w) = -c_M y_S(w) + (r - p_M - w)\int_a^{y_S(w)} x f_X(x)dx + (r - p_M - w) y_S(w)[1 - F_X(y_S(w))]$$
$$+ v_M \int_a^{y_S(w)} [y_S(w) - x] f_X(x)dx$$

Taking the derivative with respect to *w*,

$$\frac{d\Pi_M(w)}{dw} = \left(\frac{dy_S(w)}{dw}\right)\left((r - p_M - w - c_M) - (r - p_M - w - v_M)F_X(y_S(w))\right)$$
$$- \int_a^{y_S(w)} x f_X(x)dx - y_S(w)[1 - F_X(y_S(w))]$$

Taking the second derivative,

$$\frac{d^2\Pi_M(w)}{dw^2} = \left(\frac{d^2 y_S(w)}{dw^2}\right)\left((r - p_M - w - c_M) - (r - p_M - w - v_M)F_X(y_S(w))]\right)$$
$$-2\left(\frac{dy_S(w)}{dw}\right)[1\text{-}F_X(y_S(w)] - (r - p_M - w - v_M)\ f_X(y_S(w))\left(\frac{dy_S(w)}{dw}\right)^2$$

Now,

(i) From Lemma 4(ii), $\frac{dy_S(w)}{dw} > 0$

(ii) As $F_X(x)$ satisfies the concavity condition, then $\frac{d^2 y_S(w)}{dw^2} \leq 0$

(iii) $r>w+p_M+c_M$ in the allowable range for *w*,

(iv) $c_M>v_M$

(v) In the specified range for *w*,

$$F_X[y_S(w)] = \left(\frac{w - p_S - c_S}{w - p_S - v_S}\right) \leq F_X[y_M(w)] = \left(\frac{r - w - p_M - c_M}{r - w - p_M - v_M}\right) < 1$$

Using (i)-(v), $\frac{d^2\Pi_M(w)}{dw^2} < 0$

Lemma 11:

If $F_X(x)$ satisfies the concavity condition (eqn. (2)) given in Lemma 6, then the optimal w for Game B is given by the first order condition,

$$\begin{aligned}\frac{d\Pi_M(w)}{dw} &= \left(\frac{dy_S(w)}{dw}\right)\big((r-p_M-w-c_M)-(r-p_M-w-v_M)F_X(y_S(w))\big)\\ &\quad -\int_a^{y_S(w)} xf_X(x)dx - y_S(w)[1-F_X(y_S(w))]\\ &= 0\end{aligned}$$

and the optimal w is strictly greater than p_S+c_S and strictly less than w_{crit}.

Proof:

From Lemma 9, the manufacturer's optimal wholesale price lies within $p_S+c_S<w^*\leq w_{crit}$. From Lemma 10, $\Pi_M(w)$ is a strictly concave function so the first order condition is sufficient for optimality as long as the wholesale price, w^*, that satisfies the condition, lies in the interior of the range, $p_S+c_S<w\leq w_{\text{crit}}$. I will show that the optimal wholesale price satisfies $p_S+c_S<w^*<w_{crit}$.

If $w\leq p_S+c_S$, then the supplier does not invest in any capacity and the manufacturer's profit is be zero. At w_{crit} the manufacturer's profit is strictly positive. $\Pi_M(w)$ is a strictly concave function in the range $p_S+c_S\leq w\leq w_{crit}$, so we must have

$$\left.\frac{d\Pi_M(w)}{dw}\right|_{w=p_S+c_S} > 0$$

At $w=w_{crit}$,

…. (L11.1)

$$\left.\frac{d\Pi_M(w)}{dw}\right|_{w=w_{crit}} = -\int_a^{y_S(w_{crit})} xf_X(x)dx - y_S(w_{crit})[1-F_X(y_S(w_{crit}))] < 0$$

…. (L11.2)

$\Pi_M(w)$ is concave function. From (L11.2) the profit is decreasing at $w=w_{crit}$, so the first order condition must be satisfied for $w<w_{crit}$. From (L11.1) the profit is increasing at $w=p_S+c_S$. Therefore, the w^* that satisfies the first order condition must satisfy $p_S+c_S<w^*<w_{crit}$.

Lemma 12:

If $F_X(x)$ satisfies the concavity condition, then in *Game B*, the total channel profits, Π_D, is strictly less than the total channel profits obtained by a central decision-maker, Π_C.

Proof:

$\Pi_d=\Pi_c$ only if the channel capacity chosen is the same as that chosen by a central decision-maker. This only happens if the manufacturer chooses the wholesale price, w, such that $w=w_{crit}$. From Lemma 11, this never occurs. Therefore, the total channel profits is strictly less than those obtained by a central decision-maker.

Lemma 13:

For any allowable price schedule (w,Δ), $p_S+c_S \le w+\Delta \le r-c_M-p_M$ and $\Delta \ge 0$, then (i) the optimal capacity choice, $y_S(w,\Delta)$, for the supplier, assuming the manufacturer has infinite capacity, is given by,

$$y_S(w,\Delta) = F_X^{-1}\left(\frac{m_S+\Delta-c_S}{m_S+\Delta-v_S}\right) = F_X^{-1}\left(\frac{w+\Delta-p_S-c_S}{w+\Delta-p_S-v_S}\right) \ge y_S(w)$$

and (ii) the optimal capacity choice, $y_M(w,\Delta)$, for the manufacturer, assuming the supplier has infinite capacity, is given by,

$$y_M(w,\Delta) = F_X^{-1}\left(\frac{m_M-\Delta-c_M}{m_M-\Delta-v_M}\right) = F_X^{-1}\left(\frac{r-w-\Delta-p_M-c_M}{r-w-\Delta-p_M-v_M}\right) \le y_M(w)$$

Proof:

(i) Let $\Pi_S(y,w,\Delta)$ denote the supplier's profit as a function of y for a price schedule of (w,Δ).If $\Delta=0$, then the supplier invests in capacity, $y_S(w)$. However, $\Delta>0$, so when determining the optimal capacity, $y_S(w,\Delta)$, we only need to look at $y \ge y_S(w)$. In this range,

$$\begin{aligned}\Pi_S(y,w,\Delta) = -c_S y + m_S \int_a^{y_S(w)} x f_X(x)dx + m_S y_S(w)[1-F_X(y_S(w)] \\ + (m_S+\Delta)\int_{y_S(w)}^{y}(x-y_S(w)) f_X(x)dx + (m_S+\Delta)(y-y_S(w))][1-F_X(y)] \\ + v_S \int_a^{y}[y-x] f_X(x)dx\end{aligned}$$

Taking the first derivative with respect to y, $\frac{d\Pi_S(y,w,\Delta)}{dy} = m_S + \Delta - c_S - (m_S + \Delta - v_S)F_X(y)$

Taking the second derivative,

$$\frac{d^2\Pi_S(y,w,\Delta)}{dy^2} = -(m_S + \Delta - v_S) f_X(y) \le 0$$

So, $\Pi_S(y,w,\Delta)$ is a concave function of y and the first order condition is sufficient for optimality. Thus,

$$F_X(y_S(w,\Delta)) = \left(\frac{m_S + \Delta - c_S}{m_S + \Delta - v_S}\right) = \left(\frac{w + \Delta - p_S - c_S}{w + \Delta - p_S - v_S}\right)$$

(ii) Proof follows similarly to (i) but the quantity premium, Δ, is subtracted from the unit margin.

Important Note for Proofs of Lemmas 14, 15 and 16:

(i) $y_S(w,\Delta)=y_S(w')$ and $y_M(w,\Delta)=y_M(w')$ where $w'=w+\Delta$.

(ii) The first and second derivatives of $y_S(w,\Delta)$ with respect to either w or Δ is the same as the derivatives of $y_S(w')$ with respect to w'.

(iii) The first and second derivatives of $y_M(w,\Delta)$ with respect to either w or Δ is the same as the derivatives of $y_M(w')$ with respect to <u>w</u>'.

Given (i),(ii) and (iii), Lemmas 14, 15 and 16 follow directly from the proofs of Lemmas 4, 5 and 6. Fully worked proofs, independent of Lemmas 4, 5 and 6 are available.

Lemma 14:

(i) $y_M(w,\Delta)$ is strictly decreasing in both w and Δ.

(ii) $y_S(w,\Delta)$ is strictly increasing in both w and Δ.

(Note that $F_X(x)$ is assumed to be continuous and differentiable)

Proof:

Follows from Lemma 4 and above note.

Lemma 15:

For a given wholesale price $w \le w_{crit}$ there is a unique quantity premium Δ_{crit}, given by

$$w+\Delta_{crit} = \frac{(r-p_M)(c_S - v_S) + p_S(c_M - v_M) + c_M v_S - c_S v_M}{c_S - v_S + c_M - v_M} = w_{crit},$$

such that $y_M(w,\Delta_{crit}) = y_S(w,\Delta_{crit}) = y_I^*$.

Proof:

Follows from Lemma 5 and above note.

Lemma 16:

(i) $y_S(w,\Delta)$ is a concave function of both w and Δ iff the concavity condition given by equation (2) holds for $F_X(x)$.

(ii) If $F_X(x)$ is an IFR distribution then equation (2) is satisfied so $y_S(w,\Delta)$ is a strictly concave function of both Δ and w.

Proof:

(i) From Lemma 6 $\frac{\partial^2 y_S(w,\Delta)}{\partial w^2} \le 0$ if and only if the concavity condition holds. From the above note, $\frac{\partial^2 y_S(w,\Delta)}{\partial w \partial \Delta} = \frac{\partial^2 y_S(w,\Delta)}{\partial \Delta^2} = \frac{\partial^2 y_S(w,\Delta)}{\partial w^2}$. Therefore, $\frac{\partial^2 y_S(w,\Delta)}{\partial w^2} \le 0$, $\frac{\partial^2 y_S(w,\Delta)}{\partial \Delta^2} \le 0$ and $\left(\frac{\partial^2 y_S(w,\Delta)}{\partial w^2}\right)\left(\frac{\partial^2 y_S(w,\Delta)}{\partial \Delta^2}\right) - \left(\frac{\partial^2 y_S(w,\Delta)}{\partial w \partial \Delta}\right)^2 \le 0$. So $y_S(w,\Delta)$ is a concave function of w and Δ.

(ii) Follows directly from Lemma 7.

Lemma 17:

For Game C the manufacturer never chooses a quantity premium, Δ, such that $\Delta > \Delta_{crit}$.

Proof:

Similar to proof of Lemma 9 but adapting for Δ instead of w. Therefore, the manufacturer's wholesale price choice can be restricted to $0 \le \Delta \le \Delta_{crit}$.

Lemma 18:

For Game C, if $F_X(x)$ satisfies the concavity condition given by equation (2), then the manufacturer's profit, $\Pi_M(w,\Delta)$, restricted to $0\leq\Delta\leq\Delta_{crit}$, is a strictly concave function of Δ.

Proof:

The manufacturer's profit, $\Pi_M(w,\Delta)$, is given by,

$$\begin{aligned}\Pi_M(w,\Delta) = -c_M y_S(w,\Delta) + m_M \int_a^{y_S(w)} x f_X(x)dx + m_M y_S(w)[1 - F_X(y_S(w)]\\ + (m_M - \Delta)\int_{y_S(w)}^{y_S(w,\Delta)} (x - y_S(w)) f_X(x)dx\\ + (m_M - \Delta)(y_S(w,\Delta) - y_S(w))][1 - F_X(y_S(w,\Delta))]\\ + v_M \int_a^{y_S(w,\Delta)} [y_S(w,\Delta) - x] f_X(x)dx\end{aligned}$$

Taking the derivative with respect to Δ,

$$\begin{aligned}\frac{d\Pi_M(w,\Delta)}{d\Delta} = \left(\frac{dy_S(w,\Delta)}{d\Delta}\right)\left[(m_M - \Delta - c_M) - (m_M - \Delta - v_M)F_X(y_S(w,\Delta))\right]\\ - \int_{y_S(w)}^{y_S(w,\Delta)} (x - y_S(w)) f_X(x)dx - (y_S(w,\Delta) - y_S(w))[1 - F_X(y_S(w,\Delta))]\end{aligned}$$

Taking the second derivative,

$$\begin{aligned}\frac{d^2\Pi_M(w,\Delta)}{d\Delta^2} = \left(\frac{d^2 y_S(w,\Delta)}{d\Delta^2}\right)\left[(m_M - \Delta - c_M) - (m_M - \Delta - v_M)F_X(y_S(w,\Delta))\right]\\ -2\left(\frac{dy_S(w,\Delta)}{d\Delta}\right)[1 - F_X(y_S(w,\Delta)] - (m_M - \Delta - v_M) f_X(y_S(w,\Delta))\left(\frac{dy_S(w,\Delta)}{d\Delta}\right)^2\end{aligned}$$

Now,

(i) From Lemma 14(ii), $\frac{dy_S(w,\Delta)}{d\Delta} > 0$

(ii) As $F_X(x)$ satisfies concavity, then from Lemma 16(i), $\frac{d^2 y_S(w,\Delta)}{d\Delta^2} \leq 0$

(iii) $m_M>\Delta+c_M$ in the allowable range for Δ,

(iv) $c_M>v_M$

(v) In the specified range for w,

$$F_X\left[y_S(w)\right]=\left(\frac{m_S+\Delta-c_S}{m_S+\Delta-v_S}\right)\le F_X\left[y_M(w)\right]=\left(\frac{m_M-\Delta-c_M}{m_M-\Delta-v_M}\right)<1$$

Using (i)-(v), $\dfrac{d^2\Pi_M(w,\Delta)}{d\Delta^2}<0$

Lemma 19:

For Game C,

(i) the manufacturer chooses a positive quantity premium, i.e. $\Delta^*(w)>0$

(ii) the supplier's profit strictly increases with increasing Δ

(iii) $\Pi_T(w,\Delta^*(w))$ the expected total supply chain profit for the price schedule $(w,\Delta^*(w))$, is strictly greater than the expected total supply chain profit $\Pi_T(w)$ when no quantity premium is offered but the supply chain is not completely coordinated, i.e. $\Pi_T(w)<\Pi_T(w_{crit})$.

Proof:

(i) The derivative of $\Pi_M(w,\Delta)$ with respect to Δ is given in the proof of Lemma 18,

$$\frac{d\Pi_M(w,\Delta)}{d\Delta}=\left(\frac{dy_S(w,\Delta)}{d\Delta}\right)\left[(m_M-\Delta-c_M)-(m_M-\Delta-v_M)F_X(y_S(w,\Delta))\right]$$
$$-\int_{y_S(w)}^{y_S(w,\Delta)}(x-y_S(w))f_X(x)dx-(y_S(w,\Delta)-y_S(w))[1-F_X(y_S(w,\Delta))]$$

So, at Δ=0,

$$\left.\frac{d\Pi_M(w,\Delta)}{d\Delta}\right|_{\Delta=0}=\left(\left.\frac{dy_S(w,\Delta)}{d\Delta}\right|_{\Delta=0}\right)\left[(m_M-c_M)-(m_M-v_M)F_X(y_S(w)\right]$$

From lemma 14(ii), the first bracketed term is > 0. The second bracketed term is also >0. Therefore,

$$\left.\frac{d\Pi_M(w,\Delta)}{d\Delta}\right|_{\Delta=0}>0$$

From Lemma 18 Π_M(w, Δ) is a strictly concave function of Δ. $\Pi_S(w,\Delta)$ is increasing at Δ=0 and so the optimal Δ^* is strictly greater than 0.

(ii) For a fixed wholesale price, w, the supplier's profit as a function of Δ, $\Pi_S(w,\Delta)$, is given by,

$$\Pi_S(w,\Delta) = -c_S y_S(w,\Delta) + m_S \int_a^{y_S(w)} x f_X(x)dx + m_S y_S(w)[1 - F_X(y_S(w)]$$

$$+ (m_S + \Delta) \int_{y_S(w)}^{y_S(w,\Delta)} (x - y_S(w)) f_X(x)dx$$

$$+ (m_S + \Delta)(y_S(w,\Delta) - y_S(w))][1 - F_X(y_S(w,\Delta))]$$

$$+ v_S \int_a^{y_S(w,\Delta)} [y_S(w,\Delta) - x] f_X(x)dx$$

Taking the derivative with respect to Δ,

$$\frac{d\Pi_S(w,\Delta)}{d\Delta} = \left(\frac{dy_S(w,\Delta)}{d\Delta}\right)\left[(m_S + \Delta - c_S) - (m_S - \Delta - v_S)F_X(y_S(w,\Delta))\right]$$

$$+ \int_{y_S(w)}^{y_S(w,\Delta)} (x - y_S(w)) f_X(x)dx + (y_S(w,\Delta) - y_S(w))[1 - F_X(y_S(w,\Delta))]$$

$$> 0$$

So for a fixed wholesale price $w<w_{crit}$, the supplier's expected profit increases if the manufacturer offers a positive quantity premium.

(iii) From (i) the manufacturer chooses a $\Delta>0$ and its profit strictly increases. From (ii) the supplier's profit strictly increases. Therefore the total channel profits are greater than in Game A, i.e. $\Pi_T(w,\Delta^*(w))>\Pi_T(w)$. From (i) $0<\Delta^*(w)$. I will now show that $\Delta^*(w)<\Delta_{crit}$.

$$\left.\frac{d\Pi_M(w,\Delta)}{d\Delta}\right|_{\Delta=\Delta_{crit}} = -\int_{y_S(w)}^{y_S(w,\Delta_{crit})} (x - y_S(w)) f_X(x)dx - (y_S(w,\Delta_{crit}) - y_S(w))[1 - F_X(y_S(w,\Delta_{crit}))]$$

$$< 0$$

$\Pi_M(w,\Delta)$, restricted to $0\le\Delta\le\Delta_{crit}$, is a concave function of Δ. Therefore, because

$$\left.\frac{d\Pi_M(w,\Delta)}{d\Delta}\right|_{\Delta=0} > 0 \qquad \left.\frac{d\Pi_M(w,\Delta)}{d\Delta}\right|_{\Delta=\Delta_{crit}} < 0$$

the optimal Δ for the manufacturer is given by the first order condition and $0<\Delta^*(w)<\Delta_{crit}$. Therefore the channel capacity is strictly less than that chosen by a central decision-maker and thus the total channel profit is strictly less then those obtained by a central decision-maker, i.e.

$\Pi_T(w,\Delta^*(w))<\Pi_C=\Pi_T(w_{crit})$, where the central decision-maker expected profit is the same as the total supply chain profit in Game A if $w=w_{crit}$ (Lemma 5).

Lemma 20:

(i) If $F_X(x)$ satisfies the concavity condition, then the manufacturer's profit $\Pi_M(w,\Delta)$, restricted to $p_S+c_S<w+\Delta<w_{crit}$ and $\Delta>0$, is a strictly concave function of w and Δ.

(ii) The first order conditions for w and Δ are necessary and sufficient for (w^*,Δ^*) to be optimal.

Proof:

(i) From Lemma 18,

$$\frac{\partial^2 \Pi_M(y_S(w,\Delta))}{\partial \Delta^2}<0$$

Next, I will show that,

$$\frac{\partial^2 \Pi_M(y_S(w,\Delta))}{\partial w^2}<0$$

$$\begin{aligned}\Pi_M(w,\Delta) = &-c_M y_S(w,\Delta)+m_M\int_a^{y_S(w)} x f_X(x)dx + m_M y_S(w)[1-F_X[y_S(w)]] \\ &+(m_M-\Delta)\int_{y_S(w)}^{y_S(w,\Delta)}(x-y_S(w))f_X(x)dx \\ &+(m_M-\Delta)[y_S(w,\Delta)-y_S(w)][1-F_X[y_S(w,\Delta)]] \\ &+v_M\int_{y_S(w)}^{y_S(w,\Delta)}(y_S(w,\Delta)-x)f_X(x)dx\end{aligned}$$

Taking the partial derivative with respect to the wholesale price w,

$$\begin{aligned}\frac{\partial \Pi_M(w,\Delta)}{\partial w} = &+\Delta\left(\frac{dy_S(w)}{dw}\right)[1-F_X[y_S(w)]]-\int_a^{y_S(w,\Delta)} x f_X(x)dx - y_S(w,\Delta)[1-F_X[y_S(w,\Delta)]] \\ &+\left(\frac{\partial y_S(w,\Delta)}{\partial w}\right)((m_M-\Delta-c_M)-(m_M-\Delta-v_M)F_X[y_S(w,\Delta)])\end{aligned}$$

Taking the second derivative with respect to w,

$$\frac{\partial^2 \Pi_M(w,\Delta)}{\partial w^2} = \Delta\left(\frac{d^2 y_S(w)}{dw^2}\right)[1 - F_X[y_S(w)]] - \Delta\left(\frac{dy_S(w)}{dw}\right)^2 f_X(y_S(w))$$
$$+\left(\frac{\partial^2 y_S(w,\Delta)}{\partial^2 w}\right)\left((m_M - \Delta - c_M) - (m_M - \Delta - v_M)F_X[y_S(w,\Delta)]\right)$$
$$-2\left(\frac{\partial y_S(w,\Delta)}{\partial w}\right)[1 - F_X[y_S(w,\Delta)]]$$
$$-(m_M - \Delta - v_M) f_X(y_S(w,\Delta))\left(\frac{\partial y_S(w,\Delta)}{\partial w}\right)^2$$

Now,

(i) $\Delta \geq 0$,

(ii) $\frac{\partial y_S(w,\Delta)}{\partial w} > 0$ from Lemma 14(ii)

(iii) As $F_X(x)$ satisfies the concavity condition, $\frac{d^2 y_S(w)}{dw^2} \leq 0$ and $\frac{\partial^2 y_S(w,\Delta)}{\partial w^2} \leq 0$

(iv) $m_M \geq \Delta + c_M$ is the allowable range for w

(v) $c_M > v_M$

(vi) In the specified range for w,

$$F_X[y_S(w)] = \left(\frac{m_S + \Delta - c_S}{m_S + \Delta - v_S}\right) \leq F_X[y_M(w)] = \left(\frac{m_M - \Delta - c_M}{m_M - \Delta - v_M}\right) < 1$$

Using (i)-(vi), $\frac{\partial^2 \Pi_M(w,\Delta)}{\partial w^2} < 0$

I will now use the following theorem (Theorem 2.13 from Avriel, Diewert, Schaible and Zang ,1988): Let f be a differentiable function on the open convex set $C \subset R^n$. It is concave if and only if for every two points $\mathbf{x}^1 \in C$, $\mathbf{x}^2 \in C$,

$$(\mathbf{x}^2 - \mathbf{x}^1)^T [\nabla f(\mathbf{x}^2) - \nabla f(\mathbf{x}^1)] \leq 0$$

It is strictly concave if and only if this inequality is strict for $\mathbf{x}^1 \neq \mathbf{x}^2$.

So, for this problem the condition is that for every (w^1, Δ^1) and (w^2, Δ^2) in the allowable region ($p_S + c_S < w + \Delta < w_{crit}$ and $\Delta > 0$),

$$(w^2 - w^1)\left[\frac{\partial \Pi_M(w,\Delta)}{\partial w}\bigg|_{w=w^2} - \frac{\partial \Pi_M(w,\Delta)}{\partial w}\bigg|_{w=w^1}\right] + (\Delta^2 - \Delta^1)\left[\frac{\partial \Pi_M(w,\Delta)}{\partial \Delta}\bigg|_{\Delta=\Delta^2} - \frac{\partial \Pi_M(w,\Delta)}{\partial \Delta}\bigg|_{\Delta=\Delta^1}\right] \le 0$$

…. (L20.1)

From above,

$$\frac{\partial^2 \Pi_M(w,\Delta)}{\partial w^2} < 0$$

Then if $w^2>w^1$,

$$\frac{\partial \Pi_M(w,\Delta)}{\partial w}\bigg|_{w=w^2} < \frac{\partial \Pi_M(w,\Delta)}{\partial w}\bigg|_{w=w^1}$$

and if $w^2<w^1$,

$$\frac{\partial \Pi_M(w,\Delta)}{\partial w}\bigg|_{w=w^2} > \frac{\partial \Pi_M(w,\Delta)}{\partial w}\bigg|_{w=w^1}$$

So, the first term in L20.1 is < 0 unless $w^2=w^1$, for which it is = 0. Similarly, the second term in L20.1 is < 0 unless $\Delta^2=\Delta^1$, for which it is = 0. Therefore, L20.1 is ≤ 0 for every (w^1,Δ^1) and (w^2,Δ^2) in the allowable region and this inequality is strict for every $(w^1,\Delta^1)\neq(w^2,\Delta^2)$. Using the above theorem, $\Pi_M(w,\Delta)$, restricted to $p_S+c_S<w+\Delta<w_{crit}$ and $\Delta>0$, is a strictly concave function of w and Δ.

(ii) From Lemma 19 (iv), for any $w<w_{crit}$, $0<\Delta^*(w)<\Delta_{crit}$, so $0<\Delta^*<\Delta_{crit}$ If $w=p_S+c_S$, then in effect there is no quantity premium and Δ is the constant wholesale price per unit. From 19(i), the manufacturer chooses a positive quantity premium. Therefore $w^*>p_S+c_S$. At w_{crit}, the manufacturer does not offer any quantity premium, i.e. $\Delta^*(w_{crit})=0$, as the manufacturer's capacity dictates the supply chain capacity for any positive quantity. We therefore only need to consider the case of $w^*=w_{crit}$ and $\Delta^*=0$. However, this is the same as a constant wholesale price schedule and from Lemma 11(ii) w* is strictly less than w_{crit}. So $0<w^*<w_{crit}$.

Lemma 21:

(i) If $F_X(x)$ satisfies the concavity condition given by equation (2), then the total expected channel profit when the manufacturer chooses both a wholesale price and a quantity premium is strictly greater than the total channel profit when the manufacturer only chooses a wholesale price but it is strictly less than the total supply chain profit when the supply chain is completely coordinated.

Proof:

Letting $\Pi_M(w)$ be the manufacturer's profit as a function of w when no quantity premium is allowed, then, using the expressions for $\partial\Pi_M(w,\Delta)/\partial w$ and $d\Pi_M(w)/dw$ given in Lemmas 20 and 10 respectively,

$$\frac{\partial \Pi_M(w,\Delta)}{\partial w} = +\Delta\left(\frac{dy_S(w)}{dw}\right)[1 - F_X[y_S(w)]] + \left(\left.\frac{d\Pi_M(w)}{dw}\right|_{w'=w+\Delta}\right)$$

Let (w^*,Δ^*) be the optimal (w,Δ) pair chosen by the manufacturer and w^{**} be the optimal w chosen when no quantity premium is allowed. From Lemma 21 and Lemma 10, the following first order conditions must be satisfied,

(w^*,Δ^*): $$\left.\frac{\partial \Pi_M(w,\Delta)}{\partial w}\right|_{w=w^*,\Delta=\Delta^*} = 0 \qquad \left.\frac{\partial \Pi_M(w,\Delta)}{\partial \Delta}\right|_{w=w^*,\Delta=\Delta^*} = 0$$

w^{**}: $$\left.\frac{d\Pi_M(w)}{dw}\right|_{w=w^{**}} = 0$$

But,

$$\left.\frac{\partial \Pi_M(w,\Delta)}{\partial w}\right|_{w=w^*,\Delta=\Delta^*} = \Delta^*\left(\left.\frac{dy_S(w)}{dw}\right|_{w=w^*}\right)[1 - F_X[y_S(w^*)]] + \left(\left.\frac{d\Pi_M(w)}{dw}\right|_{w'=w^*+\Delta^*}\right)$$

Now,

(i) From Lemma 19(i) $\Delta^*>0$.

(ii) $\frac{dy_S(w)}{dw} > 0$ from Lemma 4(ii)

(iii) $F_X(y_S(w^*)) < 1$

So,

$$\left.\frac{d\Pi_M(w)}{dw}\right|_{w'=w^*+\Delta^*} < 0 \text{ and } \left.\frac{d\Pi_M(w)}{dw}\right|_{w'=w^{**}} = 0$$

But, from Lemma 10, $\Pi_M(w)$ is a strictly concave function so $w^{**}<w^*+\Delta^*$. From Lemma 20, for any w, $\Delta^*(w)<\Delta_{crit}$. Therefore at w^*, $\Delta^*=\Delta^*(w^*)<\Delta_{crit}$, or $w^*+\Delta^*<w_{crit}$. Therefore, the optimal channel capacity is strictly greater when a quantity premium can be offered and thus the total channel profits, which depend only on the channel capacity, are strictly greater when a quantity premium can be offered.

As $w^*+\Delta^*<w_{crit}$, the total supply chain capacity is strictly less than that chosen by a central decision-maker.

Lemma 22:

The following continuous quantity premium price schedule is an optimal wholesale price schedule for the manufacturer,

$$\frac{dW(Q)}{dQ} = \frac{c_S - v_S F_X(Q)}{1 - F_X(Q)} + p_S$$

Furthermore, it completely coordinates the supply chain but leaves the supplier with an expected profit of zero.

Proof:

For a wholesale price schedule specified by $W'(Q)$, let $\Pi_S(y, W'(Q))$ be the supplier's expected profit from choosing a capacity of y if the manufacturer has infinite capacity.

$$\Pi_S(y, W'(Q)) = -c_S y + \int_a^y \left(W(x) - p_S x\right) f_X(x)dx + \left(W(y) - p_S y\right)[1 - F_X(y)]$$
$$+ v_S \int_a^y [y - x] f_X(x)dx$$

So,

$$\frac{d\Pi_S(y, W'(Q))}{dy} = -\left(c_S - v_S F_X(y)\right) + \left(W'(y) - p_S\right)[1 - F_X(y)]$$

Suppose,

$$W'(Q) = \frac{c_S - v_S F_X(Q)}{1 - F_X(Q)} + p_S$$

Then,

$$\frac{d\Pi_S(y, W'(Q))}{dy} = -\left(c_S - v_S F_X(y)\right) + \left(c_S - v_S F_X(y)\right) = 0 \quad \forall y$$

So the supplier receives an expected profit of zero for all capacity choices.

For a wholesale price schedule specified by $W'(Q)$, let $\Pi_M(y, W'(Q))$ be the manufacturer's expected profit from choosing a capacity of y if the supplier has infinite capacity.

$$\Pi_M(y,W'(Q)) = -c_M y + \int_a^y \left((r-p_M)x - W(x)\right) f_X(x)dx + \left((r-p_M)y - W(y)\right)[1-F_X(y)]$$
$$+ v_M \int_a^y [y-x] f_X(x)dx$$

So,

$$\frac{d\Pi_M(y,W'(Q))}{dy} = \left(r - p_M - W'(y)\right)[1-F_X(y)] - [c_M - v_M F_X(y)]$$

and,

$$\frac{d^2\Pi_M(y,W'(Q))}{dy^2} = -\left(r - p_M - v_M - W'(y)\right) f_X(y) - W''(y)[1-F_X(y)]$$

Suppose,

$$W'(Q) = \frac{c_S - v_S F_X(Q)}{1-F_X(Q)} + p_S$$

Then,

$$W''(Q) = \frac{(c_S - v_S) f_X(Q)}{(1-F_X(Q))^2}$$

Therefore,

$$\begin{aligned}
\frac{d\Pi_M(y,W'(Q))}{dy} &= \left(r - p_M - p_S - \frac{c_S - v_S F_X(y)}{1-F_X(y)}\right)[1-F_X(y)] - [c_M - v_M F_X(y)] \\
&= [1-F_X(y)](r - p_M - p_S) - c_S + v_S F_X(y) - c_M + v_M F_X(y) \\
&= (r - p_M - p_S - c_M - c_S) - F_X(y)(r - p_M - p_S - v_M - v_S)
\end{aligned}$$

and

$$\begin{aligned}
\frac{d^2\Pi_M(y,W'(Q))}{dy^2} &= -\left(r - p_M - p_S - v_M - \frac{c_S - v_S F_X(y)}{1-F_X(y)}\right) f_X(y) \\
&\quad - \frac{(c_S - v_S) f_X(y)}{(1-F_X(y))^2}[1-F_X(y)] \\
&= -\left(r - p_M - p_S - v_M - \frac{c_S - v_S F_X(y)}{1-F_X(y)}\right) f_X(y) \\
&\quad - \frac{(c_S - v_S) f_X(y)}{1-F_X(y)} \\
&= -\frac{f_X(y)}{1-F_X(y)}\left([1-F_X(y)](r - p_M - p_S - v_M - v_S)\right) \\
&= -f_X(y)(r - p_M - p_S - v_M - v_S) < 0
\end{aligned}$$

as r>p_M+p_S+v_S+v_S.

So, $\Pi_M(y,W'(Q))$ is concave in y and the first order condition is sufficient for optimality. The first order condition is given by,

$$\frac{d\Pi_M(y,W'(Q))}{dy} = 0 \Leftrightarrow F_X(y) = \left(\frac{r - p_M - p_S - c_M - c_S}{r - p_M - p_S - v_M - v_S}\right) \Leftrightarrow y^* = y_I$$

The manufacturer's optimal choice is the same as that of a central-decision-maker. The supplier is indifferent to its capacity choice from above, so it is willing to invest in this capacity also. The supply chain is completely coordinated. The manufacturer captures the total expected supply chain profit as the supplier's expected profit is zero.

Note that one can induce the supplier to choose y_I by introducing an arbitrarily small quadratic penalty into the wholesale price schedule.

Lemma 23:

In the N supplier exogenous wholesale price game, the manufacturer and suppliers choose their capacities to be $y_M^*=y_{S1}^*=\ldots=y_{SN}^*=min\{y_S^{min}(w_1,\ldots,w_N),y_M(w_{Tot})\}$.

Proof:

Supplier N never chooses a capacity larger than the supplier N-1's capacity choice, as its total sales are limited by the this capacity. From Lemma 3, supplier N's profit, $\Pi_{SN}(y_N)$, is concave with the maximum achieved at $y_{SN}(w_N)$. For $y_N<y_{SN}(w_N)$, $\Pi_{SN}(y_N)$ is increasing in y. Therefore, if supplier N-1 announces a capacity of y_{N-1}, then supplier N chooses a capacity of $min\{y_{N-1},y_{SN}(w_N)\}$. Supplier N-1 therefore does not choose a capacity larger than $y_{SN}(w_N)$. Repeating this argument for suppliers N-1 through 1, the suppliers choose their capacities equal to $min\{y_M,y_{S1}(w_1),\ldots,y_{SN}(w_N)\}$, where y_M is the capacity announced by the manufacturer. From Lemma 2, the manufacturer's profit, $\Pi_M(y)$, is concave with the maximum achieved at $y_M(w_{Tot})$. For $y<y_M(w_{Tot})$, $\Pi_M(y)$ is increasing in y. The manufacturer does not choose a capacity larger than the suppliers would be willing to choose. Therefore, the manufacturer chooses its capacity, $y_M^*=min\{y_M(w_{Tot}),y_{S1}(w_1),\ldots,y_{SN}(w_N)\}$. The suppliers then choose their capacities such that $y_{S1}^*=\ldots=y_{SN}^*=min\{y_M(w_{Tot}),y_{S1}(w_1),\ldots,y_{SN}(w_N)\}$. By definition $y_S^{min}(w_1,\ldots,w_N)=min_n\{y_{Sn}(w_n)$.

Lemma 24:

(i) $y_M(w_{Tot})=y_I^*$, the coordinated channel capacity, iff $w_{Tot} = w_{Tot}^{crit}$, where

$$w_{Tot}^{crit} = \frac{(r - p_M)(c_s^{Tot} - v_s^{Tot}) + (c_M - v_M)p_s^{Tot} + c_M v_s^{Tot} - v_M c_s^{Tot}}{c_M - v_M + c_s^{Tot} - v_s^{Tot}}$$

and $y_M(w_{Tot}) \geq y_I^*$ iff $w_{Tot} \leq w_{Tot}^{crit}$

(ii) $y_{Sn}(w_n)=y_I^*$ iff $w_n = w_n^{crit}$,where

$$w_n^{crit} = \frac{(r - p_M - p_s^{Tot})(c_{S_n} - v_{S_n}) + p_{S_n}(c_M - v_M + c_s^{Tot} - v_s^{Tot}) - c_S(v_M + v_s^{Tot}) + v_S(c_M + c_s^{Tot})}{c_M - v_M + c_s^{Tot} - v_s^{Tot}}$$

and $y_{Sn}(w_n) \geq y_I^*$ iff $w_n \geq w_n^{crit}$.

(iii) $\sum_{n=1}^{N} w_n^{crit} = w_{Tot}^{crit}$

(iv) Complete channel coordination occurs iff $w_n=w_n^{crit}$ $n=1,\ldots,N$.

Proof:

(i) $y_M(w_{Tot}) = F_X^{-1}\left(\frac{m_M - c_M}{m_M - v_M}\right) \quad y_{\mathrm{I}}^* = F_X^{-1}\left(\frac{m_I - c_I}{m_I - v_I}\right)$ where $m_M=r-w_{Tot}-p_M$, $m_I=r-p_M-p_S^{Tot}$, $c_I=c_M+p_S^{Tot}$, and $v_I=v_M+v_S^{Tot}$. So, $y_M(w) \geq y_I^*$ iff

$$\left(\frac{m_M - c_M}{m_M - v_M}\right) \geq \left(\frac{m_I - c_I}{m_I - v_I}\right)$$

or,

$$w_{Tot} \leq \frac{(r - p_M)(c_s^{Tot} - v_s^{Tot}) + (c_M - v_M)p_s^{Tot} + c_M v_s^{Tot} - v_M c_s^{Tot}}{c_M - v_M + c_s^{Tot} - v_s^{Tot}}$$

(ii)

$y_{S_n}(w_n) = F_X^{-1}\left(\frac{m_{S_n} - c_{S_n}}{m_{S_n} - v_{S_n}}\right) \quad y_{\mathrm{I}}^* = F_X^{-1}\left(\frac{m_I - c_I}{m_I - v_I}\right)$ where $m_{Sn}=w_n-p_n$. So, $y_S(w) \geq y_I^*$ iff

$$\left(\frac{m_{S_n} - c_{S_n}}{m_{S_n} - v_{S_n}}\right) \geq \left(\frac{m_I - c_I}{m_I - v_I}\right)$$

or,

$$w_n \geq \frac{(r - p_M - p_s^{Tot})(c_{S_n} - v_{S_n}) + p_{S_n}(c_M - v_M + c_s^{Tot} - v_s^{Tot}) - c_{S_n}(v_M + v_s^{Tot}) + v_{S_n}(c_M + c_s^{Tot})}{c_M - v_M + c_s^{Tot} - v_s^{Tot}}$$

(iii)

$$\sum_{n=1}^{N} w_n^{crit}$$

$$= \sum_{n=1}^{N}\left(\frac{(r-p_M-p_s^{Tot})(c_{S_n}-v_{S_n})+p_{S_n}(c_M-v_M+c_s^{Tot}-v_s^{Tot})-c_{S_n}(v_M+v_s^{Tot})+v_{S_n}(c_M+c_s^{Tot})}{c_M-v_M+c_s^{Tot}-v_s^{Tot}}\right)$$

$$=\frac{(r-p_M)(c_S^{Tot}-v_S^{Tot})+(c_M-v_M)p_S^{Tot}+c_M v_S^{Tot}-v_M c_S^{Tot}}{c_M-v_M+c_s^{Tot}-v_s^{Tot}}=w_{Tot}^{crit}$$

(iv) Complete channel coordination occurs iff the supply chain capacity choice is equal to y_I*. In other words, from Lemma 37, one needs $\min\{y_{S1}(w_1),..,y_{SN}(w_N),y_M(w_{Tot})\}=y_I^*$. This is true iff (a) $y_M(w_{Tot})\geq y_I^*$, (b) $y_{Sn}(w_n)\geq y_I^*$, $n=1,\ldots,N$ and (c) one of the inequalities in (a) or (b) to hold with equality. From (i) above, $y_M(w_{Tot})\geq y_I^*$ iff $w_{Tot}\leq w_{Tot}^{crit}$. So (a) holds iff $w_{Tot}\leq w_{Tot}^{crit}$. From (ii) above, $y_{Sn}(w_n)\geq y_I^*$, $n=1,\ldots,N$ iff $w_n\geq w_n^{crit}$ $n=1,\ldots,N$. So (b) holds iff $w_n\geq w_n^{crit}$ $n=1,\ldots,N$. Therefore, using (iii) above, (a) and (b) both hold iff $w_n\geq w_n^{crit}$ $n=1,\ldots,N$ and $w_{Tot}=w_{Tot}^{crit}$. However, if some $w_n>w_n^{crit}$, then $w_{Tot}=w_{Tot}^{crit}$ iff some other $w_n<w_n^{crit}$. This cannot be if (a) and (b) both hold. Therefore (a) and (b) both hold iff $w_n=w_n^{crit}$ $n=1,\ldots,N$, which implies $w_{Tot}=w_{Tot}^{crit}$. For these wholesale prices the inequalities in (a) and (b) all hold with equality so (c) holds.

Lemma 25:

In the N supplier wholesale price game, if $F_X(x)$ satisfies the concavity condition, then

(i) The manufacturer's expected profit as a function of the suppler 1 wholesale price w, $\Pi^N_M(w)$, restricted to $p_{S1}+c_{S1}<w\leq w_1^{crit}$, is a strictly concave function.

(ii) The optimal wholesale prices w_n* are strictly less than w_n^{crit}, n=1,…,N.

(iii) The channel fails to be completely coordinated.

Proof

(i)

$$\Pi_M^N(w) = -c_M y_{S_1}(w) + (r - p_M - G_1(w)) \int_a^{y_{S_1}(w)} x f_X(x)dx$$
$$+ (r - p_M - G_1(w)) y_S(w)[1 - F_X(y_{S_1}(w))] + v_M \int_a^{y_{S_1}(w)} [y_{S_1}(w) - x] f_X(x)dx$$

Taking the derivative with respect to w,

$$\frac{d\Pi_M^N(w)}{dw} = \left(\frac{dy_{S_1}(w)}{dw}\right)\big((r - p_M - G(w) - c_M) - (r - p_M - G(w) - v_M)F_X(y_{S_1}(w))\big)$$
$$- \mathrm{A}_1 \int_{\mathrm{a}}^{y_{S_1}(w)} x f_X(x)dx - \mathrm{A}_1 y_{S_1}(w)[1 - F_X(y_{S_1}(w))]$$

Taking the second derivative,

$$\frac{d^2\Pi_M^N(w)}{dw^2} = \left(\frac{d^2 y_{S_1}(w)}{dw^2}\right)\big((r - p_M - G(w) - c_M) - (r - p_M - G(w) - v_M)F_X(y_{S_1}(w))]\big)$$
$$- 2A_1\left(\frac{dy_{S_1}(w)}{dw}\right)[1\text{-}F_X(y_{S_1}(w)] - (r - p_M - G(w) - v_M)\ f_X(y_{S_1}(w))\left(\frac{dy_{S_1}(w)}{dw}\right)^2$$

Now,

(i) From Lemma 4(ii), $\frac{dy_{S_1}(w)}{dw} > 0$

(ii) As $F_X(x)$ satisfies the concavity condition, then $\frac{d^2 y_{S_1}(w)}{dw^2} \leq 0$

(iii) $r>w+p_M+c_M$ in the allowable range for w,

(iv) $c_M>v_M$

(v) $F_X(y_S(w))\leq(r-p_M-G(w)-c_M)/(r-p_M-G(w)-v_M)<1$

(vi) $A_1>0$

Using (i)-(vi), $\frac{d^2\Pi_M^N(w)}{dw^2} < 0$

(ii)

$$\left.\frac{d\Pi_M^N(w)}{dw}\right|_{w=w_1^{crit}} = -\mathrm{A}_1 \int_{\mathrm{a}}^{y_{S_1}(w_1^{crit})} x f_X(x)dx - \mathrm{A}_1 y_{S_1}(w_1^{crit})[1 - F_X(y_{S_1}(w_1^{crit}))] < 0$$

As $\Pi^N_M(w)$ is concave function, w^* must be strictly less than w_1^{crit}. The optimal wholesale price for supplier 1 is $w_1^*=w^*$. The optimal wholesale prices for $w_2,\ldots,w_N$, are given by $T^n{}_1(w_n^*)=w_1^*$, $n=2,\ldots,N$. $T^n{}_1(w_n)$ is a one-to-one mapping such that $y_{Sn}(w_n)=y_{S1}(T^n{}_1(w_n))$. $T^n{}_1(w_n)$ is strictly increasing in w^n. $y_{Sn}(w_n{}^{crit})=y_{S1}(w_1{}^{crit})$, therefore $T^n{}_1(w_n{}^{crit})=w_1{}^{crit}$. As $w_1^*<w_1{}^{crit}$, $w_n^*<w_1{}^{crit}$, $n=1,\ldots,N$.

(iii) From Lemma 38(iv), the channel is only coordinated if $w_n=w_n{}^{crit}$ $n=1,\ldots,N$. Therefore, the channel is not completely coordinated when the manufacturer sets the prices to maximize its expected profit

Lemma 26:

(i) If $\beta_S<(r-c)/(r-v)$, then a central decision maker would invest in either medium or large capacity, investing in medium capacity if $\beta_L\leq(c-v)/(r-v)$ and in large capacity otherwise.

(ii) If $\beta_S<(r-c)/(r-v)$ and $\beta_L\leq(c-v)/(r-v)$, then a quantity premium price schedule of (c,Δ_M^*) induces the supplier to invest in medium capacity, where $\Delta^*_M=\left(\dfrac{\beta_S}{1-\beta_S}\right)(c-v)$. Furthermore, the channel is completely coordinated and the manufacturer captures all the expected supply chain profit.

(iii) If $\beta_S<(r-c)/(r-v)$ and $\beta_L>(c-v)/(r-v)$, then a quantity premium price schedule of $(c,\Delta_M^*,\Delta_L^*)$ induces the supplier to invest in large capacity, where Δ_M^* is given above and $\Delta^*_L=\left(\dfrac{1-\beta_L}{\beta_L}-\dfrac{\beta_S}{1-\beta_S}\right)(c-v)$. Furthermore, the channel is completely coordinated and the manufacturer captures all the expected supply chain profit.

Proof:

(i) Let $\Pi_I(K)$ denote the expected profit obtained by a central decision-maker that invests in a capacity of K. Then,

$$\Pi_I(S)=-cS+rS$$
$$\Pi_I(M)=-cM+\beta_S\left[rS+v(M-S)\right]+(1-\beta_S)rM$$
$$\Pi_I(L)=-cL+\beta_S\left[rS+v(L-S)\right]+\beta_M\left[rM+v(L-M)\right]+\beta_L L$$

A central decision-maker does not invest in small capacity if $\Pi_I(M)>\Pi_I(S)$.

$$\Pi_I(M) > \Pi_I(S) \Leftrightarrow -cM + \beta_S\left[rS + v(M-S)\right] + (1-\beta_S)rM > -cS + rS$$
$$\Leftrightarrow \beta_S < \frac{r-c}{r-v}$$

If $\beta_S<(r\text{-}c)/(r\text{-}v)$, then a central decision maker invests in either medium or large capacity. For this range of β_S it invests in medium capacity iff $\Pi_I(M)>\Pi_I(L)$.

$$\Pi_I(M) > \Pi_I(L) \Leftrightarrow -c(L-M) + (\beta_S + \beta_M)v(L-M) + \beta_L r(L-M) < 0$$
$$\Leftrightarrow \beta_L < \frac{c-v}{r-v}$$

(ii) Let the manufacturer offer the following quantity premium price schedule, (c,Δ_M,Δ_L). Let $\Pi_S(K,c,\Delta_M,\Delta_L)$ denote the expected profit obtained by the supplier if it invests in a capacity of K, for this price schedule.

$$\begin{aligned}
\Pi_S(S,c,\Delta_M,\Delta_L) &= -cS + cS = 0 \\
\Pi_S(M,c,\Delta_M,\Delta_L) &= -cM + \beta_S\left[cS + v(M-S)\right] + (1-\beta_S)\left[cS + (c+\Delta_M)(M-S)\right] \\
\Pi_S(L,c,\Delta_M,\Delta_L) &= -cL + \beta_S\left[cS + v(L-S)\right] + \beta_M\left[cS + (c+\Delta_M)(M-S) + v(L-M)\right] \\
&\quad + \beta_L\left[cS + (c+\Delta_M)(M-S) + (c+\Delta_M+\Delta_L)(L-M)\right] \\
&= (r-v)M - (c-v)B - \beta_S(r-v)(M-S) + \beta_L(r-v)(B-M)
\end{aligned}$$

The supplier prefers (or be indifferent to) M to S, if $\Pi_S(M,c,\Delta_M,\Delta_L)\geq\Pi_S(S,c,\Delta_M,\Delta_L)$.

$$\Pi_S(M,c,\Delta_M,\Delta_L) \geq \Pi_S(S,c,\Delta_M,\Delta_L) \Leftrightarrow \Delta_M \geq \left(\frac{\beta_S}{1-\beta_S}\right)(c-v) \geq 0$$

The supplier prefers (or be indifferent to) L to M, if $\Pi_S(L,c,\Delta_M,\Delta_L)\geq\Pi_S(M,c,\Delta_M,\Delta_L)$.

$$\Pi_S(L,c,\Delta_M,\Delta_L) \geq \Pi_S(M,c,\Delta_M,\Delta_L) \Leftrightarrow \Delta_M + \Delta_L \geq \left(\frac{1-\beta_L}{\beta_L}\right)(c-v)$$
$$\Leftrightarrow \Delta_L \geq \left(\frac{1-\beta_L}{\beta_L} - \frac{\beta_S}{1-\beta_S}\right)(c-v) \geq 0$$

If $\beta_S<(r\text{-}c)/(r\text{-}v)$, $\beta_B\leq(c\text{-}v)/(r\text{-}v)$ and the manufacturer offers a quantity premium price schedule of $(c,\Delta_M^*,0)$, then $\Pi_S(S,c,\Delta_M,0)=\Pi_S(M,c,\Delta_M,0)>\Pi_S(L,c,\Delta_M,0)$, and the supplier invests in medium capacity. A central decision-maker would also invest in medium capacity and thus the channel is completely coordinated. The manufacturer's expected profit is given by,

$$\begin{aligned}
\Pi_M(M,c,\Delta_M,\Delta_L) &= \beta_S(r-c)S + (1-\beta_S)\left[(r-c)S + (r-c-\Delta_M^*)(M-S)\right] \\
&= -cM + \beta_S\left[rS + v(M-S)\right] + (1-\beta_S)rM \\
&= \Pi_I(M)
\end{aligned}$$

and thus the manufacturer's expected profit is the same as the total expected supply chain profit obtained by a central decision-maker. The manufacturer can do no better than this and is thus this quantity premium price schedule of $(c,\Delta_M^*,0)$ [or (c,Δ_M^*)] is optimal for the manufacturer.

(iii) If $\beta_S<(r-c)/(r-v)$, $\beta_L\leq(c-v)/(r-v)$ and the manufacturer offers a quantity premium price schedule of $(c,\Delta_M^*,\Delta_L^*)$, then $\Pi_S(S,c,\Delta_M^*,\Delta_L^*)=\Pi_S(c,\Delta_M^*,\Delta_L^*)=\Pi_S(c,\Delta_M^*,\Delta_L^*)$, and the supplier invests in large capacity. A central decision-maker would also invest in large capacity and thus the channel is completely coordinated. The manufacturer's expected profit is given by,

$$\begin{aligned}
\Pi_M(L,c,\Delta_M,\Delta_L) &= \beta_S(r-c)S+\beta_M\left[(r-c)S+(r-c-\Delta_M^*)(M-S)\right] \\
&\quad +\beta_L\left[(r-c)S+(r-c-\Delta_M^*)(M-S)+(r-c-\Delta_M^*-\Delta_L^*)(L-M)\right] \\
&= (r-v)M-(c-v)B-\beta_S(r-v)(M-S)-(1-\beta_L)(r-v)(L-M) \\
&= \Pi_M(L)
\end{aligned}$$

and thus the manufacturer's expected profit is the same as the total expected supply chain profit obtained by a central decision-maker. The manufacturer can do no better than this and is thus this quantity premium price schedule of $(c,\Delta_M^*,\Delta_L^*)$ is optimal for the manufacturer.

This page intentionally left blank

5.2 Appendix 2

This appendix contains proofs of the lemmas from Chapter 3.

Lemma 1:

(i) A lower bound for the minimum shortfall in problem **P1** is given by problem **P2**,

$$\underset{M}{\text{Max}}\ \{\sum_{i\in M} d_i - \min_{L_1,..L_K}\{\sum_{k=1}^{K}\sum_{j\in P^k(L^k)} c_j^k\}\}$$

subject to

(i) $M \subseteq \{\varnothing,1,\ldots,I\}$

(ii) $L_k \cap L_{k'} = \varnothing \quad \forall k \neq k'$

(iii) $\bigcup_{k=1}^{K} L_k = M$

(ii) If either the number of stages K or the number of products I is less than three, then the minimum shortfall in problem **P1** is equal to the lower bound in (i).

Proof:

(i) **P1** is given by,

$$\min_{\mathbf{x,s}}\{\sum_{i=1}^{I} s_i\}$$

subject to

1. $\sum_{j\in P^k(i)} x_{ij}^k + s_i \geq d_i \qquad i=1,\ldots,I \quad k=1,\ldots,K$

2. $\sum_{i\in Q^k(j)} x_{ij}^k \leq c_j^k \qquad j=1,\ldots,J_k \quad k=1,\ldots,K$

$\mathbf{x,s} \geq \mathbf{0}$

Let π_i^k be the dual variables for the Type 1 constraints and μ_j^k be the dual variables for the Type 2 constraints. Letting $\nu_j^k=-\mu_j^k$ gives us the following dual problem **D1**,

$$\underset{\boldsymbol{\pi},\boldsymbol{\nu}}{\text{Max}}\ \{\sum_{k=1}^{K}\sum_{i=1}^{I}\pi_i^k d_i - \sum_{k=1}^{K}\sum_{j=1}^{J_k}\nu_j^k c_j^k\}$$

subject to

1. $\sum_{k=1}^{K}\pi_i^k \leq 1 \qquad i=1,\ldots,I$

2. $\pi_i^k \leq \nu_j^k \qquad j\in P^k(i) \quad i=1,\ldots,I \quad k=1,\ldots,K$

3. $\boldsymbol{\pi} \geq \mathbf{0} \quad , \quad \boldsymbol{\nu} \geq \mathbf{0}$

Let ***C*** be the set of solutions to **D1** that meets the following two conditions,

(i) $\boldsymbol{\pi}, \boldsymbol{\upsilon} \in \{0,1\}$

(ii) $\nu_j^k = \begin{cases} 1 \text{ if } \pi_i^k = 1 \text{ for some } i \in Q^k(j) \\ 0 \text{ otherwise} \end{cases}$

Each element in ***C*** is a feasible solution to **D1**.

Let ***E*** be the set of feasible solutions to **P2**. A feasible solution ($M, L_1, .., L_K$) to **P2** has the following objective value,

$$\sum_{i \in M} d_i - \sum_{k=1}^{K} \sum_{j \in P^k(L^k)} c_j^k$$

There is a one-to-one correspondence between elements in ***C*** and elements in ***E***; each element in ***C*** has a corresponding element in ***E*** with the same objective value and vice versa. To see this, consider an element ($M, L_1, .., L_K$) of ***E***. This can be mapped into an element of ***C*** as follows. For k=1,…,K set $\pi_i^k = 1 \ \forall \ i \in L_k$, $\pi_i^k = 0 \ \forall \ i \notin L_k$, $\nu_j^k = 1 \ \forall \ j \in P^k(L_k)$ and $\nu_j^k = 0 \ \forall \ j \notin P^k(L_k)$.

The objective value this element of ***C*** is,

$$\sum_{k=1}^{K} \sum_{i \in L^k} d_i - \sum_{k=1}^{K} \sum_{j \in P^k(L^k)} c_j^k = \sum_{i \in M} d_i - \sum_{k=1}^{K} \sum_{j \in P^k(L^k)} c_j^k$$

This is the same as the objective value for ($M, L_1, .., L_K$). Similarly any element of ***C*** can be mapped into an element of ***E*** by setting L_k to be the set of all products i with $\pi_i^k = 1$ k=1,…,K. Note that because of the Type 1 constraints, at most one π_i^k can equal 1 for each i=1,...,I, so that $L_k \cap L_{k'} = \varnothing \ \forall k \neq k'$. Again the objective value of this element of ***E*** is equal to the objective value of the element of ***C***.

Therefore, each feasible solution to **P2** corresponds to a feasible solution to **D1.** The objective value of such a solution gives a lower bound on the optimal value of **D1,** and hence from duality a lower bound on the minimum shortfall objective value of **P1**. The optimum value to **P2** is the maximum such lower bound.

(ii) From duality, the optimal solution to **P1** must equal the optimal solution to **D1**. From part (i), **P2** gives the optimal solution to **D1** subject to,

(i) $\boldsymbol{\pi}, \boldsymbol{\upsilon} \in \{0,1\}$

(ii) $\nu_j^k = \begin{cases} 1 \text{ if } \pi_i^k = 1 \text{ for some } i \in Q^k(j) \\ 0 \text{ otherwise} \end{cases}$

If the optimal solution to **D1** can be shown to satisfy both (i) and (ii), then **P2** gives the optimal solution to **D1** and hence the optimal solution to **P1**.

Let (π,ν) be a feasible solution to **D1** such that some $\nu_j^k > \pi_i^k \; \forall \; i \in Q^k(j)$. The objective function can increased by decreasing ν_j^k until $\nu_j^k = \max_{i \in Q^k(j)} \{\pi_i^k\}$ without violating any constraint. Decreasing ν_j^k any further violates a Type 2 constraint. Therefore, an optimal solution must satisfy $\nu_j^k * = \max_{i \in Q^k(j)} \{\pi_i^k *\}$. Assume for the moment that the optimal solution to **D1** is binary, i.e. satisfies condition (i). Then, condition (ii) must hold as $\nu_j^k * = \max_{i \in Q^k(j)} \{\pi_i^k *\}$.

All that remains to be shown is that the optimal solution to **D1** is binary. Substitute $y_j^k = 1 - \nu_j^k$ into problem **D1**. The following problem **D2** is obtained.

$$\underset{\boldsymbol{\pi},\mathbf{v}}{\text{Max}} \; \{\sum_{k=1}^{K}\sum_{i=1}^{I} \pi_i^k d_i + \sum_{k=1}^{K}\sum_{j=1}^{J_k} y_j^k c_j^k - \sum_{k=1}^{K}\sum_{j=1}^{J_k} c_j^k\}$$

subject to

1. $\sum_{k=1}^{K} \pi_i^k \le 1 \qquad i = 1,\ldots,I$
2. $\pi_i^k + y_j^k \le 1 \qquad j \in P^k(i) \quad i = 1,\ldots,I \quad k = 1,\ldots,K$
3. $y_j^k \le 1 \qquad j = 1,\ldots,J_k \quad k = 1,\ldots,K$
4. $\boldsymbol{\pi} \ge \mathbf{0}$

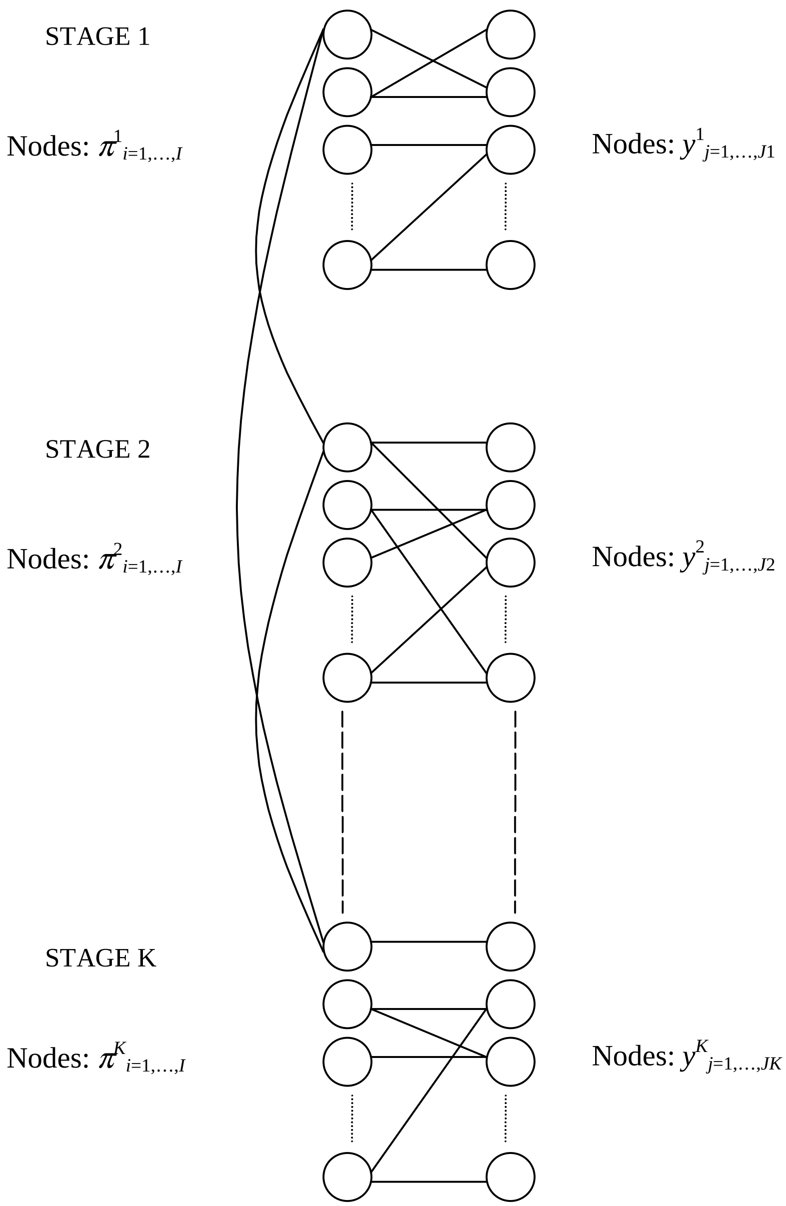
If y_j^k is negative, then the Type 2 constraints in which it appears are satisfied with strict inequality because from the Type 1 constraints $\pi_i^k \le 1$. Set y_j^k to zero. This solution remains feasible. The objective function is strictly increased. Therefore **D2** can be restricted to $\mathbf{y} \ge \mathbf{0}$ without any affect on the optimal solution. The upper bound constraints on the **y** variables (Type 3) can be ignored as the non-negativity of the $\boldsymbol{\pi}$ variables along with the Type 2 constraints ensure that the **y** upper bounds are not exceeded. Therefore **D2** can be solved by the following problem **D3**,

$$\underset{\boldsymbol{\pi},\mathbf{v}}{\text{Max}} \; \{\sum_{k=1}^{K}\sum_{i=1}^{I} \pi_i^k d_i + \sum_{k=1}^{K}\sum_{j=1}^{J_k} y_j^k c_j^k - \sum_{k=1}^{K}\sum_{j=1}^{J_k} c_j^k\}$$

subject to

1. $\sum_{k=1}^{K} \pi_i^k \le 1 \qquad i = 1,\ldots,I$
2. $\pi_i^k + y_j^k \le 1 \qquad j \in P^k(i) \quad i = 1,\ldots,I \quad k = 1,\ldots,K$
3. $\boldsymbol{\pi} \ge \mathbf{0} \;\; , \;\; \mathbf{y} \ge \mathbf{0}$

Let **A** be the constraint matrix of the linear program **D3**. **A** is the clique matrix of the undirected graph **G**, in which the π and **y** variables are nodes. Figure 1 shows an example of the graph **G**.



Note: The $(\pi^k_i, \pi^{k'}_i)$ arcs are only shown for i=1

Figure 1
The Constraint Matrix A is the Clique Matrix of the above Graph, G

A π_i^k variable is a node for product i at stage k and a y_j^k variable is a node for the j^{th} plant of stage k. An arc joins π_i^k to y_j^k $\forall j \in P^k(i)$, i.e. the product node i at stage k to all plants that can process product i at stage k. For each product i, there is an arc $(\pi_i^k, \pi_i^{k'})$ from node π_i^k to $\pi_i^{k'}$ $k'>k$, k=1,…,K. In other words the $\boldsymbol{\pi}$ variables for product i have arcs to all other $\boldsymbol{\pi}$ variables for product i. Note that I use the convention that $(\pi_i^k, \pi_i^{k'})$ arcs always have the smaller k value as the first node of the arc. There are no arcs joining the $\boldsymbol{\pi}$ variables of two different products. The π_i^k k=1,…,K for each product i form a clique as does each π_i^k to y_j^k arc.

If the number of stages K is less than three, then **G** is a bipartite graph. To see this, group all π_i^1 and y_j^2 nodes in Set 1. Group all π_i^2 and y_j^1 nodes in Set 2. The only arcs in **G** are those joining a node in Set 1 to a node in Set 2. As **G** is bipartite, it is a perfect graph (Nemhauser and Wolsley, 1988).

If the number of products I is less than three, then the graph G is again perfect. The proof of for this case is a little more involved and is most easily understood by referring to Figure 1.

- Consider a cycle with no $(\pi_1^k, \pi_1^{k'})$ type arcs and no $(\pi_2^k, \pi_2^{k'})$ type arcs. This cycle must contain only (π_i^k, y_j^k) type arcs for a single stage k. Such a cycle must have an even number of arcs as each stage k's subgraph is a bipartite.
- There is no cycle with exactly one $(\pi_1^k, \pi_1^{k'})$ type arc and no $(\pi_2^k, \pi_2^{k'})$ type arcs as such a cycle would leave the set of $\boldsymbol{\pi}$ nodes for stage k and never return. Likewise there is no cycle with exactly one $(\pi_2^k, \pi_2^{k'})$ type arc and no $(\pi_1^k, \pi_1^{k'})$ type arcs.
- Consider a cycle exactly one $(\pi_1^{k_1}, \pi_1^{k_1'})$ type arc and exactly one $(\pi_2^{k_2}, \pi_2^{k_2'})$ type arc, where the subscript on the stage k denotes the product i to which it refers. We must have k_1=k_2 and k_1'=k_2' as otherwise the "cycle" would leave the set of $\boldsymbol{\pi}$ nodes for stage k_1 and never return. Clearly this would not be a cycle. Any cycle in which k_1=k_2 and k_1'=k_2' must have an even number of arcs.

Therefore, any cycle that contains no more than one $(\pi_1^k, \pi_1^{k'})$ type arc and no more than one $(\pi_2^k, \pi_2^{k'})$ type arcs must have an even number of arcs. So any odd length cycle must contain at least two $(\pi_i^k, \pi_i^{k'})$ type arcs for i=1 or 2. Consider a cycle containing at least two $(\pi_i^k, \pi_i^{k'})$ type arcs. The only clique of graph **G** that contains $(\pi_i^k, \pi_i^{k'})$ arcs is the clique $\{\pi_i^1, \pi_i^2, \ldots, \pi_i^K\}$. This clique contains all such $(\pi_i^k, \pi_i^{k'})$ arcs and therefore contains another arc of the cycle. Therefore from Theorem 5.17 of Nemhauser and Wolsley (1988), the graph **G** is perfect.

So if either the number of products K or number of stages I is less than three, then **G** is a perfect graph. Therefore the polyhedron defined by **D3** ($Ax \le 1$) is integral. The optimal solution to **D3** is thus integral. In fact it is binary, because of the right hand side values. This in turn implies that the optimal solution to **D1** is binary.

Lemma 2:

(i) For any subset of products, M, define the problem **P3(*M*)** as

$$\underset{L_1,\ldots,L_k}{\text{Min}} \{\sum_{k=1}^{K} \sum_{j\in P^k(L^k)} c_j^k\}$$

subject to

(i) $L_k \cap L_{k'} = \emptyset \quad \forall k \neq k'$

(ii) $\bigcup_{k=1}^{K} L_k = M$

If for every possible M, there exists an optimal solution to **P3**(*M*) with only one non-empty L_k*, then a stage-spanning bottleneck can never occur in this case.

(ii) If $A_{\min} \geq \frac{TC_{\max}}{2}$, where $A_{\min} = \underset{\substack{i=1,\ldots,I \\ k=1,\ldots,K}}{\text{Min}} \{ \sum_{j\in P^k(i)} c_j^k \}$ and $TC_{\max} = \underset{k=1,\ldots,K}{\text{Max}} \{TC_k\}$,then a stage-spanning bottleneck can never occur. Note that $A_{\min}$ is the minimum total capacity available to any product at any stage and $TC_{\max}$ is the maximum total stage capacity across all stages.

Proof:

(i) Let **d**={$d_1,\ldots,d_I$} be any demand realization. Let *M** be the optimal M set for problem **P2** given this demand realization **d**. If for every possible subset of products, *M*, there exists an optimal solution to **P3**(*M*) with only one non-empty L_k*, then for the optimal set *M** there exists an optimal solution to **P3**(*M**) with only one non-empty L_k*. **P3**(*M*) is the internal minimization in **P2** and therefore if **P3**(*M**) has only one non-empty L_k*, then **P2** has only one non-empty L_k*. By definition a stage-spanning bottleneck does not occur for this demand realization, **d**. This is true for any demand realization and so a stage-spanning bottleneck can never occur.

(ii) For any subset of products, *M*, let Λ(*M*) be the set of stages with non-empty L_k*, where L_k* are the minimizing sets in **P3**(M), subject to there being at least two non-empty L_k*. Then |Λ(*M*)|≥2.

$$\sum_{k=1}^{K} \sum_{j\in P^{k'}(L_k^*)} c_j^k \geq \sum_{k\notin\Lambda(M)} 0 + \sum_{k\in\Lambda(M)} A_{\min} \geq 2A_{\min}$$

Construct a set of L_k, k=1,..K, as follows. Set $L_1^{new} = \bigcup_{k=1}^{K} L_k^*$ and $L_k^{new} = \varnothing$, k=2,…,K. The objective value of this new set of L_k is bounded above by $TC_{\max}$, the maximum total capacity of any stage. As $TC_{\max} \le 2A_{\min}$, then the new set has an objective value for **P3**(M) at least as small as the original L_k* set, and from part (i) a stage-spanning bottlenecks never occur if $TC_{\max} \le 2A_{\min}$.

Note that if one defines an N-stage-spanning bottleneck as one in which there are exactly N non-empty L_k*, then one can adapt the above proof directly to show that such a bottleneck can never occur if $TC_{\max} \le NA_{\min}$.

Lemma 3:

If a supply chain is $g_{\min}$-type, then (i) a stage-spanning bottleneck can never occur if the total number of products, I, is less than or equal to $2g_{\min}$. Furthermore, if at each stage each individual product is connected to the same total capacity, then (ii) a stage-spanning bottleneck can never occur if the total number of products, I, is less than or equal to $2(g_{\min}+1)$

Proof:

(i) Let $W^k(L_k)$ be the total capacity available to a subset L_k of product at stage k. As each stage k has a g-type configuration with $g_k \ge g_{\min}$, then $W^k(L_k)=0$ iff $L_k=\{\varnothing\}$ and $W^k(L_k) \ge \min\{TC_k, (|L_k|+g_{\min}-1)C_k\}$ iff $L_k \ne \{\varnothing\}$, where as defined earlier TC_k is the total capacity of the stage and $C_k=TC_k/I$. As $TC_{\min}$ is the minimum total stage capacity, $TC_k \ge TC_{\min}$ and $C_k \ge C_{\min}$, therefore $W^k(L_k) \ge \min\{TC_{\min}, (|L_k|+g_{\min}-1)C_{\min}\}$ iff $L_k \ne \{\varnothing\}$. For any subset of products, M, let $\Lambda(M)$ be the set of stages with non-empty L_k*, where L_k* are the minimizing sets in **P3**(M), subject to there being at least two non-empty L_k*. Then $|\Lambda(M)| \ge 2$. For this set of L_k*'s, the objective value for **P3**(M) is given by,

$$\sum_{k=1}^{K} \sum_{j \in P^{k'}(L_k^*)} c_j^k \ge \sum_{k \notin \Lambda(M^*)} W^k(L_k^*) + \sum_{k \in \Lambda(M^*)} W^k(L_k^*) \ge \sum_{k \in \Lambda(M^*)} \min\left\{TC_{\min}, \left(\left|L_k^*\right| + g_{\min} - 1\right)C_{\min}\right\} \tag{L3.1}$$

As the L_k* are non-empty for all $k \in \Lambda(M)$, then $|L_k^*| \ge 1$ for all $k \in \Lambda(M)$. Therefore,

$$\begin{aligned}\sum_{k=1}^{K} \sum_{j \in P^{k'}(L_k^*)} c_j^k &\ge \sum_{k \in \Lambda(M^*)} \min\{TC_{\min}, g_{\min} C_{\min}\} \\ &\ge 2\min\{TC_{\min}, g_{\min} C_{\min}\} \\ &\ge \min\{TC_{\min}, 2g_{\min} C_{\min}\}\end{aligned}$$

Without loss of generality, assume the first stage has the minimum total stage capacity. Construct a new set of L_k, k=1,…,K, as follows. Set $L_1^{new} = \bigcup_{k=1}^{K} L_k^*$ and $L_k^{new} = \varnothing$, k=2,…,K. The objective value of this new set of L_k is bounded above by $TC_{\min}$, the total capacity of stage 1. If $TC_{\min} \leq 2g_{\min}C_{\min}$, or alternatively $I \leq 2g_{\min}$ as $TC_{\min}=IC_{\min}$, then the new set has an objective value for **P3**(M) at least as small as the original L_k* set. Therefore for every possible M, there exists an optimal solution to **P3**(M) with only one non-empty L_k*. Following Lemma 2(a), a stage-spanning bottlenecks can never occur if $I \leq 2g_{\min}$.

Note that if one defines an N-stage-spanning bottleneck as one in which there are exactly N non-empty L_k*, then one can adapt the above proof directly to show that such a bottleneck can never occur if $I \leq Ng_{\min}$.

(ii) If $L_{k'}$* and $L_{k''}$* are non-empty with $L_{k''}$*={i}, i.e. it contains exactly one product, set $L_{k'}^{\text{new}} = L_{k'}^* \cup L_{k''}^*$ and $L_{k''}^{\text{new}} = \{\varnothing\}$. Then,

$$\sum_{j \in P^{k'}(L_{k'}^{new})} c_j^k \leq \sum_{j \in P^{k'}(L_{k'}^*)} c_j^k + \sum_{j \in P^{k'}(i)} c_j^k = \sum_{j \in P^{k'}(L_{k'}^*)} c_j^k + \sum_{j \in P^{k''}(L_{k''}^*)} c_j^k$$

where the equality occurs because at each stage each individual product is connected to the same total capacity. So, a new set of L_k can be constructed with an optimal value to **P3**(M) at least as small as the original optimum. This is true for any $|L_k^*|=1$. Therefore any optimal set for **P2** with N>1 non-empty L_k* can be transformed into a new optimal set with N-1 non-empty L_k^{new} if some $|L_k^*|=1$. Repeat this process until all non-empty L_k* have $|L_k^*| \geq 2$. So, $|L_k^*| \geq 2$ for any stage-spanning bottleneck. Substituting this into (L3.1) yields a lower bound of $2(g_{\min}+1)C_{\min}$. The rest of proof follows as above.

Note that if one defines an N-stage-spanning bottleneck as one in which there are exactly N non-empty L_k*, then one can adapt the above proof directly to show that such a bottleneck can never occur if $I \leq N(g_{\min}+1)$.

Lemma 4:

If a supply chain is a g_{min}-type and has the following properties,

(i) each stage has a total capacity of at least the total expected demand

(ii) the demands for the I products are independent and identically distributed $N(\mu,\sigma)$

then the probability of any particular LB stage-spanning bottleneck is bounded above by $\Omega_S(I,g_{min})$, where,

$$\Omega(I,g_{\min}) = \Phi\left(\frac{-2(g_{\min}-1)\mu}{\sigma\sqrt{I/2}}\right)^2$$

Proof:

From Section 3.3.1.1, an upper bound on the probability of $(M,L_1,\ldots,L_K)$ being a stage-spanning bottleneck is given by,

$$\Omega_S(M,L_1,\ldots,L_K) = [1-\Phi(z_1)]\Phi(z_2)$$

where

$$z_1 = \frac{\sum_{n=1}^{N}\sum_{j\in P^{k_n}(L_{k_n})} c_j^{k_n} - \sum_{i\in M}\mu_i}{\sqrt{\sum_{i\in M}\sigma_i}} \text{ and } z_2 = \frac{TC_{\min} - \sum_{n=1}^{N}\sum_{P^{k_n}(L_{k_n})} c_j^{k_n} - \sum_{i\notin M}\mu_i}{\sqrt{\sum_{i\notin M}\sigma_i}}$$

and the N stages with non-empty L_k are denoted by $k_1,\ldots,k_N$.

From Lemma 5 below, as each stage in the supply chain has a g-value greater than or equal to g_{min}, then

$$\Omega_S(M,L_1,\ldots,L_K) = [1-\Phi(z_1)]\Phi(z_2) \le [1-\Phi(y_1)]\Phi(y_2) = \Omega_S(x,N,I,g_{\min})$$

where,

$$y_1 = \frac{C_{\min}(x+N(g_{\min}-1)) - \sum_{i\in M}\mu_i}{\sqrt{\sum_{i\in M}\sigma_i}}$$

$$y_2 = \frac{C_{\min}(I-x-N(g_{\min}-1)) - \sum_{i\notin M}\mu_i}{\sqrt{\sum_{i\notin M}\sigma_i}}$$

and x is the number of products in M, i.e. $x=|M|$ and C_{min} is equal to TC_{min}/I, where TC_{min} is the minimum total stage capacity.

As the product demands are iid, then,

$$y_1 = \frac{C_{\min}\left(x + N\left(g_{\min} - 1\right)\right) - x\mu}{\sigma\sqrt{x}} = \frac{N\left(g_{\min} - 1\right)C_{\min} + x(C_{\min} - \mu)}{\sigma\sqrt{x}}$$

$$y_2 = \frac{C_{\min}\left(I - x - N\left(g_{\min} - 1\right)\right)c - (I - x)\mu}{\sigma\sqrt{(I - x)}} = \frac{(I - x)(C_{\min} - \mu) - N\left(g_{\min} - 1\right)C_{\min}}{\sigma\sqrt{(I - x)}}$$

$\Omega_S(x,N,I,g_{\min})=[1-\Phi(y_1)]\Phi(y_2)$ provides an upper bound on $\Omega_S(M,L_1,\ldots,L_K)$, which itself is an upper bound on the probability that $(M,L_1,\ldots,L_K)$ is a stage-spanning bottleneck. Note that this upper bound does not depend on the actual M set and L_k subsets, only on the number of products, x, in M and the number of non-empty L_k subsets, N. As such it is valid for any $(M,L_1,\ldots,L_K)$ for which $|M|=x$ and for which there are N non-empty L_k subsets.

By maximizing $\Omega_S(x,N,I,g_{\min})$ over all possible x for which there can be N non-empty L_k subsets, the dependence of $\Omega_S(x,N,I,g_{\min})$ on x can be removed, to give $\Omega_S(N,I,g_{\min})$, an upper bound on the probability of occurrence of any particular stage-spanning bottleneck with N non-empty L_k subsets (in a supply chain that processes I products and that has a g-value of $g_{\min}$). From Lemma 6 below, x, the number of products in M, must be less than or equal to $I-N(g_{\min}-1)-1$, if M is to be a stage-spanning bottleneck with N non-empty L_k subsets. Therefore,

$$\Omega_S(N,I,g_{\min}) \leq \underset{x \leq I - N\left(g_{\min} - 1\right) - 1}{\mathrm{Max}} \left\{\Omega_S(x,N,I,g_{\min})\right\} = \underset{x \leq I - N\left(g_{\min} - 1\right) - 1}{\mathrm{Max}} \left\{\left[1 - \Phi\left(y_1\right)\right]\Phi\left(y_2\right)\right\}$$

If the minimum total stage capacity equals the expected total demand, then $C_{\min}$ equals the mean product demand, μ, and the maximum occurs at $x^*=I/2$, assuming that $I/2 \geq N(g_{\min}-1)+1$. In this case,

$$\Omega_S(N,I,g_{\min}) = \left[1 - \Phi\left(\frac{N\left(g_{\min} - 1\right)\mu}{\sigma\sqrt{I/2}}\right)\right]\Phi\left(\frac{-N\left(g_{\min} - 1\right)\mu}{\sigma\sqrt{I/2}}\right) = \Phi\left(\frac{-N\left(g_{\min} - 1\right)\mu}{\sigma\sqrt{I/2}}\right)^2 \tag{L4.1}$$

Note that as $\Omega_S(x,N,I,g_{min})$ is decreasing in $C_{\min}$, (L4.1) provides an upper bound on $\Omega_S(N,I,g_{min})$ for any $C_{\min} \geq \mu$ (i.e. any supply chain in which the minimum total stage capacity is greater than or equal to the total expected demand).

$\Omega_S(N,I,g_{min})$ is decreasing in N, i.e. $\Omega_S(2,I,g_{min}) > \Omega_S(N,I,g_{min}) \forall N>2$. By definition a stage-spanning bottleneck must have at least two non-empty L_k subsets, that is $N \geq 2$. Therefore, $\Omega_S(2,I,g_{min})$ gives and upper bound on the probability of occurrence of any particular stage-spanning bottleneck, regardless of the number of non-empty L_k subsets. So, setting $\Omega_S(I,g_{min})=\Omega_S(2,I,g_{min})$ proves the lemma.

Note, if one is interested in the probability of occurrence of an N-stage-spanning bottleneck, then $\Omega_S(N,I,g_{min})$ provides an upper bound on this probability.

Lemma 5:

If each stage in the supply chain has a g-value greater than or equal to $g_{\min}$, then

$$\Omega_S(M,L_1,\ldots,L_K) = [1-\Phi(z_1)]\Phi(z_2) \le [1-\Phi(y_1)]\Phi(y_2) = \Omega_S(x,N,I,g_{\min})$$

where,

$$y_1 = \frac{C_{\min}(x+N(g_{\min}-1)) - \sum_{i\in M}\mu_i}{\sqrt{\sum_{i\in M}\sigma_i}}$$

$$y_2 = \frac{C_{\min}(I-x-N(g_{\min}-1)) - \sum_{i\notin M}\mu_i}{\sqrt{\sum_{i\notin M}\sigma_i}}$$

and x is the number of products in M, i.e. $x=|M|$. $C_{\min}$ is equal to $TC_{\min}/I$, where $TC_{\min}$ is the minimum total stage capacity.

Proof:

Let $W^k(L_k)$ be the total capacity available to a non-empty subset, L_k, of products at stage k. As each stage k has a g-type configuration with $g_k \ge g_{\min}$, then from equation (11) $W^k(L_k) \ge \min\{TC_k,(|L_k|+g_{\min}-1)C_k\}$. As $TC_{\min}$ is the minimum total stage capacity, $TC_k \ge TC_{\min}$ and $C_k \ge C_{\min}$, therefore $W^k(L_k) \ge \min\{TC_{\min},(|L_k|+g_{\min}-1)C_{\min}\}$. From Lemma 6 below, the number of products in M, must be less than $I-N(g_{\min}-1)$, if M is to be a stage-spanning bottleneck with N non-empty L_k subsets. For each non-empty subset L_{kn}, $n=1,\ldots,N$, $|L_{kn}| \le |M| < I-N(g_{\min}-1)$. Therefore $|L_{kn}|+g_{min}-1<I$ and $(|L_{kn}|+g_{min}-1)C_{\min}<TC_{\min}$, so $\min\{TC_k,(|L_k|+g_{\min}-1)C_k\}=(|L_{kn}|+g_{min}-1)C_{\min}$. Therefore, $W^k(L_{kn}) \ge (|L_{kn}|+g_{\min}-1)C_{\min}$ for $n=1,\ldots,N$. So,

$$\sum_{n=1}^{N}\sum_{j\in P^{k_n}(L_{k_n})} c_j^{k_n} = \sum_{n=1}^{N} W^{k_n}(L_{k_n}) \ge \sum_{n=1}^{N}\left[\left(\left|L_{k_n}\right| + g_{\min}-1\right)C_{\min}\right] = \left(|M| + N(g_{\min}-1)\right)C_{\min}$$

Let $|M|=x$. Therefore

$$z_1 = \frac{\sum_{n=1}^{N}\sum_{j\in P^{k_n}(L_{k_n})} c_j^{k_n} - \sum_{i\in M}\mu_i}{\sqrt{\sum_{i\in M}\sigma_i}} \ge \frac{C_{\min}(x+N(g_{\min}-1)) - \sum_{i\in M}\mu_i}{\sqrt{\sum_{i\in M}\sigma_i}} = y_1$$

$$z_2 = \frac{TC_{\min} - \sum_{n=1}^{N} \sum_{P^{k_n}(L_{k_n})} c_j^{k_n} - \sum_{i \notin M^*} \mu_i}{\sqrt{\sum_{i \notin M} \sigma_i}} \le \frac{TC_{\min} - C_{\min}\left(x + N\left(g_{\min} - 1\right)\right) - \sum_{i \notin M} \mu_i}{\sqrt{\sum_{i \notin M^{**}} \sigma_i}}$$

$$= \frac{C_{\min}\left(I - x - N\left(g_{\min} - 1\right)\right) - \sum_{i \notin M} \mu_i}{\sqrt{\sum_{i \notin M} \sigma_i}} = y_2$$

as $TC_{\min}=IC_{\min}$. Now,

$$\Omega_S(M, L_1, \ldots, L_K) = [1 - \Phi(z_1)]\Phi(z_2) \le [1 - \Phi(y_1)]\Phi(y_2) = \Omega_S(x, N, I, g_{\min})$$

as $y_1 \le z_1$, $y_2 \ge z_2$ and $\Phi(z)$ is increasing in z.

Lemma 6:

If a supply chain is $g_{\min}$-type, then a stage-spanning bottleneck with N non-empty L_k* and $|M^*| \ge I - N(g_{\min}-1)$ can never occur.

Proof:

Let $(M^*, L_1^*, \ldots, L_K^*)$ be an LB stage-spanning bottleneck with N non-empty L_k* and $|M^*| > I - N(h-1)$. Let $W^k(L_k)$ be the total capacity available to a subset L_k of product at stage k. As each stage k has a g-type configuration with $g_k \ge g_{\min}$, then $W^k(L_k)=0$ iff $L_k=\{\varnothing\}$ and $W^k(L_k) \ge \min\{TC_k, (|L_k|+g_{\min}-1)C_k\}$ iff $L_k \ne \{\varnothing\}$, where as defined earlier TC_k is the total capacity of the stage and $C_k = TC_k/I$. As $TC_{\min}$ is the minimum total stage capacity, $TC_k \ge TC_{\min}$ and $C_k \ge C_{\min}$, therefore $W^k(L_k) \ge \min\{TC_{\min}, (|L_k|+g_{\min}-1)C_{\min}\}$ iff $L_k \ne \{\varnothing\}$. Let $\Lambda(M^*)$ be the set of stages with non-empty L_k*. The objective value for **P3**(M*) is,

$$\sum_{k=1}^{K} \sum_{j \in P^{k'}(L_k^*)} c_j^k \ge \sum_{k \notin \Lambda(M^*)} W^k(L_k^*) + \sum_{k \in \Lambda(M^*)} W^k(L_k^*)$$

$$\ge \sum_{k \in \Lambda(M^*)} \min\left\{TC_{\min}, \left(\left|L_k^*\right| + g_{\min} - 1\right)C_{\min}\right\}$$

$$\ge \min\left\{TC_{\min}, \sum_{k \in \Lambda(M^*)} \left[\left(\left|L_k^*\right| + g_{\min} - 1\right)C_{\min}\right]\right\}$$

$$= \min\left\{TC_{\min}, \left(\left|M^*\right| + N\left(g_{\min} - 1\right)\right)C_{\min}\right\}$$

Without loss of generality, assume the first stage has the minimum total stage capacity. Construct a new set of L_k, k=1,..K, as follows. Set $L_1^{new} = \bigcup_{k=1}^{K} L_k^*$ and $L_k^{new} = \varnothing$, k=2,…,K. The objective value of this new set of L_k is bounded above by $TC_{\min}$, the total capacity of stage 1. If $TC_{\min} \le (|M^*|+N(g_{\min}-1))C_{\min}$, then the new set has an objective value for **P3**(M) at least as small as the original L_k* set. Therefore for every possible M, there exists an optimal solution to **P3**(M) with only one non-empty L_k*. Following the addendum to Lemma 2(i) regarding N-stage-spanning bottlenecks, if $TC_{\min} \le (|M^*|+N(g_{\min}-1))C_{\min}$, then a stage-spanning bottlenecks with N non-empty L_k* can never occur. $TC_{\min}=IC_{\min}$ and therefore a stage-spanning bottlenecks with N non-empty L_k* can never occur if $I \le |M^*|+N(g_{\min}-1)$.

Lemma 7:

For a two-stage 4-product supply chain with each stage having 4 plants and a type h=2 chain configuration, and all plant capacities being equal (=c), if the product demands are iid $N(\mu,\sigma)$, then the probability that the stand alone shortfalls for the two stages are the same is greater than or equal to,

$$1-\left(8\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\left[\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)+\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\right]+8\Phi\left(\frac{\mu-2c}{\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu-c}{\sigma}\right)\right)$$

Proof:

Let SF_k denote the random variable for the shortfall for stage k, k=1,2. For a given demand realization $d_1,\ldots,d_I$, let m_k* be the maximizing set for,

$$\max_{m_k}\{\sum_{i\in m_k} d_i - \sum_{j\in P^k(m_k)} c_j^k\}$$
subject to $m \subseteq \{1,\ldots,I\}$

and let sf_k be the optimal value for stage k, k=1,2 (i.e. the shortfall for this realization). Let the indicator function I_{12} be such that I_{12}=0 if $m_1^*=m_2$ and I_{12}=1 if $m_1^*\neq m_2$. As $c^1_j=c^2_j$, j=1,…,J, $sf_1\neq sf_2$ only if $m_1^*\neq m_2^*$. Therefore $P[SF_1\neq SF_2]\le P[I_{12}=1]$. Because the demand distribution is continuous the probability that $M_1^*\neq M_2^*$ and $SF_1=SF_2$ is zero. Therefore in this case, $P[SF_1\neq SF_2]=[I_{12}=1]$. Note that the upper bound in the lemma is still valid if $P[SF_1\neq SF_2]\le P[I_{12}=1]$.

The only possible chains for a 4-product 4-plant stage are {1,2,3,4}, {1,2,4,3} and {1,3,2,4}. Let stage 1 have a {1,2,3,4} chain and stage 2 have a {1,2,4,3} chain. For any demand realization, the possible m_1* sets are {∅},{1},{2},{3},{4},{1,2},{2,3}, {3,4},{4,1},{1,2,3,4} and the possible m_2* sets are {∅},{1},{2},{3},{4},{1,2}, {2,4},{3,4},{1,3},{1,2,3,4}. The events in which I_{12}=1 ($m_1^* \neq m_2^*$) can be partitioned into the following mutually exclusive (and exhaustive) events:

Event 1: $m_1^*=\{\varnothing\} \cap m_2^* \neq \{\varnothing\}$

Event 2: $m_1^* \neq \{\varnothing\} \cap m_2^*=\{\varnothing\}$

Event 3: $m_1^*=\{1,2,3,4\} \cap m_2^* \neq \{\varnothing\} \cap m_2^* \neq \{1,2,3,4\}$

Event 4: $m_1^* \neq \{\varnothing\} \cap m_1^* \neq \{1,2,3,4\} \cap m_2^*=\{1,2,3,4\}$

Event 5: $m_1^*=\{i\} \cap m_2^* \neq \{i\} \cap m_2^* \neq \{1,2,3,4\} \cap m_2^* \neq \{\varnothing\}$ i=1,2,3,4

Event 6: $(m_1^*=\{1,2,3,4\} \cup m_1^*=\{\varnothing\}) \cap m_2^*=\{i\}$ i=1,2,3,4

Event 7: (a) $m_1^*=\{1,2\} \cap (m_2^* \in \{1,3\} \cup \{3,4\} \cup \{2,4\})$

(b) $m_1^*=\{2,3\} \cap (m_2^* \in \{1,3\} \cup \{3,4\} \cup \{2,4\} \cup \{1,2\})$

(c) $m_1^*=\{3,4\} \cap (m_2^* \in \{1,3\} \cup \{2,4\} \cup \{1,2\})$

(d) $m_1^*=\{4,1\} \cap (m_2^* \in \{1,3\} \cup \{3,4\} \cup \{2,4\} \cup \{1,2\})$

I now develop upper-bounds for the probability of each event.

Event 1: $\boldsymbol{m_1^*=\{\varnothing\}, m_2^* \neq \{\varnothing\}}$

$m_1^*=\{\varnothing\}$ implies:

$$d_i - 2c \le 0 \qquad \text{i} = 1,2,3,4$$
$$d_i + d_{i+1} - 3c \le 0 \qquad \text{i} = 1,2,3,4$$
$$d_1 + d_2 + d_3 + d_4 - 4c \le 0$$

where i+1=1 if i =4. Therefore, the only possible m_2* (≠{∅}) sets are {1,3} or {2,4}. For $m_1^*=\{\varnothing\}$ and $m_2^*=\{1,3\}$, the following four conditions are necessary (but not sufficient),

$$\text{(i) } d_i - 2c \le 0 \qquad \text{i} = 1,2,3,4$$
$$\text{(ii) } d_i + d_{i+1} - 3c \le 0 \qquad \text{i} = 1,2,3,4$$
$$\text{(iii) } d_1 + d_3 - 3c > 0$$
$$\text{(iv) } d_1 + d_2 + d_3 + d_4 - 4c \le 0$$

(iii) and (iv) imply that the following is a necessary condition for $m_1^*=\{\varnothing\}$ and $m_2^*=\{1,3\}$,

$$d_1 + d_3 > 3c \quad \text{and} \quad d_2 + d_4 \le c$$

As the demands are iid $N(\mu,\sigma)$, the probability of this event is given by,

$$\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)$$

and this is an upper-bound on P[m_1*={∅} and m_2*={1,3}]. As above, it can also be shown to be an upper-bound on P[m_1*={∅} and m_2*={2,4}]. Therefore,

$$\text{P[Event 1]} \leq 2\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)$$

Event 2: $\boldsymbol{m_1^* \neq \{\varnothing\} \cap m_2^* = \{\varnothing\}}$

Using the same derivation as for Event 1, but adapting for Event 2, the same upper-bound as 1 can be developed. Therefore,

$$\text{P[Event 2]} \leq 2\Phi\left(\frac{2\mu-3C}{\sqrt{2}\sigma}\right)\Phi\left(\frac{C-2\mu}{\sqrt{2}\sigma}\right)$$

Event 3: $\boldsymbol{m_1^* = \{1,2,3,4\} \cap m_2^* \neq \{\varnothing\} \cap m_2^* \neq \{1,2,3,4\}}$

m_1*={1,2,3,4} implies:

$$d_1 + d_2 + d_3 + d_4 - 4C > 0$$
$$d_1 + d_2 + d_3 + d_4 - 4C > d_i - 2c \qquad \text{i} = 1,2,3,4$$
$$d_1 + d_2 + d_3 + d_4 - 4C > d_i + d_{i+1} - 3c \qquad \text{i} = 1,2,3,4$$

Therefore, the only possible m_2* sets (≠{∅},≠{1,2,3,4}) or are {1,3} or {2,4}.

For m_1*={1,2,3,4} and m_2*={1,3}, the following two conditions are necessary (but not sufficient),

(i) $d_1 + d_3 - 3c > 0$
(ii) $d_1 + d_3 - 3c \geq d_1 + d_2 + d_3 + d_4 - 4c \Rightarrow d_2 + d_4 - c \leq 0$

As the demands are iid $N(\mu,\sigma)$, the probability of this event is given by,

$$\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)$$

and this is an upper-bound on P[m_1*={1,2,3,4} and m_2*={1,3}]. As above, it can also be shown to be an upper-bound on P[m_1*={1,2,3,4} and m_2*={2,4}]. Therefore,

$$\text{P[Event 3]} \leq 2\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)$$

Event 4: $\boldsymbol{m_1^* \neq \{\varnothing\} \cap m_1^* \neq \{1,2,3,4\} \cap m_2^* = \{1,2,3,4\}}$

Using the same derivation as for Event 3, but adapting for Event 4, the same upper-bound as 1 can be developed. Therefore,

$$\text{P[Event 4]} \leq 2\Phi\left(\frac{2\mu - 3c}{\sqrt{2}\sigma}\right)\Phi\left(\frac{c - 2\mu}{\sqrt{2}\sigma}\right)$$

Event 5: $\boldsymbol{m_1^* = \{i\} \cap m_2^* \neq \{i\} \cap m_2^* \neq \{1,2,3,4\} \cap m_2^* \neq \{\varnothing\}}$ **i=1,2,3,4**

m_1*={1} implies

$$\begin{aligned}
&d_1 - 2c \geq 0 \\
&d_i - 2c \leq d_1 - 2c && \text{i} = 2,3,4 \\
&d_i + d_{i+1} - 3c \leq d_1 - 2c && \text{i} = 1,2,3,4 \\
&d_1 + d_2 + d_3 + d_4 - 4c \leq d_1 - 2c
\end{aligned}$$

Canceling terms, then

(i) $d_1 - 2c \geq 0$

(ii) $d_i \leq d_1$ $\quad$ i = 2,3,4

(iii)(a) $d_2 - c \leq 0$ (b) $d_2 + d_3 - d_1 - c \leq 0$ (c) $d_3 + d_4 - d_1 - c \leq 0$ (d) $d_4 - c \leq 0$

(iv) $d_2 + d_3 + d_4 - 2c \leq 0$

Therefore, the only possible m_2* set (≠{1},{1,2,3,4},{∅}) is {1,3}. m_2* cannot be ={j}, j≠1, from (ii). m_2* cannot be {2,4} from (iv). For m_1*={1} and m_2*={1,3}, the following conditions are necessary (but not sufficient),

(v) $d_1 - 2c \geq 0$

(vi) $d_1 + d_3 - 3c \geq d_1 + d_2 + d_3 + d_4 - 4c \Rightarrow d_2 + d_4 - c \leq 0$

(vii) $d_1 + d_3 - 3c \geq d_1 - 2c \Rightarrow d_3 - c \geq 0$

As the demands are iid $N(\mu,\sigma)$, the probability of this event is given by,

$$\Phi\left(\frac{\mu - 2c}{\sigma}\right)\Phi\left(\frac{c - 2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu - c}{\sigma}\right)$$

and this is an upper-bound on P[m_1*={1} and m_2*≠{1},{1,2,3,4},{∅})]. The same upper-bound can be developed for P[m_1*={i} and m_2*≠{i},{1,2,3,4},{∅})], i=2,3,4. Therefore,

$$\text{P[Event 5]} \leq 4\Phi\left(\frac{\mu - 2c}{\sigma}\right)\Phi\left(\frac{c - 2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu - c}{\sigma}\right)$$

Event 6: $(m_1^*=\{1,2,3,4\} \cup m_1^*=\{\varnothing\}) \cap m_2^*=\{i\} \quad i=1,2,3,4$

Using the same derivation as for Event 3, but adapting for Event 4, the same upper-bound as 1 can be developed. Therefore,

$$\text{P[Event 6]} \leq 4\Phi\left(\frac{\mu-2c}{\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu-c}{\sigma}\right)$$

Event 7: **(a)** $m_1^*=\{1,2\} \cap (m_2^* \in \{1,3\} \cup \{3,4\} \cup \{2,4\})$

(b) $m_1^*=\{2,3\} \cap (m_2^* \in \{1,3\} \cup \{3,4\} \cup \{2,4\} \cup \{2,1\})$

(c) $m_1^*=\{3,4\} \cap (m_2^* \in \{1,3\} \cup \{2,4\} \cup \{2,1\})$

(d) $m_1^*=\{4,1\} \cap (m_2^* \in \{1,3\} \cup \{3,4\} \cup \{2,4\} \cup \{2,1\})$

(a) $m_1^*=\{1,2\}$ implies

$$\begin{aligned}
&d_1 + d_2 - 3c \geq 0 \\
&d_i - 2c \leq d_1 + d_2 - 3c && \text{i} = 2,3,4 \\
&d_i + d_{i+1} - 3c \leq d_1 + d_2 - 3c && \text{i} = 2,3,4 \\
&d_1 + d_2 + d_3 + d_4 - 4c \leq d_1 + d_2 - 3c
\end{aligned}$$

Therefore $m_2^* \neq \{3,4\}$. Only need to consider $m_2^*=\{1,3\}$ or $\{2,4\}$.

<u>$m_1^*=\{1,2\}$ $m_2^*=\{1,3\}$:</u>

Canceling terms in the above equations, then

(i) $d_1 + d_2 - 3c \geq 0$

(ii) (a) $d_2 - c \geq 0$ (b) $d_1 - c \geq 0$ (c) $d_1 + d_2 - d_3 - c \geq 0$ (c) $d_1 + d_2 - d_4 - c \geq 0$

(iii) (a) $d_1 - d_3 \geq 0$ (b) $d_1 + d_2 - d_3 - d_4 \geq 0$ (c) $d_2 - d_4 \geq 0$

(iv) $d_3 + d_4 - c \leq 0$

But, $m_2^*=\{1,3\}$ implies

(v) $d_1 + d_2 + d_3 + d_4 - 4c \leq d_1 + d_3 - 3c \Rightarrow d_2 + d_4 - c \leq 0$

Using (iv) and (v), then $d_2 + d_3 + 2d_4 - 2c \leq 0$. From (ii)(a) $d_2 \geq c$. Therefore,

(vi) $c + d_3 + 2d_4 - 2c \leq 0 \Rightarrow d_3 + 2d_4 - c \leq 0$.

(i) and (vi) are thus necessary conditions for $m_1^*=\{1,2\}$ and $m_1^*=\{1,3\}$. As the demands are iid $N(\mu,\sigma)$, the probability of this event is given by,

$$\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)$$

This is an upper-bound on the probability that m_1*={1,2} m_2*={1,3}. It can also be shown to be an upper-bound on m_1*={1,2} m_2*={2,4}. Therefore,

$$\text{P[Event 7(a)]} \le 2\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)$$

Events 7(b),(c) and (d) have similar upper bounds. Therefore,

$$\text{P[Event 7]} \le 8\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)$$

The probability of $I_{12}=1$ ($m_1^* \neq m_2^*$) is the sum of the probability of the above events. Therefore,

$\text{P}[I_{12}] \le \sum_{\text{n}=1}^{7} \text{P[Event n]}$. So,

$$\text{P}[I_{12}] \le 8\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\left[\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)+\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\right]+8\Phi\left(\frac{\mu-2c}{\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu-c}{\sigma}\right)$$

As, $\text{P}[SF_1 \neq SF_2] \le \text{P}[I_{12}=1]$,

$$\text{P}[SF_1 \neq SF_2] \le 8\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\left[\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)+\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\right]+8\Phi\left(\frac{\mu-2c}{\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu-c}{\sigma}\right)$$

and so,

$$\text{P}[SF_1 = SF_2] \ge 1-\left(8\Phi\left(\frac{2\mu-3c}{\sqrt{2}\sigma}\right)\left[\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)+\Phi\left(\frac{2c-3\mu}{\sqrt{5}\sigma}\right)\right]+8\Phi\left(\frac{\mu-2c}{\sigma}\right)\Phi\left(\frac{c-2\mu}{\sqrt{2}\sigma}\right)\Phi\left(\frac{\mu-c}{\sigma}\right)\right)$$

This proof can be repeated for the other possible chain pairings.

5.3 Appendix 3

This appendix contains proofs of the lemmas from Chapter 4 and an algorithm for generating the set of possible stage subsets over which the maximization in Lemma 1 is evaluated (c.f. Lemma 3).

Lemma 1:

(i) A lower bound on the minimum total shortfall in problem **P3** is given by

$$\max_{\Lambda \subseteq \{1,\ldots,K\}} \{ \sum_{f \in P(\Lambda)} d_f - \sum_{k \in \Lambda} c_k \}$$

(ii) If the path-stage matrix **B** is totally unimodular, then the minimum total shortfall in problem **P3** equals the lower bound in (i).

Proof:

Problem **P3** is given by the linear program,

$$\text{Min} \{ \sum_{f=1}^{f} sf_f \}$$

subject to

$$1.\ y_f + sf_f \geq d_f \qquad f \in F$$

$$2.\ \sum_{f \in P(k)} y_f \leq c_k \qquad k = 1,\ldots,K$$

$$3.\ \mathbf{y}, \mathbf{sf} \geq \mathbf{0}$$

Let π_f be the dual variable for the Type 1 constraints and μ_k be the dual variable for the Type 2 constraints. Letting ν_k=-μ_k gives us the following dual formulation (**D3**),

$$\underset{\boldsymbol{\pi},\mathbf{v}}{\text{Max}} \{ \sum_{f \in F} \pi_f d_f - \sum_{k=1}^{K} \nu_k c_k \}$$

subject to

$$1.\qquad \pi_f \leq 1 \qquad f \in F$$

$$2.\qquad \pi_f \leq \sum_{k \in Q(f)} \nu_k \qquad f \in F$$

$$3.\qquad \boldsymbol{\pi} \geq \mathbf{0} \quad , \quad \mathbf{v} \geq \mathbf{0}$$

Remember that $P(\Lambda)$ is the set of flow paths that are processed by at least one stage $k \in \Lambda$ and $Q(f)$ is the set of stages that process flow path f.

Let problem **D4** be the same as **D3** but with the additional constraint that the solution be binary. For a feasible solution $(\boldsymbol{\pi}, \mathbf{v})$, let Λ be the subset of stages k=1,…,K for which $v_k = 1$. This solution can be optimal only if $\pi_f = 1$ $\forall f \in P(\Lambda)$ and $\pi_f = 0$ $\forall f \notin P(\Lambda)$, where the second condition is required for feasibility. To see this, consider a solution to **D4** in which $\pi_f = 0$ for some flow path $f \in P(\Lambda)$. This solution can be improved upon by setting $\pi_f = 1$. This is a feasible solution with an increased objective function. Each of the possible optimal solutions is therefore completely specified by the subset Λ. The objective value for such a solution is given by,

$$\sum_{f \in P(\Lambda)} d_f - \sum_{k \in \Lambda} c_k$$

Any subset Λ of stages 1,…,K is a possible candidate for optimality and therefore the optimum objective value to **D4** is given by

$$\max_{\Lambda \subseteq \{1,\ldots,K\}} \{ \sum_{f \in P(\Lambda)} d_f - \sum_{k \in \Lambda} c_k \}$$

As all solutions to **D4** are feasible for **D3**, the optimal objective for **D4** is a lower bound on the optimal objective value to **D3**. From duality the minimum shortfall for **P3** is equal to the optimal objective value to **D3**.

(ii) The constraint matrix, **A**, for **D3** is given by,

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{B} \end{bmatrix}$$

where **B** is the path-stage matrix. The path-stage matrix is the matrix in which there is a row for each flow path, a column for each stage k and b_{fk}=1 if $k \in Q(f)$ and b_{fk}=0 if $k \notin Q(f)$. In other words, element (f,k) is 1 if flow path f requires stage k and 0 otherwise.

A is totally unimodular (TU) if **B** is TU. This follows from the fact that total unimodularity is preserved under the following operations (Schrijver, 1987)

(a) multiplying a column by –1

(b) adding a row or column with at most one nonzero, being +/-1.

If **B** is TU then **–B** is TU using (a). Then **[I –B]** is TU using (b) and **A** is TU using (b) again.

So, if the path-stage matrix **B** is TU, then the constraint matrix **A** for **D3** is TU and therefore optimal solution to **D3** is integral. The Type 1 constraints ensure that $\pi_f \leq 1 \; f \in F$. No optimal solution can have any $v_k > 1$ as the objective function can be decreased by setting such a $v_k = 1$

while still maintaining feasibility. The optimal solution to **D3** is therefore binary and thus from part (i), the optimal objective value for **D3** and **P3** is given by,

$$\max_{\Lambda \subseteq \{1,\ldots,K\}} \{ \sum_{f \in P(\Lambda)} d_f - \sum_{k \in \Lambda} c_k \}$$

Lemma 2:

The path-stage matrices for the Alcalde Job Shop and for Work Center A are both totally unimodular (TU).

Proof:

The path-stage matrix for the Alcalde job shop $\mathbf{B}_{Alcalde}$ is given by,

$$\mathbf{B}_{Alcalde} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Total unimodularity is preserved under addition of a column with at most one nonzero, being +/-1 (Schrijver, 1987). Therefore $\mathbf{B}_{Alcalde}$ is TU if the following submatrix **S** of $\mathbf{B}_{Alcalde}$ is TU.

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

S is TU if each collection of columns of **S** can be split into two parts so that the sum of the columns in one part minus the sum of columns in the other part is a vector with entries only 0, +1, and –1 (Schrijver, 1987).

Clearly any collection of two columns of **S** can be split into two such parts as each column contains only 0's or +1's. It only remains to show that a collection of all three columns can be split in this manner. Assign the first and second column to one part and the third column to the other part. The sum of columns in the first part is a vector with entries only +1 and +2. The sum of columns in the second part is a vector with every entry equal to +1. Therefore the sum of columns in one part minus the sum of columns in the other part is a vector with entries only 0, +1, and –1. **S** is TU and thus $\mathbf{B}_{Alcalde}$ is TU.

The path-stage matrix for Work Center A $\mathbf{B}_{WCA}$ is given by,

$$\mathbf{B}_{WCA} = \left[\begin{array}{c|cccccccccccc}
 & A & B & C & D & E & F & G & I & K & L & N & P \\
\hline
BP & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
ABP & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
ABFP & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
ACP & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
ACFP & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
ADFP & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
ADGP & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
ADIP & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
ADKP & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
ADLP & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
ADNP & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
ADP & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
AEP & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]$$

where I have added the stages and flowpaths for clarity. As discussed in Section 4.3.1, stages *I* and *N* refer to the aggregated stages *IJ* and *NO*. Column vectors are denoted by the associated stage letter, e.g. *A* or *P*.

Total unimodularity is preserved under both column permutation and the addition of a column with at most one nonzero, being +/-1 (Schrijver, 1987). Therefore $\mathbf{B}_{WCA}$ is TU if the following submatrix **S** of $\mathbf{B}_{WCA}$ is TU.

$$\mathbf{S} = \left[\begin{array}{cccccc} P & A & F & B & C & D \\ \hline 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{array}\right]$$

where the columns *G-N* have been removed and the column order has been rearranged. The matrix **S** has the following two properties.

[1] Any collection of columns from {P,A,F} can be split into two parts so that the sum of columns in the first part minus the sum of columns in the second part is a vector with entries only 0 or +1. To see this *P-A*, *P-F* and *P+F-A* are all vectors with entries only 0 or +1.

[2] The sum of any collection of columns from {B,C,D} is a vector containing only 0 or +1 entries.

From [1] and [2], any collection of columns of **S** can be split into two parts so that the sum of the columns in one part minus the sum of columns in the other part is a vector with entries only 0, +1, and –1. Therefore **S** is TU (Schrijver, 1987) and thus $\mathbf{B}_{WCA}$ is TU.

Lemma 3:

Any subset Λ that can be partitioned into two disjoint subsets Λ_m and Λ_n such that $P(\Lambda_m) \subseteq P(\Lambda_n)$, can be omitted from the set of subsets over which the maximum in Lemma 1 is evaluated.

Proof:

If Λ can be partitioned into two disjoint subsets Λ_m and Λ_n such that $P(\Lambda_m)\subseteq P(\Lambda_n)$, then $P(\Lambda)=P(\Lambda_n)$. This implies

$$\sum_{f\in P(\Lambda)} d_f - \sum_{k\in\Lambda} c_k = \sum_{f\in P(\Lambda_n)} d_f - \sum_{k\in\Lambda_m} c_k - \sum_{k\in\Lambda_n} c_k \leq \sum_{f\in P(\Lambda_n)} d_f - \sum_{k\in\Lambda_n} c_k$$

and therefore Λ cannot be the optimum in the above maximization.

Algorithm for generating the set *L* of possible stage combinations in the maximization of Lemma 1.

The set of stage combinations (or subsets) over which the maximum in Lemma 1 is evaluated does not contain every single possible combination of stages. Lemma 3 states that some stage combinations (or subsets) can be removed. An equivalent statement to Lemma 3 is given by,

If $P(\Lambda_m)\subseteq P(\Lambda_n)$, then $\Lambda=\Lambda_m\cup\Lambda_n$ can be omitted from the set of subsets over which the maximum in Lemma 1 is evaluated.

Denote the set of possible subsets by the set *L*. By using this version of Lemma 3, it is possible to compare two subsets to see whether the union of the two subsets can is contained in *L*. Remember a subset Λ corresponds to a combination of stages and $P(\Lambda)$ corresponds to the set of flow paths processed by any stage in Λ.

This comparison can be done using matrix algebra as follows. Let $P(\Lambda_m)$ be specified by a row vector

$$\mathbf{R^m} = \left[r_1^m, \cdots, r_F^m\right]$$

where the element r_f^m=1if $f\in P(\Lambda_m)$ and equals 0 otherwise for $f = 1,\ldots,F$. In other words, if a flow path is processed by some stage in Λ_m, then the corresponding flow path entry equals one. Let $\mathbf{R^n}$ be the flow path row vector for Λ_n. Let $\mathbf{R^{n-m}} = \mathbf{R^n} - \mathbf{R^m}$. If $\mathbf{R^{n-m}} \geq \mathbf{0}$, i.e. each element is non-negative, then $P(\Lambda_m)\subseteq P(\Lambda_n)$. Likewise if $\mathbf{R^{m-n}} \geq \mathbf{0}$, then $P(\Lambda_n)\subseteq P(\Lambda_n)$. An algorithm for comparing two subsets to see if the union of the subsets can be one of the subsets in L can be specified as follows.

ALGORITHM: COMPARE($\mathbf{R^n}, \mathbf{R^m}$)

$\mathbf{R^{n-m}} = \mathbf{R^n} - \mathbf{R^m}$

$\mathbf{R^{m-n}} = \mathbf{R^m} - \mathbf{R^n}$

IF $\mathbf{R^{n-m}} \geq \mathbf{0}$

RETURN FALSE

ELSE IF $\mathbf{R^{m-n}} \geq \mathbf{0}$

RETURN FALSE

ELSE

RETURN TRUE

END

The *COMPARE*($\mathbf{R^n}, \mathbf{R^m}$) algorithm determines whether the union of two subsets is contained in the set *L*. However, we still need to generate the set *L*. The algorithm for doing this is called *GENERATE L*.

This algorithm works as follows. The set *L* starts as an empty set. All subsets containing only one stage are then added to *L*. Next, the possible subsets containing two stages are added to L. Then the possible subsets with three stages are added, then four stages, etc. up until *K* stages. At this stage all possible subsets have been evaluated. If a set with *j*-1 stages, say Λ_{j-1}, is not in *L*, then no set that contains Λ_{j-1} will be in *L*. Therefore, when evaluating the subsets containing *j* stages, one only needs to consider the subsets that are formed by the union of a single stage and a subset containing *j*-1 stages that is already in *L*.

The algorithm is an iterative algorithm. Each iteration corresponds to the evaluation of the subsets containing *j* stages. At the start of the iteration, the subsets in *L* that contain *j*-1 stages will have been identified. Each of these "*j*-1" subsets will have a row vector corresponding to the flow paths processed by any stage in this subset (See above). These row vectors will form the matrix $\mathbf{M^{j-1}}$. The number of *j*-1 subsets will be specified by N^{j-1}. The n^{th} row vector of $\mathbf{M^{j-1}}$ will be identified by $\left[\mathbf{M^{j-1}}\right]_n$ .In the algorithm, the stage subset corresponding to the n^{th} row vector of $\mathbf{M^{j-1}}$ will be identified by *stage_subset* $\left[\mathbf{M^{j-1}}\right]_n$. In general, for any set Λ $\mathbf{R}\{\Lambda\}$ is the row vector corresponding to the flow path subset P(Λ).

ALGORITHM: GENERATE L

Initialization: Adding the single stage subsets to L

L=EMPTY SET

FOR k=1 TO K

    INCLUDE {k} in L

    ATTACH row vector $\mathbf{R}\{k\}$ to bottom of matrix $\mathbf{M^1}$

END k LOOP

Iteration: Generating the subsets with j stages that are in L, j=2,…,K

FOR j=2 TO K

    FOR k=1 TO K

        FOR n=1 TO N^{j-1}

            COMPARE $(\mathbf{R}\{k\}, [\mathbf{M^{j-1}}]_n)$

            IF RETURN = TRUE

            $\{newset\}=\{k\}\cup\{\, stage\_subset[\mathbf{M^{j-1}}]_n \}$

            INCLUDE {*newset*} in L

            ATTACH row vector $\mathbf{R}\{newset\}$ to bottom of matrix $\mathbf{M^1}$

        END n LOOP

    END k LOOP

END j LOOP

END

5.4 References

Agarwal, M., S. Nahmias, and A. Tsay, "Modeling Supply Chain Contracts: A Review," In Tayur, S., M. Magazine, and R. Ganeshan, eds., *Quantitative Models for Supply Chain Management* (Chapter 10). Kluwer Academic Publishers, 1998

Alcalde, J.L., "The Design and Implementation of a Synchronous Manufacturing System in a Job Shop Environment," LFM Thesis, MIT, 1997

Avriel, M, W.E. Diewert, S. Schaible and I. Zang, *Generalized Concavity*, Plenum Press, 1988

Banerjee, A., A Joint Economic-Lot-Size Model for Purchaser and Vendor, *Decision Sci.*, 17 (1986) 292-311

Birge, J.R. and F. Louveaux, *Introduction to Stochastic Programming*, Springer, 1997

Cachon, G.P., "Stock Wars: Inventory Competition in a Two Echelon Supply Chain," Working Paper, The Fuqua School of Business, Duke University, 1997

Cachon, G.P., "Competitive Supply Chain Management," In Tayur, S., M. Magazine, and R. Ganeshan, eds., *Quantitative Models for Supply Chain Management* (Chapter 5). Kluwer Academic Publishers, 1998

Cachon, G.P. and M.A. Lariviere, "Contracting to Assure Supply or What Did the Supplier Know and When Did He Know It?," Working Paper, The Fuqua School of Business, Duke University, 1997

Cachon, G.P. and M.A. Lariviere, "Contracting to Assure Supply: How to Share Demand Forecasts in a Supply Chain," Working Paper, The Fuqua School of Business, Duke University, 1999

Cachon, G.P. and P.H. Zipkin, "Competitive and Cooperative Inventory Policies in a Two-Stage Supply Chain," Working Paper, The Fuqua School of Business, Duke University, 1997

De Toni, A. and S. Tonchia, "Manufacturing Flexibility: A Literature Review," *Int. J. Prod. Res.*, 36, 6, (1998) 1587-1617

Eppen, G.D., R.K. Martin and L. Schrage, "A Scenario Approach to Capacity Planning," *Operations Research*, 37, 4 (1989) 517-527

Fine, C.H. and R.M. Freund, "Optimal Investment in Product Flexible Manufacturing Capacity," *Management Science*, 36, 4 (1990) 449-466

Gavish, R., Flexibility in Multi-Product Multi-Assembly Multi-Component Systems, Masters Thesis, MIT, 1994

Gupta D., Y. Gerchak and J.A. Buzacott, “The Optimal Mix of Flexible and Dedicated Manufacturing Capacities: Hedging against Demand Uncertainty,” *Int. J. Prod. Ec.*, 28, (1992) 309-319

Harrison, J.M. and J.A. Van Mieghem, “Multi-resource Investment Strategies: Operational Hedging under Demand Uncertainty,” *Eur. J. Oper. Res*., 113 (1999) 17-29

Hopp, W.J. and M.L. Spearman, *Factory Physic: Foundations of Manufacturing Management*, Irwin, 1996

Jeuland A.P. and S.M. Shugan, “Managing Channel Profits,” *Marketing Sci.*, 2, 3 (1983) 239-272

Jordan, W.C. and S.C. Graves, “Principles on the Benefits of Manufacturing Process Flexibility,” Management Science, 41, 4 (1995) 577-594

Lal, R. and R. Staelin, “An Approach to Developing an Optimal Discount Pricing Policy,” *Management Sci.*, 30, 12 (1984) 1524-1539

Lariviere, M.A., “Supply Chain Contracting and Coordination with Stochastic Demand,” In Tayur, S., M. Magazine, and R. Ganeshan, eds., *Quantitative Models for Supply Chain Management* (Chapter 8). Kluwer Academic Publishers, 1998

Law, A.M. and W.D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill, 1991

Lee, H.L. and M.J. Rosenblatt, “A Generalized Quantity Discount Pricing Model to Increase Supplier’s Profits,” *Management Sci.*, 32, 9 (1986) 1177-1185

Li, S. and D. Tirupati, “Dynamic Capacity Expansion Problem With Multiple Products: Technology Selection and Timing of Capacity Additions,” *Operations Research*, 42, 5 (1994) 958-976

Li, S. and D. Tirupati, “Impact of Product Mix Flexibility and Allocation Policies on Technology,” *Computers Ops. Res.*, 24, 7 (1997) 611-626

Li, S. and D. Tirupati, “Technology Choice with Stochastic Demands and Dynamic Capacity Allocation: A Two Product Analysis,” *Journal of Operations Management*, 12 (1995) 239-258

Monahan, J.P., “A Quantity Discount Pricing Model to Increase Vendor Profits,” *Management. Sci.*, 30, 6 (1984) 720-726

Moorthy, K.S., “Managing Channel Profits: A Comment,” *Marketing Sci.*, 6, 4 (1987) 375-379

Nahmias, S., *Production and Operations Analysis*, 3rd Edition, Irwin, 1997

Nemhauser, G.L. and L.A. Wolsey, *Integer and Combinatorial Optimization*, John Wiley & Sons, 1988

Parlar, M. and Z.K. Weng, “Designing a Firm’s Coordinated Manufacturing and Supply Decisions with Short Product Life Cycles,” *Management Sci.*, 43, 10 (1997) 1329-1344

Pasternack, B.A., “Optimal Pricing and returns Policies for Perishable Commodities,” *Marketing Science*, 4, (1985) 166-176

Schrijver, A., *Theory of Linear and Integer Programming*, John Wiley & Sons, 1987

Spengler, J.J., “Vertical Restraints and Antitrust Policy,” *Journal of Political Economy*, 58 (1950) 347-352

Van Mieghem, J.A., “Investment Strategies for Flexible Resources,” Management Science, 44, 8 (1998) 1071-1078

Weng, Z.K., “Channel Coordination and Quantity Discounts,” *Management Sci.*, 41, 9 (1995) 1509-1522